

Soil Moisture Prediction using Deep Neural Network Approach.

Hiren J. Modha¹ , Ashish M. Kothari² ¹Research Scholar, Electronics & Communication, Atmiya University, Rajkot, Gujarat- India Email: hiren_modha2001@yahoo.com ²Associate Professor, Electronics & Communication, Atmiya University, Rajkot, Gujarat- India Email: amkothari.ec@gmail.com

Abstract

Soil moisture content is the most significant element in fit farming output and circulation of water, and its accurate forecast is critical regarding water resource management. Mostly soil moisture is complicated through structural features in addition to climatic complications. It's tough to come up with an optimal mathematical model for soil moisture since there are so many variables for calculation. Presented forecasting models have issues with prediction accuracy, generality, plus other factors like Prediction performance, as well as multi-feature processing capabilities etc. considering all these factors taking Gallipoli, Turkey as a reference site for developing a deep neural network model to forecast moisture with good accuracy and minimum error. The dataset contains entities since 2008 to 2021. Doing quite a bit of mathematical analysis and establishing the correlation between selected features with the spearman coefficient, the appropriate weather data is able to give proper weight to forecast soil moisture. The output of the proposed method proves that the deep learning approach is realistic as well as efficient for the prediction of moisture. Also, deep learning technique is able to make model generalizations with excellent accuracy and minimum errors which is used to save irrigation water with controlling drought.

Keywords: Deep Neural Network, Machine Learning, Multilayer Layer Perceptron (MLP), Support Vector Machine (SVM), Rectified Linear Activation Function (ReLU).

Introduction

*e***ISSN**1303-5150 www.neuroquantology.com Water is the most important asset for the existence as well as the growth of all living things on the planet. Soil moisture is not only necessary for plant growth, but it is also an important thing to water a farm with a system that has a correlation among soil-plant-atmosphere systems [1-3]. Groundwater resources, on the other hand, decrease in water quality when human activities increase and the quantity of excavation is greatly surpassed. Groundwater levels continue to fall, causing a drop-in water level on farm and a reduction in the

soil's ability to store the water [4-5]. The absence of precipitation, especially in arid locations, creates drought on farm and there is no refilling of water at a particular time, which has a detrimental impact on crop development [6]. It is very critical in this scenario to create a suitable irrigation system at the proper time [7, 12]. Water consumption and crop growth are directly influenced by the increase and regression of soil moisture. Dryness of farmland, flood management plus the determination of proper watering of the farm is all key indicators in agricultural production. To

appropriately utilize water assets regarding farming, and boost harvesting, a precise forecast of soil water regression normal decision is critical.

To build prediction models, most current soil moisture prediction approaches rely on empirical formulae, linear regression, and neural networks. The earliest model is the empirical formula model. Shu Sufang constructed a linear regression model to forecast soil moisture. In this research, the model is able to display the trends, based on a grey correlation examination of climate statistics [13]. When the dataset is unstructured the linear regression provides a huge amount of errors and poor accuracy compared to other statistics approach. Thus, it is very difficult to achieve required prediction. Research scientist had started to utilize the different approach with neural network as it has improved training algorithms in forecasting. J.W. Hummel used a nearinfrared reflection sensor to gather soil moisture data, which was then analyzed using multiple linear regressions, providing a predicted standard deviation of 5.31 percent [14]. Hou Xiaoli et al. utilized the ANN algorithm to forecast soil moisture index at dissimilar depths in addition to many input climate entities [15]. Here the outcomes were best matched by means of actual entities. With the foundation of soil moisture data properties, Li Ning enhanced the neural network optimization technique. Because the network's starting parameters are randomly set, the BP method has a poor training pace and is prone to local optima [16].Ji Ronghua improved the function of neural network activation. The traditional activation function was dropped in favor of a new complex number domain, and the network was trained using the MLP approach [9].When compared to a typical back-propagation (BP) neural network, the forecast truthfulness improved by 9.1%, offering an additional precise notional foundation forsoil moisture forecasting. M. Kashif Gill used an SVM technique to forecast soil moisture along with raised accuracy to 89 percent, avoiding the curse of dimensionality difficulty in neural networks [17]. As a result, the evolutionary technique was applied to determine the large-scale ideal starting parameters prior to training, greatly speeds up the training and improves the model's prediction accuracy.However, because soil moisture is affected by structural and meteorological factors, it is difficult to construct an optimal mathematical model for predicting soil moisture.The architectural properties and methods of classic neural networks are inadequate for processing large amounts of information, prediction accuracy is not easy to increase additionally, and simplification with scalability is restricted. Chen Xiaofeng proposes a technique for forecasting soil moisture, precipitation, and drought focused on the multidimensional linear correlation of soil moisture [18]. This approach forecasts drought for the next several days by assessing the preliminary moisture proportion, precipitation per day, mean temperature, and saturation change, as well as the multivariate linear connection of soil moisture. Jackson employs an empirical method and a time domain reflectometry device to measure the soil moisture flux in his model, which presents recommendations for drought-resistant irrigation systems (TDR) [8]. Although the outcomes are thesame, the method is less complicated.Although the empirical method is straightforward and straightforward, the model parameters are very regionally dependent and must be recalculated when transplanting to other locations, which is both lengthy and wasteful. Various prediction models have arisen as a result of the fast advancement of computer technology. Hinton created Deep Learning (DL) in 2006, which uses a multiple hidden layer to improve the categorization and fitting skills of large amounts of data with numerous features [19].It has been effectively employed in image recognition, search engines, stock price forecasts, and other domains, and has shown great computational power when compared to typical neural networks.Because of the nonlinear and exceedingly complex character of farms, few researchers have recently used Deep Learning to improve prediction accuracy in soil physical structure analysis [10]. The prediction of the moisture with use of other parameters requires high accuracy in measurement of relative parameters.The suggested research study aims to create a deep neural network model for predicting soil moisture using a meteorological dataset given by the Gallipoli-Turkey metrological department. For generalizations of model, a dataset with fine accuracy is generated with use ofa raspberry pi board and few sensors, which contain different features. Then build a model based on the same algorithm, with its robust data analysisabilities to attain excellent accuracy of soil moisture forecastingfor the self-generated dataset.Chantal Saad Hajjar et al. has proposed a different approach based on image processing. It utilized images of different soil with different moisture level. To predictmoisture in the vineyard soils from the

digital images using machine learning techniques. In this research it should be found that SVR performs excellent as compared to MLP [11].

Materials and Methods

Dataset summary

The sight map is positioned in Gallipoli, Turkey (40.3333° N, 26.5000° E). The sea temperature is low from November to April, while it is moderate from May to October. January is the coldest month and July is the warmest month while the August is driest month and December is the wettest month in Gallipoli. The metrological department of Gallipoli Turkey provides a weather dataset since 2008 to 2021.The dataset contains approximately1, 20,000 entities for different parameters as mentioned below.This research establishes a strong foundation of theoretical analysis for Gallipoli-Turkey.

Table 1 Gallipoli Weather data entities.

Dataset Preprocessing and Performance Evaluation

The CSV file contains total of 1, 22,732 entries of all parameters. Splitting them into training and testing with a standard ratio is 80% to 20%. In this case, 80% of the data is used for training whereas 20% for testing.So, there are 98,185 entries are used to train the model and reaming 24,547 entries for testing the model. By taking this amount of entries the deep neural networkbased model is developed with the concept of transfer learning. The different features are plotted in figure 1. From the figure, it can be obvious that wind direction and sunshine duration covers the whole plane means both parameters are almost constant. The remaining features are varying in nature.Figure 1 shows the graphical analysis for all parameters given in table 1 from 2008 to 2021. Table 2 shows the statistical analysis of all parameters which helps to develop a prediction model. To analyze the nature of data we simply plot a heat map among all parameters. It is used to determine the most distinct clusters or the best combination of characteristics to describe a connection between two variables. Creating some straightforward linear separations or basic lines in our data set, it also helps to create some straightforward classification models.A straightforward heat map shows adirectgraphical review of information. The viewer can better grasp complicated data sets with more intricate heat maps.From the heat map, it is quite obvious that data are varying in nature.

NeuroQuantology|July2022|Volume20|Issue8|Page4217-4229|doi:10.14704/nq.2022.20.8.NQ44456 Hiren J. Modha et al / Soil Moisture Prediction using Deep Neural Network Approach

Table 2 Analysis of all parameters.

From the heat map, humidity levels, average sea level pressures, wind speed, and prevailing winds all show a positive link with our goal characteristic soil moisture, but temperature, sunshine length, shortwave radio radiation, and soil temperature all have a negative correlation.Table 3 shows the correlation of different parameters with soil moisture. From that, it is possible to get some initiative regarding consequent weight for developing model with high accuracy and minimum errors. The various features utilized over here improve the universal potential of the model. The analysis proves that the dataset is fine for preparing a reliable model.

Spearman Correlation

Figure 3 Heat map among all parameters

Table 3 Soil moisture's Correlation with all other parameters.

To measure model performance, a person can calculate Mean Absolute Error, Mean Squared Error, and Root Mean Square Error.They are mentioned in Table 4.

Parameter	Initial	Equation
Mean Absolute Error	MAE	n MAE = $1/n(\sum Y' - Y'')$
Mean Squared Error (MSE)	MSE	$MSE = 1/n(\sum_{i=1}^{n} Y^{'} - Y^{''})^2$
Root Mean Squared Error	RMSE	n RMSE = Sqrt[1/n($\sum Y' - Y''$)]

Table 4 Performance Evaluation Parameters.

Model Preparation

Figure 4 Basic architecture of deep neural network.

Figure 4 depicts the basic architecture of deep neural network regression. It has one input layer, one output layer, and several hidden (at least two) layers. If we are going to compare, it witha single hidden layer perceptron for the same input data, The multi hidden layer approach exhibits excellent performance by reducing the nodes in each hidden layer. Deep learning techniques are based on neural networks, sometimes referred to as artificial neural networks (ANNs) or simulated neural networks (SNNs), which are a part of

machine learning. Their structure and nomenclature are modeled after the human brain, mirroring the communication between organic neurons. It has an input layer, followed by hidden layers, and finally an output layer. With a certain threshold and weight, all neurons are connected with one another.Any node whose output exceeds the defined threshold value is activated and begins providing data to the network's uppermost layer. If not, there is nothing to send to the

network's subsequent tier.The numerical explanation of a deep neural network is:

- Total no. of nodes in input layer = Total no. of parameters
- To remove problems of overfitting and underfitting it requires proper selection of no. of hidden layers.
- Neurons comprise the hidden layers. Both rectifier activation and aggregation tasks are performed by the neurons.
- ReLU is by the default activation function
- With ReLU the neurons exhibit sparse characteristics and also remove the problem of overfitting. It also increases the depth of the

 $F[Y; Z, C] = \sum_{j} [Z_{j}^{T}Y_{j} + C]$ (2)

architecture and hence the speed of training model. It is defined as

```
G(x) =Maximum (0, x) (1)
```
- The output layer of both Regression and classification are dissimilar. The regression model's output layer is a distinct node. To produce the regression prediction assessment, the result of the preceding hidden layer is multiplied by the weight, which is then added to the bias on the output node.Equation 2 explains the procedure where bias is denoted with C and j is no. nodes in thepreceding layer.
- The result of the preceding layer becomes the input feed for the subsequent layer. Here Y is the input parameter, and Z representsthe weight for a particular layer. Here C as well as C' are biases of the node.

F [Y; Z, C, z, b'] = $Z^{T} \sum_{j} [Z_{j}^{T} Y_{j} + C] + C'$ (3)

 The different optimization techniques like AdaGrade, ADAM, etc. are used which automatically adjust the learning rate which isdecreased with regularly occurring parameters to evade fluctuation of parameters.

To decrease error, model parameters (weights and biases) are adjusted based on comparisons of predicted values outcomes and labeling for both the training and test set characteristics. For the deeper neural network regression model, this is known as supervised learning. Training is ended when the maximum number of training directions is achieved or the intended outcome is met.

The model's training time and prediction accuracy can both be directly impacted by the number of veiled layers and unseen layer nodes. In order to forecast soil moisture, this study analyses eight meteorological data elements and one variable related to wetness inside the soil.The output layer limits the number of nodes (based on the regression features) to 1; the no. of input layer nodes, therefore, is 8, which is the same as the no. of features; and as per the data set, 4 hidden layers in the hidden layer structure are adequate to satisfy the necessities. Multiple testing iterations must be performed in order to assess and choose the number of each layer'sconcealed nodes.

Table 5Outcomes from different hidden layer architecture

The different combinations of nodes for four hidden layers were executed. Table 5 shows a few good combinations which provide good results for the model. Among all combinations, the last group 8-256-128-128256-1architecture is selected to prepare the model. The model weights are initialized uniformly with mean=0 and variance=1.

Algorithm

Outcomes with Hyper Parameter Tuning

To enhance the performance of the model we select two parameters for hyper parameter tuning. The first one is the optimizer function, and the second one is the learning rate. The mathematical and graphical analysis is shown below. Also, the research is executed with different activation functionsReLU and Sigmoid but it

does not make any change inside it. Table 6 shows the results for different optimizer function. It is quite observable that ADAM, RMSprop, and SGD functions perform almost equal as compared to the AdaGrad optimizer. We select the ADAM optimizer because it enhances the speed of the model.

NeuroQuantology|July2022|Volume20|Issue8|Page4217-4229|doi:10.14704/nq.2022.20.8.NQ44456 Hiren J. Modha et al / Soil Moisture Prediction using Deep Neural Network Approach

Soil Moisture Prediction

Figure 5 [A to D]Prediction accuracy analysis of Model for Gallipoli-Turkey weather dataset with different optimizer function

Table 6Model Analysis for different optimizer functions.

Table 7 Model Analysis for different learning rates.

NeuroQuantology|July2022|Volume20|Issue8|Page4217-4229|doi:10.14704/nq.2022.20.8.NQ44456 Hiren J. Modha et al / Soil Moisture Prediction using Deep Neural Network Approach

[A] Learning Rate = 0.00001 [B] Learning Rate = 0.0005

Figure 6 [A to G] Prediction accuracy analysis of Model for Gallipoli-Turkey weather dataset with different Learning Rate.

In particular, the learning rate is a hyperparameter that can be customized and is used to train neural networks. Its value is typically modest and positive, falling between 0.0 and 1.0. How quickly the model adapts to the challenge is determined by the learning rate. Table 7 shows the outcomes of model with different learning rates. The learning rate with a value of 0.1 fails as it gives error as compare to others. From the remaining entities we select learning rate 0.001 as it performs the best and increase the speed of model.

Model Consistency

This model is also test on other dataset to evaluate its reliability. For this task different parameters like soil moisture, soil temperature, environmental temperature, CO2 in the air were continuously measured with use of Raspberry-pi board. The outcomes from the board are clouded on and hence the dataset is prepared as shown in figure 6. The sight map is located at Vasad –Gujarat-India (22.4526° N, 73.0650° E). Approximately, there are more than 5000 entries of each parameter inside the dataset. A block diagram for the same is shown in figure 6. The sensors are connected with appropriate pins of Raspberry-Pi. The sensors sense continuously, all the parameters with delay of 30 seconds. This all parameters are clouded and from the cloud particular CSV file is taken. After this process, the proposed methodology is applied to predict future values of soil moisture. Here soil moisture is dependent value while soil temperature, environmental temperature environmental humidity and CO2 are dependent variable. The CSV file contains total 5571 entries of all parameter. Splitting them in to training and testing with a standard ratio is 80% to 20%. Here the 80% data is for training and 20% data is for testing. So, there are 4456 entries are used to train the model and reaming 1115 entries for testing the model.

Figure 6[A]IoT based module to generate weather data&**6[B]** Graphical analysis model for soil moisture prediction

Figure 7 Prediction accuracy analysis of Model for self-generated weather dataset

In comparison to conventional machine learning methods, neural networks typically require more no. of entities in dataset. RMSE values between 0.2 and 0.5 indicate that the model can reasonably predict the data reliably. Here in proposed technique the dataset contains approximately 5000 data although the proposed model provides best results in terms of error parameters

Discussion

There were climatic extremes at Gallipoli. The men's meager water supplies seemed much more insufficient during the scorching summer months when sickness and insects were easier to spread. In these conditions water for farming is a big issue and hence soil moisture forecasting is significant. As a result, the evaluation of a soil moisture model is the major topic of discussion in this work. The major factors for the input variables of the present soil moisture prediction model are the air temperature, sunshine duration, relative humidity, mean sea level pressure, soil temperature, soil moisture, shortwave radiation, wind speed, and wind direction. One of the answers is a proper combination of variables (including those described above). Soil moisture prediction is successful because different model characteristics necessitate different input variables.The accuracy of predicting soil moisture may be greatly increased by using appropriate meteorological factors as the model's input characteristics. The types and volumes of monitoring data are continually expanding due to the agricultural Internet of Things' quick growth [12]. Therefore, a model must provide forecast accuracy while having enough data compatibility and expandability. However, soil moisture exhibits substantial geographical variations, making it challenging to directly evaluate the effectiveness of prediction models built utilizing various areas and their associated datasets. To assess the benefits and drawbacks of various models, assessment indicators must be used as qualitative and quantitative measuring criteria. As a result, there are difficulties that need to be addressed with the choice of input characteristics and models as well as the assessment of model performance once it has been fully built. The DNNR model utilized here contains multiple input features, demonstrating that it can retain prediction accuracy when enlarging the feature types.

Conclusions

Soil moisture data is a non-stationary time series with significant changes that demonstrates a cyclic pattern of normal variation. Soil temperature, air temperature, and humidity all play a role in predicting soil moisture. Each parameter feature, according to correlation theory, has a connection with the moisture parameter, which influences the estimated outcome.The multiple input variables that were mentionedin this work were preferred as inputfor the prediction model.Experiments have shown that having too many layers in the model can result in overfitting, which reduces training accuracy and generality and increases training time. But boosting the number of layers gives a shortcut to increasing the model's capacity with fewer resources. Finally, it was decided that the structure of the proposed model would be best served by a four-layer hidden layerin the first layer, input characteristics are learned. The second layer is in charge of the polynomial fitting of the learnt features,if there areseveral nodes, overfitting will occur, which will lower prediction accuracy and generalization ability. The model structure was ultimately found to be 8-256-128-128-256-1. Table 8 shows the generalization ability of the proposed DNNR model. The proposed model is also performing well with self-generated dataset with different features. According to the research, the DNNR model has a high degree of generalization and scalability. This model may be used to anticipate soil moisture and provide technical support for irrigation methods as well as drought control.

4228

References

- [1] Yu Cai, Wengang Zheng, Xin Zhang, Lili Zhangzhong, XuzhangXue. Research on soil moisture prediction model based on deep learning.PLoS ONE 14(4): e0214508. https://doi.org/10.1371/journal. pone.0214508, April 2019.
- [2] Liao RK, Yang PL, Wang ZH, Wu WY, Ren SM. Development of a soil water movement model for the superabsorbent polymer application. Soil Science Society of America Journal, 2018a, 82(2): 436–446.
- [3] Li X.; Huo Z.; Xu B. Optimal allocation method of irrigation water from river and lake by considering the field water cycle process. Water 2017, 9(12), 911.
- [4] Shahdany SMH, Firoozfar A, Maestre JM, Mallakpour I, Taghvaeian S, Karimi P. Operational performance improvements in irrigation canals to

overcome groundwater overexploitation. Agricultural Water Management, 2018, 204:234– 246.

- [5] Liao RK, Yang PL, Wu WY, Luo D, Yang DY. A DNA tracer system for hydrological environment investigations. Environmental Science & Technology, 2018b, 52(4): 1695–1703.
- [6] Liao RK, Yang PL, Yu HL, Wu WY, Ren SM. Establishing and validating a root water uptake model under the effects of Superabsorbent polymers. Land Degradation & Development, 2018c:1–11.
- [7] Feki M, Ravazzani G, Ceppi A, Milleo G, Mancini M. Impact of infiltration process modeling on soil water content simulations for irrigation management. Water 2018, 10(7), 850;
- [8] Li N, Zhang Q, Yang FX, Deng ZL. Research of adaptive genetic neural network algorithm in soil moisture prediction. Computer Engineering and Applications, 2018, 54(01): 54–59+69.
- [9] Ji RH, Zhang SL, Zheng LH, Liu QX. Prediction of soil moisture based on multilayer neural network with multi-valued neurons. Transactions of the Chinese Society of Agricultural Engineering, 2017, 33(S1): 126–131.
- [10] Wang JR, Chen TJ, Wang YB, Wang LS, Xie CJ. Soil near-infrared spectroscopy prediction model based on deep sparse learning. Chinese Journal of Luminescence, 2017, 38(01): 109–116.
- [11] Chantal Saad Hajjar, Celine Hajjar, Michel Esta and YollaGhorraChamoun.Machine learning methods for soil moisture prediction in vineyards using digital images.https://doi.org/10.1051/e3sconf/20201670 2004, E3S Web of Conferences 167, 02004 (2020),

ICESD 2020

[12] Bright Keswani, Ambarish G. Mohapatra, Amarjeet Mohanty, Ashish Khanna, Joel J. P. C. Rodrigues, Deepak Gupta, Victor Hugo C. De Albuquerque.Adapting weather conditions based IoT enabled smart irrigation technique in precision agriculture mechanisms.Neural Computing and Application[s.Neural Computing and Applications.](https://dl.acm.org/toc/ncaa/2019/31/1) [Volume 31,Issue:1J](https://dl.acm.org/toc/ncaa/2019/31/1)anuary 2019, pp 277– 292[,https://doi.org/10.1007/s00521-018-3737-1.](https://doi.org/10.1007/s00521-018-3737-1)

- [13] Shu SF, Qian HF, Qiu XW. Soil moisture forecast model based on meteorological factors in Jinhua City. Chinese Journal of Agrometeorology, 2009, 30(02):180–184.
- [14] Hummel JW, Sudduth KA, Hollinger SE. "Soil moisture and organic matter prediction of surface and subsurface soils using an NIR soil sensor." Computers and electronics in agriculture 32.2 (2001): 149– 165.

- [15] Hou XL, Feng YH, Wu GH, He YX, Chang DM. Application research on artificial neural network dynamic prediction model of soil moisture. Water Saving Irrigation, 2016(07):70–72+76.
- [16] Li N, Zhang Q, Yang FX, Deng ZL. Research of adaptive genetic neural network algorithm in soil moisture prediction. Computer Engineering and Applications, 2018, 54(01): 54–59+69.
- [17] Gill MK, Asefa T, Kemblowski MW, Mckee M. "Soil moisture prediction using support vector machines 1." JAWRA Journal of the American Water Resources Association 42.4 (2006): 1033–1046.
- [18] Chen XF, Wang ZM, Wang ZL, Li R. Drought evaluation and forecast model based on soil moisture simulation. China Rural Water and Hydropower, 2014(05): 165–169.
- [19] Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. Neural computation, 2006, 18(7): 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527 PMID: 16764513

 \blacksquare