Educational Administration: Theory and Practice
2024, 30(6)(s), 278 - 286
ISSN: 2148-2403
https://kuey.net/
Research Article

Unlocking The Potential Of Machine Learning For Diabetes Prediction

Mr Nisarg Kishorchandra Atkotiya[1*], Dr Ramani Jaydeep Ramniklal[2], Dr Jayesh N Zalavadia[3]

[1*]Department of Statistics, Saurashtra University, Rajkot, INDIA Email: nisargatkotiya@gmail.com
[2]CS & IT Department, Atmiya University, Rajkot, INDIA, Email: jaydeep.ramani@atmiyauni.ac.in
[3]Department of Comm. & Mngt. Atmiya University, Rajkot, INDIA. Email: jayesh.zalavadia@atmiyauni.ac.in

Citation: Mr Nisarg Kishorchandra Atkotiya , et.al (2024) Unlocking The Potential Of Machine Learning For Diabetes Prediction, Educational *Administration: Theory And Practice, 30(6)(s), 278 - 286*
Doi: 10.53555/kuey.v30i6(S).5372

ARTICLE INFO | **ABSTRACT**

Millions of individuals throughout the world suffer with diabetes, a chronic condition that if unchecked can have catastrophic health repercussions. In order to forecast diabetes risk and aid healthcare professionals in managing or preventing the condition, machine learning algorithms have become increasingly effective. The goal of our work is to inspect the achievement of machine learning techniques in predicting diabetes. The dataset used in previous study consists of demographic and clinical data of patients who have been diagnosed with diabetes and those who have not. Different classification and Neural Network algorithms, such logistic regression, Artificial Neural Network, XGBoost Random Forest, Voting Classifier and Naïve bays were employed to forecast the occurrence of diabetic in patients. The findings of the study indicate that these machine learning algorithms achieved significant accuracy rates in diabetes prediction. Among the algorithms utilized, the Random Forest algorithm achieved the best accuracy rate of 86.5The study also discovered that a range of parameters, such as hypertension, age, body weight, and levels of glucose, were valid markers of diabetes. For individuals who have a greater chance of acquiring diabetes, these factors can help medical experts act early and provide unique treatment strategies.

Keywords—Machine Learning, Classification Algorithms, Prediction, Accuracy, Precision, Random Forest, Naive Bayes Decision Tree.

I. INTRODUCTION

The word "diabetes" is now commonly known and presents serious difficulties for both industrialized and developing nations. When we eat, the pancreas releases the chemical insulin, which allows glucose to enter the bloodstream. Diabetes is a medical condition characterized by insufficient production of insulin in the pancreas, leading to various complications. Diabetes is a condition characterized by an inability of the pancreas to produce insulin, which results in a variety of complications. These complications include obviousness, kidney damage, pancreatic beta cell degeneration, heart and cerebrovascular diseases, peripheral venous disease, sex issues, joint pain, weight loss, ulcers, and a compromised immune system. High glucose levels are one of the initial signs of type 2 diabetes, among others including incessant urination, extreme thirst, increased hunger, and weight loss. Diabetes can develop as a result of low insulin release or a body that does not respond to insulin effectively, resulting in elevated blood glucose levels. [1].

It is impossible to fully recover from a prolonged disease, necessitating constant managerial oversight. The prevention of illness, however, may be aided by early detection and forecasting. The non-straightness and uniqueness of most boundaries can make it challenging to deconstruct diabetic information and exactly estimate its consequences for individuals. Numerous research projects have focused on developing diabetes expectation models using various AI techniques. One such evaluation suggested using the PIMA dataset with an artificial neural network (ANN) for the early detection of diabetes. [2].

The primary objective of treating diabetes is to manage blood sugar levels effectively and mitigate the risks of associated complications. Treatment options encompass lifestyle adjustments, including maintaining a healthy weight, involved in regular physical activity, and following a nutritious diet. Additionally, healthcare professionals may prescribe medications such as insulin, oral hypoglycemic agents, or other glucose-lowering drugs. In certain situations, surgical interventions such as bariatric surgery might be suggested as part of the treatment plan.

The management of diabetes requires a multidisciplinary approach involving healthcare providers such as physicians, nurses, dietitians, and pharmacists. Diabetes self-management education and support programs are also essential in helping individuals with diabetes to manage their condition effectively and improve their quality of life Machine learning algorithms have recently shown promise in detecting diabetes risk and assisting in the treatment of the condition. These algorithms can evaluate huge volumes of clinical and demographic data to find trends and predict the likelihood of developing diabetes. By identifying high-risk patients early, healthcare providers can develop personalized treatment plans and interventions to prevent or manage diabetes.

The capacity of machine learning algorithms to effectively and quickly assess massive datasets is one of the key benefits of utilizing them to forecast diabetes. These algorithms can spot links and trends in the data that human analyst would miss, enabling more accurate forecasts and actions.

Diabetes is a complex and growing public health concern that has significant health and economic consequences. Machine learning algorithms have emerged as a promising tool for predicting diabetes risk and aiding in the management of the disease.

A. Types and Importance of Machine Learning

The goal of machine learning, a subfield of artificial intelligence, is to develop models and methods that allow computers to analyses past information, forecast future events, and even make choices without explicit programming. There are many different machine learning models in it, and each one has unique properties and uses.

ML algorithms may be divided into three main groups: supervised learning, unsupervised learning, transfer learning, and reinforcement learning. Below, we briefly touch on these.

B. Supervised Learning

In supervised learning, the computer is trained on data that has already been labelled, and it makes predictions about the future based on that data [3]. Regression and classification are the two other subtypes of supervised learning.

C. Regression

This kind of supervised learning is used when the projected outcome is a continuous variable and there is a relationship between the input and output variables [4]. It may be found in many different forms, such as polynomial, linear, and non-linear regression as well as regression trees.

D. Classification

This type of Supervised Learning is used When there is a
link between the input and output characteristics and the predicted outcome is a categorical variable, such as TrueFalse, Male-Female, and Apple-Banana [5]. Some examples of Classification algorithms are SVM, K-NN, NB, RF and XGBoost.

E. Unsupervised Learning

In the case of unsupervised learning, the data that we provide the computer for training is unlabeled, and the computer uses machine learning (ML) techniques to figure out the underlying pattern and predict the outcome [6]. Unsupervised algorithms come in the form of association and clustering. Unsupervised learning, as the name implies, relies on algorithms that are capable of self-learning. These algorithms are able to learn without the need for labels or prior training. On the contrary, the model receives unprocessed, unidentified data and is required to deduce its own principles and organize the information by identifying resemblances, distinctions, and patterns, without any specific guidance on how to handle each individual data point.

F. Clustering

Clustering is a kind of supervised machine learning that divides unlabeled data into groups [7]. The method divides the data into two groups: one for similar data and the other for less comparable data, hence the name "clustering."

G. Reinforcement Learining

It is a form of the machine learning algorithm in which an agent requires a sequence of decisions to achieve a final goal [8]. Artificial Agent interacts with the environment and for every correct and the bad action it does, it will earn either reward or punishment dependent on the action it does. The main purpose is to improve the chances of earning the prize.

II. LITERATURE SURVEY

Govindraj. M et al. [9]. In their work used random forest, neural network and decision tree for prediction of diabetes mellitus (DM). Results of medical tests for their dataset was taken from a hospital in Luzhou, China. In their study, the models were assessed using fivefold crossvalidation. Several of the best approaches were selected for this study based on their adaptability to independent test trials. 68994 individuals with diabetes and 68994 persons in good health made up the training set of 68994 data points. In addition, the facts considered are practically randomly distributed throughout 5 times. The outcome is calculated as the mean of these five tests. PCA and mRMR were employed in this work to minimize the dimensionality. When all the factors are considered, the RF forecast's accuracy rises to 80.84%.
Using a unique genetic method based on stacking generalization, Dr. Sandeep Kumar Hegde et al. [10] were able to predict chronic illnesses, such female breast cancer, cardiovascular disease, diabetes, and cancer of the lungs at an early stage. The researchers used the UCI sickness dataset for their study. A symmetrized feature choice approach was first used to extract the top features from the chronic disease dataset. The suggested approach used a variety of computations for stacking speculation, including SVM, Random Woodland, Decision Tree, Naive Bayes Classifier, and Multilayer Perceptron. The prediction ability of the model was increased even more using a genetic algorithm. According to the findings, the recommended stacking generalization model and genetic algorithm had an accuracy of 90.21% for diabetes, 91.25% for breast cancer, 92.50% for lung cancer, and 93.25% for other diseases.
Bade Kranthi Priya et al. [11], in their study, proposed a robust framework for diabetes prediction that includes feature selection, outlier rejection, handling missing values, data normalization, and several ensemble strategies. It can produce superior results on their dataset, and they found that the XGBoost classifier achieved an outstanding accuracy of 81% as compared to other classifiers.
S. Nava Bharath Reddy et al.'s study [12] focused on applying the “extreme learning machine” (DP-ELM) model to investigate diabetes prediction. They applied pre-processing techniques to solve the issue of missing data in the Pima Indian diabetes dataset. The pre-processed dataset was used to extract statistical features using principal component analysis (PCA). The ELM model was told to employ PCA features to generate the training feature dataset. Through the use of a random test combination, an ELM test was carried out to categorise both positive as well as negative stages of diabetes. The simulations showed that the suggested DP-ELM model worked better than prior methods.
N. Satheesh Kumar and Dr. N. Sathyanarayana [13]. The purpose of their work is to build a framework for making decisions that will allow doctors to recognize diabetic patients who are at risk of an early readmission. To do this, numerous data analysis techniques have been used. An innovative model is developed in this work using computer vision. Prioritising those who are most likely to experience difficulties and require readmission early on lowers healthcare expenses and enhances the hospital's reputation, improving the quality of the medical service and reducing expenditures. It is more accurate to make predictions using machine learning than it is to make predictions using more conventional techniques. The study focused on predicting the likelihood of patient readmission to the hospital. Various classifiers, including a standard scaler, Random Forest and decision tree, CATboost for categorical features, and XGBoost, were employed. However, the deep learning approach utilizing a machine learning methodology outperformed the other models when evaluated with realworld data. The analysis was enhanced, and a more practical framework was developed, addressing different modules such as feature extraction. The dataset revealed that more than 98% of patients did not receive medications during their hospital stay. This observation may be attributed to the limited information available on medication administration in the hospital, as most prescriptions are typically meant for home use.
CS Manikandababu et.al [14] This study aims to create a machine learning algorithm that can identify type 2 diabetes in humans. The PIMA Indian Diabetic Dataset is submitted to the logistic regression classification approach with a total of 8 parameters. The PIDD dataset contains 768 entries, of which 268 are of diabetics and 500 are not. In this case, the stacking strategy is used to create a new model that outperforms and is more precise than the basic classifier alone. The model is constructed using machine learning techniques such as NB, SVM and DT. For easier comparisons, efficiency and timing of execution for each strategy are recorded. With a 94% accuracy rate, the proposed approach was used in the Python IDE.
Rashiduzzaman Shakil et.al [15], used an updated Convolution Neural Network (CNN) model on 130 USA Hospitals diabetes dataset, various hyper-parameters and layer topologies. They employed ADAMAX, Nesterov-accelerated adaptive moment (NADAM), adaptive moment estimation (ADAM), stochastic gradient descent (SGD), and RMSprop root mean square propagation (RMSprop), which are five different types of optimizers. The ADAMAX optimizer was specifically used to achieve an accuracy of 99.98%.
In their work, Saravanan Alagarsamy et al. [16] used a combination of unsupervised and supervised AI models to give patient-explicit ongoing estimates and screen sickness forecasts. They collected the PIDD dataset from the American National Institute of Diabetes from 1965–1976. Numerous machine learning algorithms were tried in this study, but Naive Bayes outperformed them all with an accuracy rate of 94%.
MALINI M. et. al. [17], used different machine learning algorithm to predict an early stage of Diabetes. In this work, the data from patient medical records is combined with the various machine learning techniques, and the accuracy of these techniques is assessed through performance analysis. Using techniques like random forests, decision trees, and Knearest Neighbour classifiers, it is feasible to assess either a patient has diabetes

or not. The use of “machine” learning approaches to help doctors identify diabetic patients in their early stages. In order to compare classifier models, the accuracy value was used. In these, we employ performance evaluation criteria, including precision, recall, and F1-Score.In which the XGB model classifier outperforms other classifier models by 82% in terms of accuracy.
P. T. Siva Gurunathan et. al. [18] contrasting the KNN with Random Forest Classifier, for the anticipation of diabetes, they prepared an online application using AI. PIMA India diabetes dataset was taken from Kaggle. Compared to KNearest Neighbour (KNN), that only forecasts with an accuracy of 75% on the cleaned data, Random Forest (RF) has a prediction accuracy of 83%
In a study on how to enhance type 1 diabetics' selfmanagement behaviours, Zhu et al. [19] explored the use of wearables and deep learning. Six participants wore wearable devices that the researchers used to collect physiological data from; they then analysed the data using deep learning techniques to produce individualized alerts and suggestions. The outcomes demonstrated the system's efficacy in making prompt and precise recommendations that could improve type 1 diabetes self-management practises.
Atef Hadi Ataya [20] used six classification algorithms, namely, SVM, RF, XGBoost, LR, K-NN, and LightGBM for prediction of diabetes in the early stage. The PIMA India Diabetes Datasets (PIDD) was taken from the Kaggle repository. On comparison of these algorithms LightGBM algorithm achieved 85.5% accuracy which is better than other algorithms.
The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

III. PROPOSED ARCHITECURE

The prediction of diabetes is done using an architecture for machine learning that performs a number of tasks. Firstly, we collected the PIDD dataset from Kaggle. Pre-processing is done on the collected data to remove useless features, and then feature extraction is applied to important features like blood sugar, glucose level, etc. In the end, we employed machine learning techniques such as DT, NB, RF, and SVM to forecast the disease.
The machine learning platform employed four different kernels, namely linear, polynomial, RBF, and sigmoid, in order to forecast diabetes. The authors achieved varied accuracies in various kernels, which fell in the range of 0.69 to 0.82. The radial basis kernel function combined with SVM technique achieved the highest accuracy of 0.82.
A machine learning approach and the Pima Indian dataset are utilized to develop a prediction system capable of detecting various diseases, including diabetes. As per the authors, the Naive Bayes method outperformed the random forest approach by achieving a higher accuracy, with improvements of 0.43%.

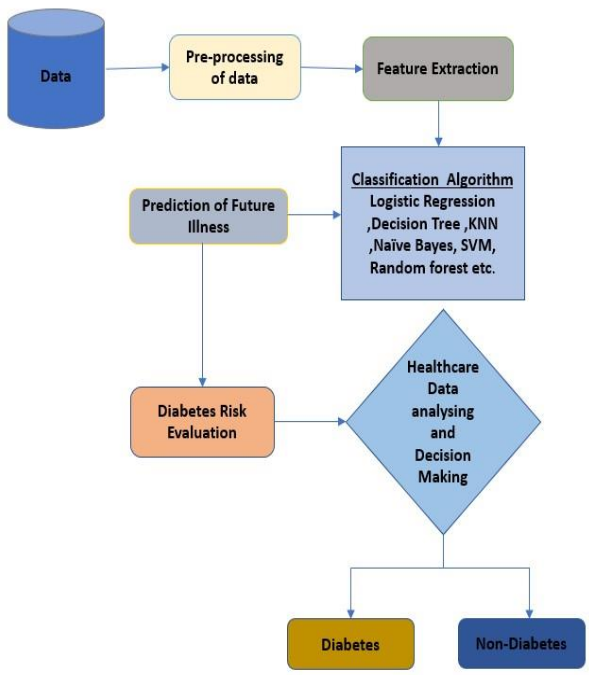

Fig 1. Proposed framework for Diabetes prediction

Author/Year	Algorithm Used	Dataset	Optimal Result
C. S. Manikandababu et. al. (2022)	NB, SVM, DT, LR, and Ensemble.	PIMA India Diabetic Dataset (PIDD)	Accuarcy- 94% by Ensemble
N. Satheesh Kumar et.al. (2023)	Random Forest (RF), DT, Gradient Boost (XGB) and Support Vector Machine – Radial Basis function Kernel (SVM-RBF)	130 Hospital of USA	Accuracy- 89% by SVM-RBF
Bade Kranthi Priya et. Al. (2023)	DT Classifier, RF Classifier and XGB classifier	John Diabetes Dataset	Accuracy81% by XGB
S. Nava Bharath Reddy et. al. (2023)	Linear Support Vector Machine, NB, RF, LR, K-Nearest Neighbour (KNN), Distributed and Parallel Extreme Learning Machine (DP-ELM)	PIDD	Accuracy- 92% by DPELM.
Govind raj M .et. al. (2023)	KNN, RF, XGB, LR and DT	Kaggle Diabetes Dataset	Accuracy- 85% by RF
Saravanan Alagarsamy et. al. (2023)	NB, KNN, DT and SVM	PIDD	Accuracy- 94% by NB
Nicole D'Souza et. al. (2022)	KNN, LR, DT, NB, RF, XGB	PIDD	Accuracy- 98% by RF
MALINI M. et. al. (2023)	RF, DT, LR SVM, KNN and XGB	PIDD	Accuracy82% by XGB

TABLE 1: A comparative study of various papers on diabetes prediction using ML model.

TABLE 2: The Result comparison of different studies on diabetes Dataset.

Study	Algo rithm	Performance Matric	Preci sion	Re-call	F1- scor e	Accuracy
Govind raj Met. al. [3]	Random Forest	Accuracy	90%	_	85 %	85%
Bade Kranthi Priya et. Al. [5]	XGBoos t	Accuracy	_	_	_	81%
S.Nava Bharath Reddye t. al. [6]	DP-ELM	Accuracy	_	_	_	92%
N.Satheesh Kumar et.al [7]	SVMRBF	Accuracy	91%	80 %	82 %	89%
Sarava- nan Alagar samy et. al. [10]	Naïve Bayes	Accuracy	82%	81 %	89 %	94%
MAL	XGBoos	Accuracy				

INI M. et. (2023) [11]	t		85%	88 %	86 %	82%
Nicole D'Souz a et. [13]	Random Forest	Accuracy	97%	98 %	_	98%

IV. RESULT AND ANALYSIS

This evaluation covered a total of 15 studies and was focused on the use of predictive approaches for identifying individuals who are at risk of developing diabetes. The population investigated, the outcome measures considered, and the predictive models employed varied throughout these investigations. The review's conclusions show that predictive models are useful for spotting those who are at risk of developing diabetes. Out of 15, 12 of the studies successfully used their prediction models to differentiate between those who were at elevated risk and those who were not at danger, exhibiting good or excellent discrimination. The majority of research examined in the study combined clinical, demographic, and perhaps genetic data to create their predictive models.

These models frequently considered variables like age, place of residence, BMI, gender, and smoking status. In order to improve the precision of their forecasts, some studies additionally included further information, such as blood tests or genetic markers. In Table 2, which comprised multiple performance measures, the findings of various research based on the same dataset, such as the Healthcare Kaggle dataset on diabetes, were summarized. The Comparative study of these studies, depicted in Figures 2, 3, 4, and 5, revealed variations in the findings of various studies and the findings that followed their application. Analysis of this study indicates important classification metrics, including accuracy, precision, recall, and F1-Score, that might influence how well diabetes risk prediction models perform.

More predictors in a model often led to greater accuracy than fewer predictors in a model. In addition, machine learning-based models, such as boosting or random forest, typically outperformed models created using conventional statistical techniques. The results of this review are consistent with those of other reviews and meta-analyses, however, there are some restrictions to consider. Some of the studies had a high risk of bias and varied in terms of quality. More study is required to understand the long-term efficacy of predictive models for the prevention of diabetes, as the majority of studies concentrated on short-term results.

Optimum ways to incorporate these indicators into clinical practice as well as the combination of predictors that works best are still to be determined through additional research. We came across some remarkable instances of zero values. For instance, it is not possible for skin thickness and Body Mass Index (BMI) to be zero. The mean value has been substituted in place of the zero value. The holdout validation technique was employed to split the dataset into training and test sets, with 80% allocated as training data and 20% as test data.

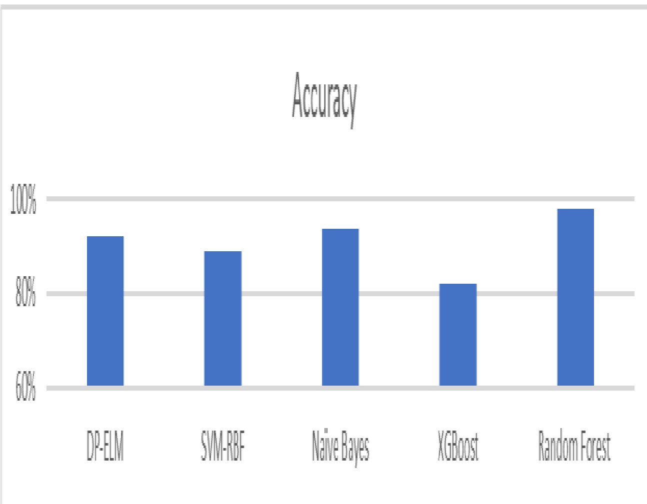

Figure 2. Comparison of Different Algorithm based on Accuracy.

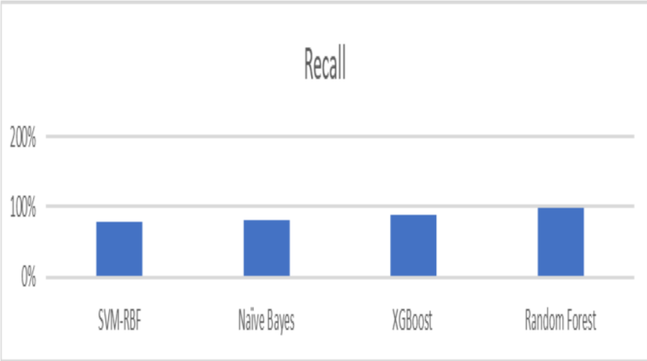

Figure 3. Comparison of Different Algorithm based on Recall.

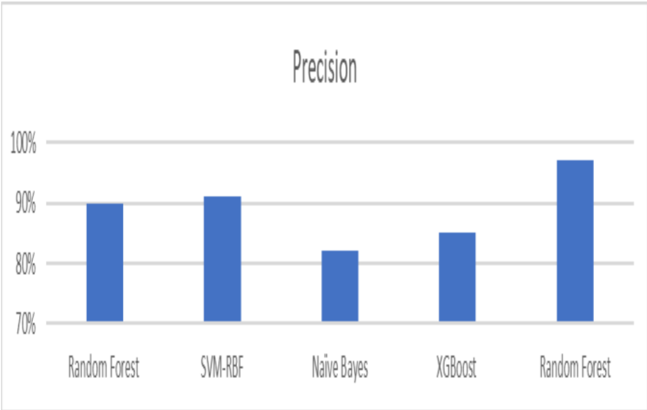

Figure 4. Comparison of Different Algorithm based on Precision.

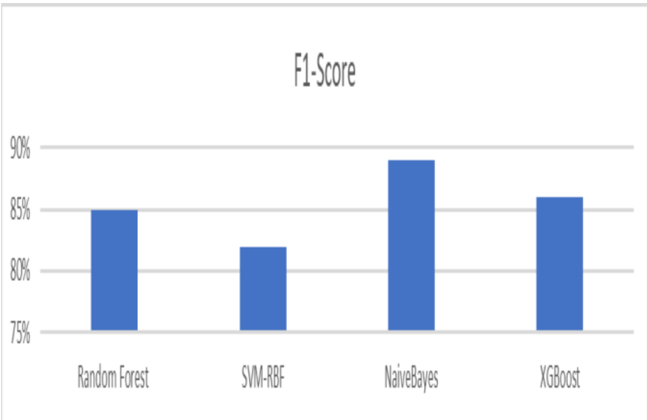

Figure 5. Comparison of Different Algorithm based on F1-Score.

V. CONCLUSION

Early detection of diabetes can help high-risk individuals make lifestyle changes, reducing complications and potentially saving lives. Diabetes is a significant health concern that requires prompt intervention to prevent further complications. Diabetes is a serious health concern that requires immediate attention to prevent additional miscommunication. As part of this study, we have conducted a detailed literature analysis on the different methods of diagnosing diabetes. As a result of our investigation into the use of AI, the top models for forecasting and precision that we found were neural networks, XGBoost, decision trees, random forests, and K-Nearest Neighbor. In view of the findings of this study, our goal is to develop an AI model for the early detection of diabetes, focusing on achieving high accuracy while minimizing cost and complexity. This experiment may mark a significant advance in medicine by enabling practical treatments and diabetes prevention strategies for those who are at risk. By utilizing AI techniques, we want to improve the effectiveness and viability of the diabetic location, resulting in improved health outcomes for those who are affected.

REFERENCES

[1] D'Souza, Nicole, Kunjal Shah, and Pranav Singh. "Diabetes Detection Using Machine Learning Algorithms". In 2022 IEEE Bombay Section Signature Conference (IBSSC) 2022 Dec 8 (pp. 1-5), IEEE.

[2] Fakih, A. H., Venkatesh, A. N., Vani, V., Naved, M., Kshirsagar, P. R., & Vijayakumar, P. (2022, May). “An efficient prediction of diabetes using artificial neural networks”. In AIP Conference Proceedings (Vol. 2393, No. 1). AIP Publishing.
[3] T. Jiang, J. L. Gradus, and A. J. Rosellini, “Supervised machine learning: A brief primer,” Behavior Therapy, vol. 51, no. 5, pp. 675–687, May 2020
[4] Chauhan, Tannu, Surbhi Rawat, Samrath Malik, and Pushpa Singh. "Supervised and unsupervised machine learning based review on diabetes care." In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), vol. 1, pp. 581585. IEEE, 2021.
[5] Abdulhadi Nour, and Amjed Al-Mousa. "Diabetes detection using machine learning classification methods." In 2021 International Conference on Information Technology (ICIT), 2021 Jul 14 (pp. 350354). IEEE.
[6] S. Mushtaq and K. S. Saini, "A Review on Predicting Brain Stroke using Machine Learning," 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, pp. 667-673, May 2023.
[7] Andrabi, Syed Abdul Basit, and Inderjeet Singh. "A Comparative Study of Machine Learning Techniques for Diabetes Prediction." In 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA),2022 Sep 21 (pp. 741-745). IEEE.
[8] Kumar JS, Kumar KH, Haleem A, Sivaranjani B, Kiran BN, Prameela S. "IBM auto AI bot: Diabetes mellitus prediction using machine learning algorithms". In 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC) 2022 May 9 (pp. 24-29) IEEE.
[9] Govindraj, M., Asha, V., Prasad, A., Kumar, S., & Rakshitha, C. . “Comparison of Machine Learning Techniques for Prediction of Diabetes”. In 2023 7th International Conference on Computing Methodologies and Communication (ICCMC) 2023 Feb 23 (pp. 73-80). IEEE.
[10] Hegde Sandeep Kumar, Rajalaxmi Hegde, Veeranna Hombalimath, D. Palanikkumar, Neha Patwari, and V. Dankan Gowda. "Symmetrized Feature Selection with Stacked Generalization based Machine Learning Algorithm for the Early Diagnosis of Chronic Diseases." In 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), pp. 838-844. IEEE, 2023.
[11] Priya, Bade Kranthi, VS Amar Koushik Tanniru, and Madhavi Katamaneni. "Ensemble Learning Model for Diabetes Prediction." In 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), pp. 33-36. IEEE, 2023.
[12] Reddy, S. Nava Bharath, KV Narasimha Reddy, SN Tirumala Rao, and K. S. M. V. Kumar. "Diabetes Prediction using Extreme Learning Machine: Application of Health Systems." In 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), pp. 993-998. IEEE, 2023.
[13] Kumar, N. Satheesh, and N. Sathyanarayana. "Prediction of Diabetic Patients with High Risk of Readmission using Smart Decision Support Framework." In 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), pp. 1664-1671. IEEE, 2023.
[14] Manikandababu, C. S., IndhuLekha, S., Jeniefer, J., & Theodora, T. A. (2022, October). Prediction of Diabetes using Machine Learning. In 2022 International Conference on Edge Computing and Applications (ICECAA) 2022 Oct 13 (pp. 1121-1127). IEEE.
[15] Shakil R., Akter, B., Faisal, F., Chowdhury, T. R., Roy, T., & Khater, A. (2022, March). A promising prediction of diabetes
[16] using a deep learning approach. In 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), 2022 Mar 29 (pp. 923-927). IEEE.
[17] Alagarsamy, S., Muthuvel, P., Kumar, K.S., Reddy, M.M., Naidu, M.M. and Kumar, M.V., 2023, March. “System for Predicting the Diabetes using Machine Learning Techniques”. In 2023 Second International Conference on Electronics and Renewable Systems (ICEARS) (pp. 1333-1338). IEEE
[18] Malini, M., Gopalakrishnan, B. and Naveena, S., 2022, May 9. "Prediction and Detection of Diabetics Mellitus using Different Machine Learning Approaches". In 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC) (pp. 323-330). IEEE.
[19] Gurunathan PS, Raghuraj S, Roahit S, Nithishkumar S. “Web Application-based Diabetes Prediction using Machine Learning”. In 2023 7th International Conference on Computing Methodologies and Communication (ICCMC) 2023 Feb 23 (pp. 296-302). IEEE.
[20] Zhu, Taiyu, Chukwuma Uduku, Kezhi Li, Pau Herrero, Nick Oliver, and Pantelis Georgiou. "Enhancing self-management in type 1 diabetes with wearables and deep learning." npj Digital Medicine 5, no. 1 (2022): 78.
[21] Ataya, Atef Hadi. "Early detection of Diabetes using Machine Learning Techniques." In 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS),(ICAIS) 2023 Feb 2 (pp. 886-891). IEEE
[22] Gayatri, Gottumukkala, Aitha Sai Varun, Manne Jaswanth Sail, Ch Pratyusha Chowdari, and K. Jamal. "Diabetic Foot Ulcer Identification using Machine Learning." In 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), 2023 Mar 23 (pp. 309-315). IEEE.
[23] V. Rajyaguru, C. Vithalani, and R. Thanki, “A literature review: Various learning techniques and its applications for eye disease identification using retinal images,” International Journal of Information Technology, vol. 14, no. 2, pp. 713–724, 2020.

[24] M. A. Ramessur and S. D. Nagowah, "A predictive model to estimate effort in a sprint using machine learning techniques," International Journal of Information Technology, vol. 13, no. 3, pp. 1101–1110, 2021.
[25] J. N. Heo, J. G. Yoon, H. Park, Y. D. Kim, H. S. Nam, and J. H. Heo, "Machine learning–based model for prediction of outcomes in acute stroke," Stroke, vol. 50, no. 5, pp. 1263–1265, May 2019.
[26] R. S. Jeena and S. Kumar, "Stroke prediction using SVM," 2016 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), pp. 600–602, Dec. 2016.
[27] F. S. Alotaibi, "Implementation of machine learning model to predict heart failure disease," International Journal of Advanced s2019.
[28] G. Sailasya and G. L. Kumari, "Analyzing the performance of stroke prediction using ML classification algorithms," International Journal of Advanced Computer Science and Applications, vol. 12, no. 6, pp. 53