A Machine Learning Approach for Automated Irrigation Management System

Pulin Prabhu[1], Anuj Purandare[2], Abhishek Nagvekar[2], and Aditi Kandoi[1]
Sunil Ghane[1] Mahendra Mehra[2]

[1] Department of Computer Engineering, Sardar Patel Institute of Technology, Mumbai, India
[2] Department of Computer Engineering, Fr. CRCE , Mumbai, India
{pulin.prabhu,aditi.kandoi,sunil_ghane}@spit.ac.in
{anujpurandare,abhisheknagvekar8,prof.mahendramehra}@gmail.com

Abstract. The scarcity of food and water has become a serious issue in the country and requires an immediate solution. This can be overcome by ensuring efficient distribution of water in a variety of sectors but most importantly, the agricultural sector. The current systems use hardware sensors to determine the moisture in real-time, which are expensive and can produce an error with wear and tear. Therefore, an automated system for irrigation mechanism for water conservation is required. The said research deals with an automated irrigation system that analyzes the moisture content of the soil using the weather data, time of the day, and past irrigation data to predict the soil moisture which helps to in turn determine an optimal irrigation schedule. The said system, based on various factors, predicts moisture values with the help of an artificial neural network eliminating the use of hardware.

Keywords: Automated Irrigation System · Moisture content · Irrigation schedule · Machine Learning · Artificial Neural Network

1 Introduction

In India, due to uneven distribution and overuse of water in irrigation, the efficiency of irrigation systems is only about 25%. Moreover, India's population comprises the world's 15% population however, it has only 4% of the world's freshwater resources. This leads to dryness of soil and wastage of water on a huge scale. The existing irrigation systems use manual irrigation, drip irrigation sprinkler irrigation, etc. However, due to the irregularity of the amount of water distribution in farms, these techniques result in a lot of problems like land degradation, plant health, and waste wastage. Hence, to avoid these issues, there is an urgent need for the improvement of existing systems.

Therefore, the paper proposes an automated irrigation and nutrient management system along with a prediction model for irrigation tables using an artificial neural network that helps the other farmers. Mobile-based irrigation solutions have been designed with various components to satisfy constraints like energy and economical saving. In this implementation, the paper suggests an automated system that helps one save not only time and money but also valuable resources like water and land. Also, having a water level sensor constantly take readings or having an irrigation schedule will help in the prevention of dry running of the pump. This proposed system helps the consumers and farmers who cannot afford existing expensive systems and suggests irrigation and nutrition schedules by its prediction model.

2 Literature Review

In the described method, the authors initially perform a check of the moisture level of the soil. This gives an idea if the soil requires water or not. If required the system with the help of its ESP-01 sensor (wifi module) and relay supply required voltage for the Pump to supply water to the soil. [1]

In the proposed system the authors want to calculate the percentage of basic nutrients of the soil and get the amount for the particular type of soil. The system determines the said things in real-time.[2]

In this paper, the authors say that the system will determine the moisture content in the soil and provide appropriate water to the plant. Through this implementation, the system makes the plant more self-sufficient and maintains its health. [3]

These suggested systems work on various microcontrollers. They are set such that they detect the moisture level of the soil with additional sensors provided. The authors suggest that if the soil has enough moisture then the system will itself pause for a given time and when required will start hence saving valuable resources like manpower. [4] [5] [6][7] [8] [9]

In this system, the authors suggest a mechanism that will find the moisture content in the soil. Another suggestion made here is to constantly capture images of the plant to study the health of the plant and to determine whether it has any diseases. [10]

The author suggests having a fully automated system where one can have access to multiple features like temperature, humidity, moisture, and fertilizer levels can be monitored and triggered to get rid of abnormalities to maintain an evergreen farm/garden with good plant health. [11]

To reduce manpower the authors suggest using various ways like measuring pH levels of the soil to determine the Nutrient levels. This way the fertilizer quantity can be determined according to the requirement of the crop. With this, the cost will decrease and fertilizer efficiency will improve ensuring better returns to human health and the environment. [12]

These papers talk about how computer technologies like artificial intelligence and deep learning can be implemented in the field of agriculture to eradicate various problems like diseases of plants, pesticides, and the conservation of water and other natural resources. [13] [14]

The author suggests a system that can be pocket-friendly and at the same time could increase the efficiency of crop production. For the said system there is an Arduino-based camera that transmits images and soil moisture levels to another microcontroller for computation that in turn handles the water supply. [15]

The author suggests similar measures as others but also highlights the benefits of this system as the agriculturist who is elderly can make use of the said system and have full control over their harvest instead of handing it over to someone to take care of it. [16]

In this system, a Dual Outlet Tap is used and the rest of the system is designed upon it. A moisture sensor is connected to the circuit as each plant has different traits the system takes the reading of each plant soil and the water is supplied by considering each plant's need. [17]

3 Methodology

The data set used to check the proper functioning of the system was collected in an experiment conducted by Wazihub with 4 IoT sensors over 4 months in 4 fields in Senegal. The raw data comprises of 7 columns and 28049 rows.

The setup to obtain the readings for development of the system was done by having four separate pieces of ground planted with either maize or peanuts. An IoT sensor was placed such that the quantity of the yield was approximately the same. The plots, separated by one meter perimeter, were next to each other.

Following are the steps for the prediction of irrigation schedule:

3.1 Adjusting Data Types and removing missing values

According to the algorithm, the data is pre-processed for making the necessary calculations. Testing In case of missing values, it is tested to make the data consistent.

3.2 Analysing the Data Pattern

Frequency Histograms The analysis of the frequency of each data item was done and it was plotted as a histogram. This helped in understanding how various factors are associated with the topography of a particular region that the data-set focused upon. The IoT soil moisture and weather setup in and around the field would communicate with the following data in 5 minute intervals, given below are our findings from each of the data values :

- Soil Humidity : In Figure 1, the values mostly lie between 60 and 70, with a mean value of 63.025. The minimum value is 36 and maximum value is 88.
- Air Temperature: In Figure 2 the values mostly lie between 15 and 20, with a mean value of 24.26. The minimum value is 11.22 and maximum value is 45.56.
- Air Humidity: In Figure 3 the values mostly lie between 80 and 100, with a mean value of 58.52. The minimum value is 0.59 and maximum value is 96.

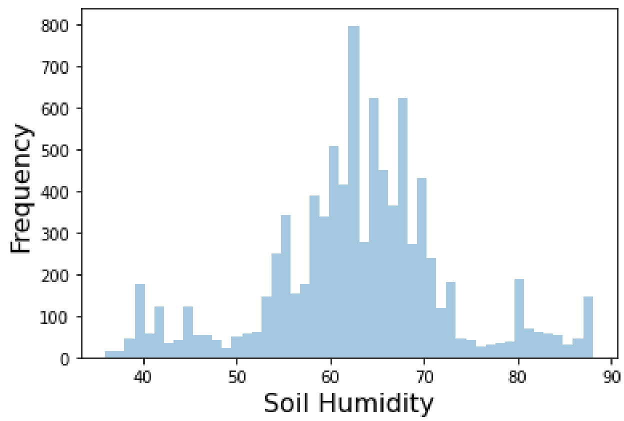

Fig. 1. Soil Humidity Analysis

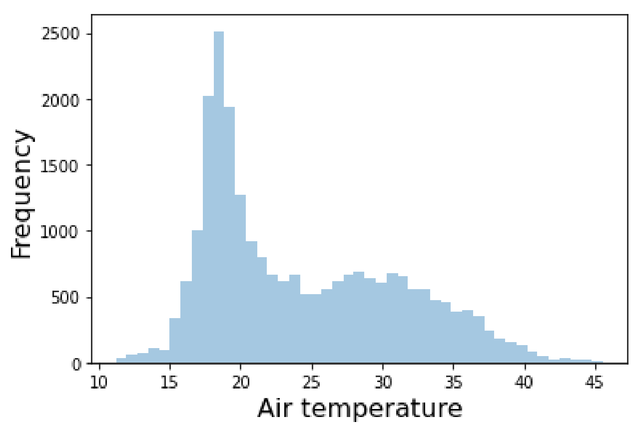

Fig. 2. Air Temperature Analysis

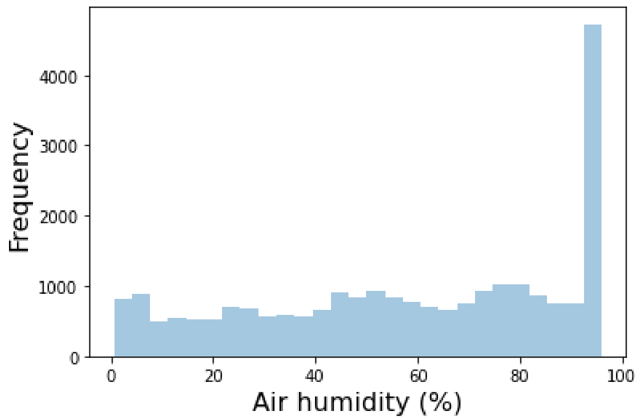

Fig. 3. Air Humidity Analysis

– Pressure: In Figure 4 the values mostly lie between 101.0 and 101.4, with a mean value of 101.13. The minimum value is 100.5 and maximum value is 101.86.

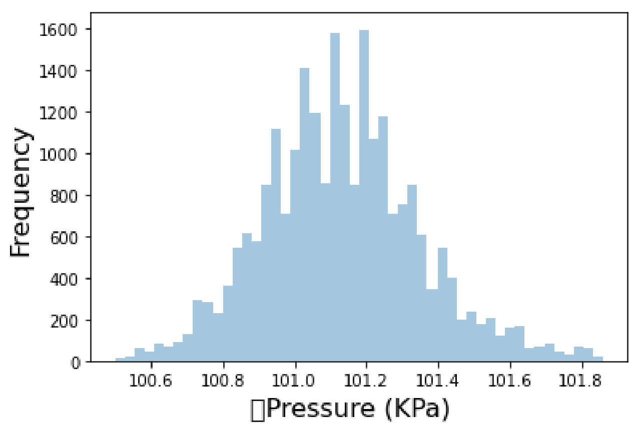

Fig. 4. Pressure Analysis

– Wind Speed: In Figure 5 the values mostly lie between 5 and 15, with a mean value of 9.89. The minimum value is 0 and maximum value is 31.36.
– Wind Gust: In Figure 6 the values mostly lie between 0 and 40, with a mean value of 41.74. The minimum value is 0 and maximum value is 133.33.
– Wind Direction: In Figure 7 the values mostly lie between 0 and 350, with a mean value of 93.98. The minimum value is 0 and maximum value is 337.

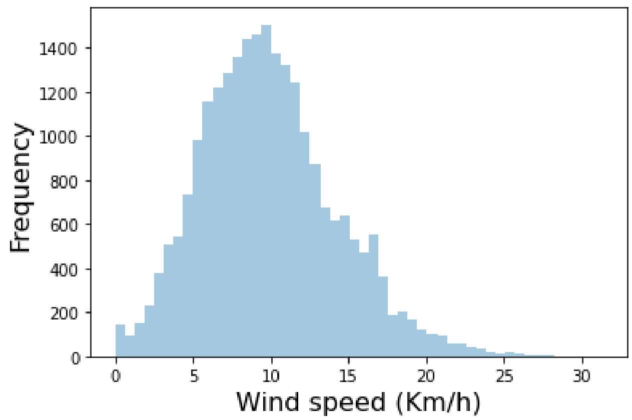

Fig. 5. Wind Speed Analysis

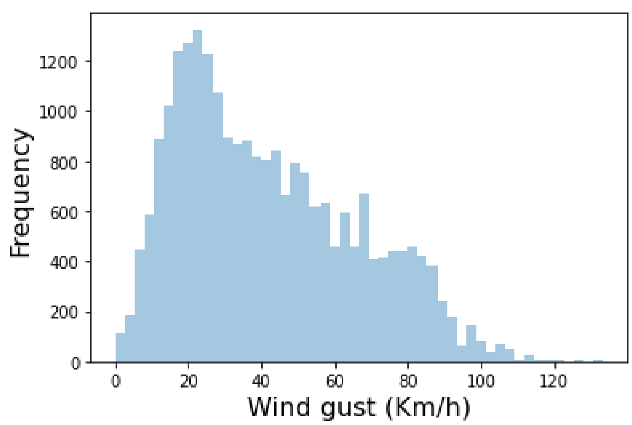

Fig. 6. Wind Gust Analysis

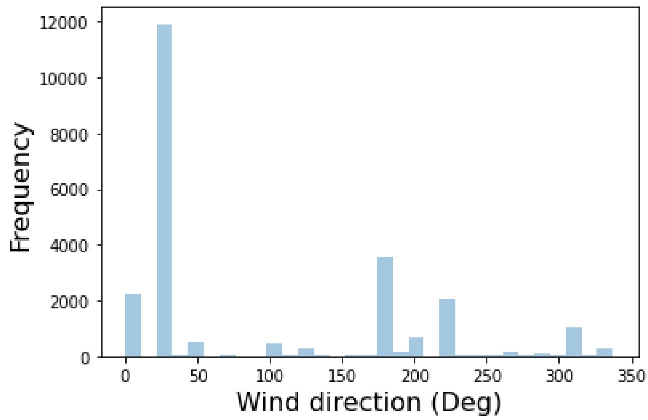

Fig. 7. Wind Direction Analysis

Heat Maps As seen in Figure 8, heatmap is a graphical representation of data that uses a system of color-coding to represent different values. A heatmap is plotted to understand the correlation between soil humidity and other features. A correlation matrix is a table showing correlation coefficients between variables. Air temperature and Air humidity show maximum correlation. On the contrary, pressure, wind gust and wind direction show minimum correlation.

For training and testing purposes, the dataset is divided in a 70-30 way, where 70% is used for training and the rest 30% for testing and validation purposes.

3.3 Feature Engineering

In order to analyse and utilize the given timestamps feature for an overall better prediction, two new features (columns) namely ”days” and ”time of day” are created to record and understand how the time of the day and days together affect the overall soil moisture reading. It turned out in the further analysis that these new features are important in predicting the moisture readings.

The time-of-day has been split into 4 buckets as given in Table 1.

Table 1. FEATURE ENGINEERING

Table Head	Splitting timestamps to time of day	
	original timestamp Values	***Assigned time of day***
1.	between 00hrs and 06hrs	Midnight
2.	between 06hrs and 12hrs	Morning
3.	between 12hrs and 18hrs	Afternoon
4.	between 18hrs and 00hrs	Night

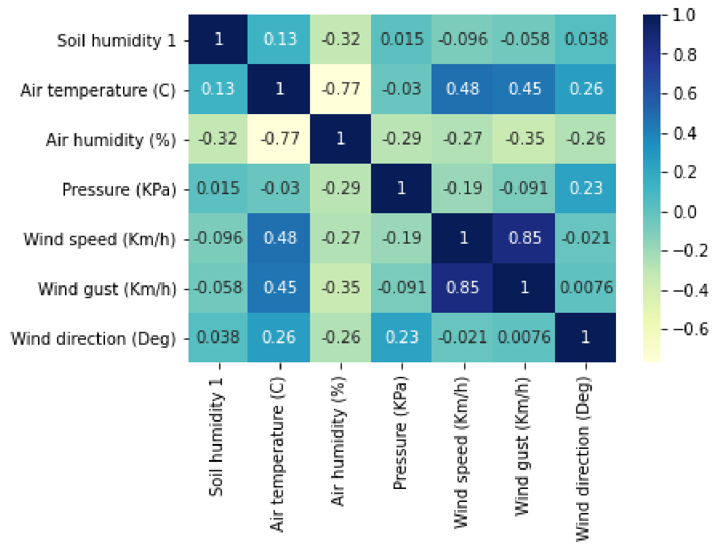

Fig. 8. Wind Direction Analysis

- New Feature, Days : This particular feature is calculated by only consider the day part of the timestamp column. We consider the earliest day of the year as day 1 and further calculate the day for each row.
- New Feature, time of day : The 24 hours of a day have been split into 4 brackets as seen in the histogram.

3.4 Hyperparameter tuning

A hyperparameter is a parameter whose value is used to control the learning process and they tend to define the architecture of the model. Here as different classification algorithms are to be used to build the model, there is a need for a process of searching for the ideal model architecture. Hyperparameter tuning is a step that can get this job done.

3.5 ML Algorithms

The approach to the solution is finding out the best Machine Learning algorithm that will give the most accurate result for detection of frauds in transactions.

Support Vector Machine Support Vector Machine uses a kernel. A kernel helps us find a hyperplane in the higher dimensional space without increasing the computational cost. Usually, the computational cost will increase if the dimension of the data increases. We are getting RMSE of 6.8.

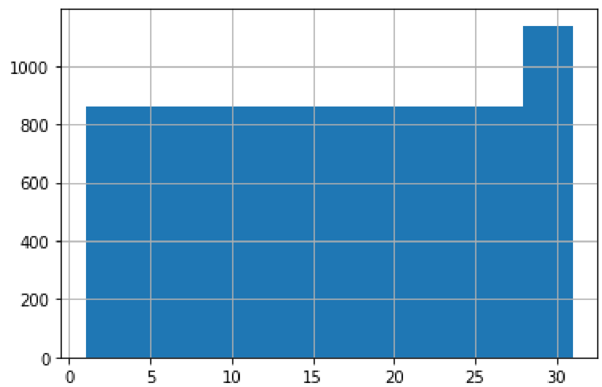

Fig. 9. Days

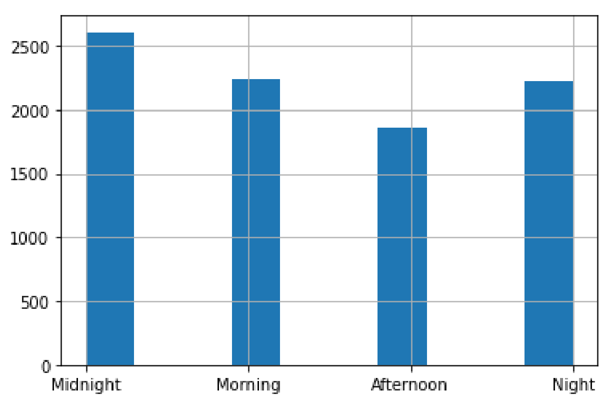

Fig. 10. Time of Day

K Nearest Neighbor K nearest neighbors is a simple algorithm that stores all available cases and predict the numerical target based on a similarity measure (e.g., distance functions). Using KNN we got a RMSE of 7.4.

XGBoost XGBoost is a decision-tree-based ensemble Machine Learning algorithm that uses a gradient boosting framework. Using XG Boost we got a RMSE of 7.4.

LGB Classifier LightGBM is a gradient boosting framework that uses tree based learning algorithm. LightGBM grows tree vertically while other tree based learning algorithms grow trees horizontally. LightGBM is a fast, distributed, high performance gradient boosting framework

3.6 Artificial Neural Network

ANN stands for artificial neural networks, It's essentially a computation model. It's based off of the biological neural network architectures and functions. Since a flow of data affects the structure of the ANN, changes rely on input and output characteristics in the neural network.We should consider ANN as data which is not linear.Which implies a dynamic relationship between input and output.

Structure : ANN takes from how the human brain makes correct decisions. In fig xx we can see how structure of ANN is shown by representing the neurons that are a part of the human brain which composes of 86 billion nerve cells which are in turn connected to thousands of cells with the help of Axons. The dendrites help in receiving inputs from various organs, as a result of which electric impulses are created. These impulses travel through the artificial neural network from a neuron to another in order to handle different tasks.

It is seen that the ANN is composed of numerous nodes which represent the neurons in the human brain. These nodes are connected via links. Just like neurons the nodes in the ANN accept data, perform operations, result of which the operations are passed on to other nodes. Furthermore there are weights attached to each node which when altered make the network capable of learning. The output at the end of the node is called the node value.

We will now see the two kinds of :

- **FeedForward ANN :** The information flow in the feed forward ANN is unidirectional. While containing fixed input and outputs, the Feed Forward ANN does not have any kind of feedback loop. Feedforward artificial neural network model will be used in the proposed system.
- **FeedBack ANN :** The Feedback ANN gives the capability for feedback loops to exist.

Working of ANN : In fig xx, every arrow represents a connection. The flow of information throughout is represented by these arrows. We also saw that there are weights attached to these connections that help in making it capable of learning but are also used controlling the signal between two nodes. To improve the results if not accurate, the system can require the weights to be altered. If output generated is accurate, no weight alteration shall be required.

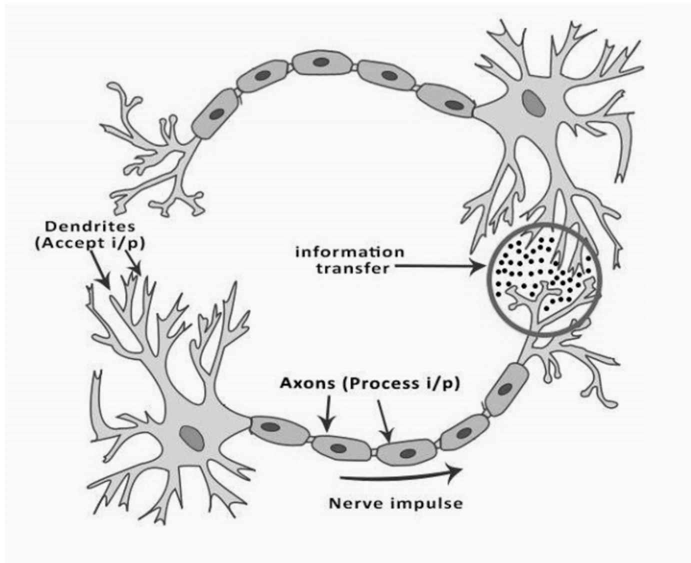

Fig. 11. Basic Structure of Artificial Neural Network

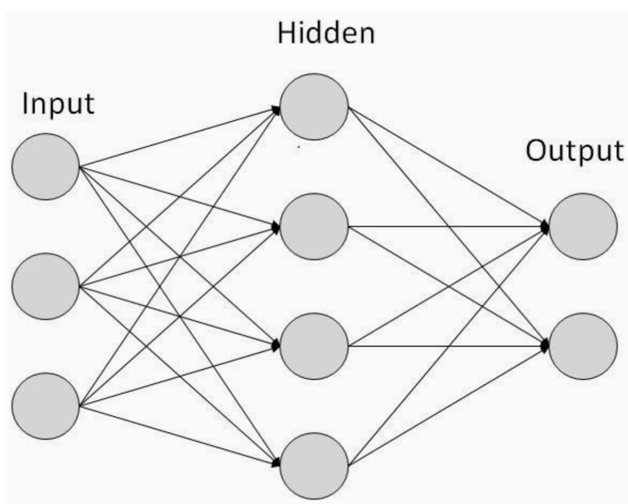

Fig. 12. Artificial Neural Network Structure

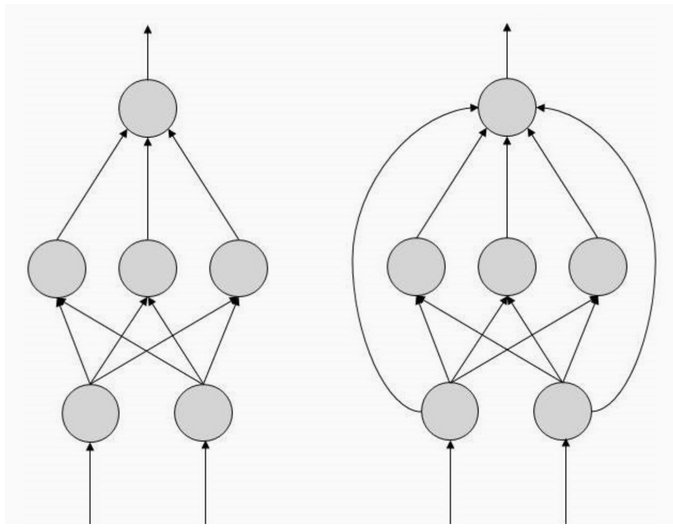

Fig. 13. Types of Artificial Neural Networks – FeedForward ANN

3.7 Experimental Setup

Table 2. Model Architecture for ANN

Table Head	**Model architecture**	
	Hyperparameters	***Values***
1.	Input Layer[a]	1/30 neurons
2.	Output Layer[a]	1/1 neurons
3.	Training Epochs[a]	10
4.	Optimizer[a]	Adam
5.	Error Propagation[a]	Gradient Descent

Network Architecture We shall now look at the various hyper parameters considered as shown in TABLE 1 and the rationale behind choosing them.

- The input layers are chosen to be at 1/30 neurons and the output layer is chosen to be at 1/1 neuron for getting a greater accuracy at higher number of data points.
- The input layers are chosen to be at 1/30 neurons and the output layer is chosen to be at 1/1 neuron for getting a greater accuracy at higher number of data points.
- Unlike Iterations epochs use one forward and one reverse pass of all training data which optimizes the time and gives accurate results hence a epoch of 10 is chosen.

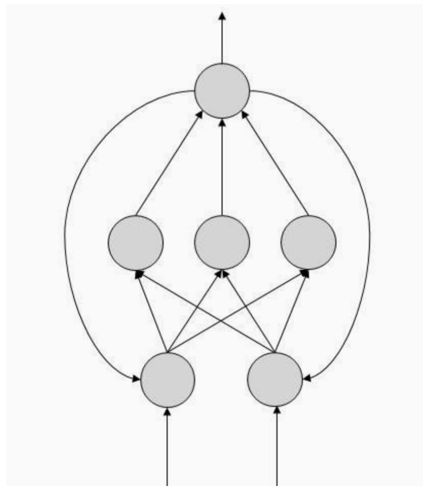

Fig. 14. Types of Artificial Neural Networks – FeedBack ANN

- A few of the reasons behind choosing SGD as an Optimiser and Gradient Descent for Error Propagation in our case are :
  - Stochastic Gradient descent works better than batch training since it performs updates more frequently
  - The whole data set is not required for Approximation of the gradient
  - Taking the advantage of vectorised operations to process the entire mini batch at once can make the training faster on a single data point.
  - In the multivariate case, the formulae are a little more complex on paper when it comes to several variables and require even more calculation when implementing it in software and is given by:
    β =(X X) 1X Y

With the above mentioned experimental setup for Training and Validation the model performs with great accuracy, RMSE being used as an error metric, and is found out to be 0.536 units. This is further discussed in the results section as this model performs best amongst the ones discussed during the course of this research.

4 Result

In this section, we present the experimental results. As mentioned earlier, our goal was predicting the soil humidity using machine learning / artificial neural networks.

The root mean square error(RMSE) is the standard deviation of all the predicted errors. The said errors are a measure of how far are the data points from the actual regression line. Strictly speaking, it shows how closely the data is clustered around the line of best fit.

Table 3. Results Table

Table	Model Comparison	
Head	***Models***	***RMSE Values***
1.	SVM	6.221
2.	KNN	2.06
3.	XG Boost	2.346
4.	LGBM	0.684
5.	ANN	0.536

As seen in the table above, Support Vector Machine performed worst among all the models with an RMSE of 6.221. KNN and XG Boost performed moderately with RMSE values 2.06 and 2.346 respectively. LGBM was almost comparable to ANN in terms of performace with an RMSE of 0.684 while ANN performed the best with an RMSE of 0.536.

4.1 Observations

Fig. 15.

As seen in the Training (Fig. 6.) and Validation (Fig. 7.) graphs the model performs with great accuracy and produces predictions that are very accurate. To prove the system accuracy, RMSE is used as an error metric, and is found out to be 0.536 units. Furthermore there hasn't been any use of hidden layers in the artificial neural network to avoid the over fitting of data.

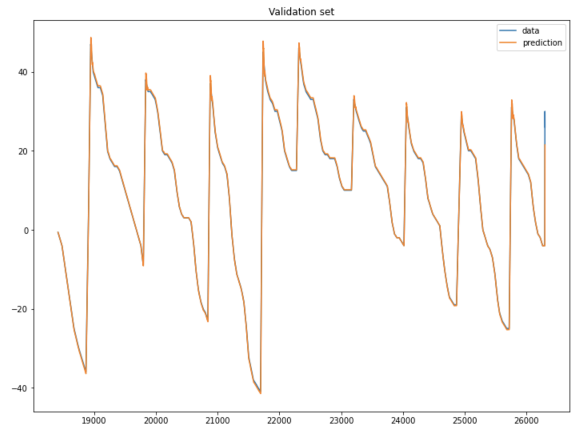

Fig. 16.

5 Conclusion

The present irrigation methods or for that matter the automated irrigation systems either require high cost hardware or waste some valuable natural resources hence, we have designed a system which works with or without hardware. The proposed system is programmed to read the data recorded to calculate the schedule, that is predicting the irrigation schedule based on the historic data values collected by the sensors and weather data that is present in the database with the help of machine learning and artificial neural network. By adopting this system there is an all round savings because of judicious amount of water usage, power supply saving without the use of hardware. With this when the Water Table and larger quantity of healthy land increases, the agricultural produce increases which in turn will make the farmer happy, the interest towards farming/agriculture would also increase and accordingly the Problem of Food Scarcity would also be addressed.

6 Future Scope

The project demonstrates a system which is able to determine a unique approach combining various features in order to perform irrigation schedule prediction. Having achieved a significant improvement in classification accuracies, the future scope of this study can be identified as deploying hardware alongside the machine learning or intelligent computer systems in order to get specific and accurate results as per the topography.

References

1. Sarode M, Shaikh A, Krishnadas S, Rao DY. Automatic Plant Watering System. AJCT [Internet]. 15Dec.2020 [cited 18Aug.2021];6(3):90-4. Available from: https://asianssr.org/index.php/ajct/article/view/977
2. Minal Vaity Daresh Vadalia and Dyananeshwar Kapse. "Real Time soil fertility analyzer and crop predic-tion". In:International Research Journal of Engineering and Technology. 2017, pp. 516–521.DOI: 10 .1109/ICCPEIC.2017.8290420.
3. S. Bhardwaj, S. Dhir, and M. Hooda. "Automatic PlantWatering System using IoT". In:2018 Second Interna-tional Conference on Green Computing and Internet ofThings (ICGCIoT). 2018, pp. 659–663.DOI: 10.1109/ICGCIoT.2018.8753100
4. C. M. Devika, K. Bose, and S. Vijayalekshmy. "Auto-matic plant irrigation system using Arduino". In:2017IEEE International Conference on Circuits and Systems(ICCS). 2017, pp. 384–387.DOI: 10.1109/ICCS1.2017.8326027.
5. A. Patil, M. Beldar, A. Naik, and S. Deshpande. "Smartfarming using Arduino and data mining". In:2016 3rdInternational Conference on Computing for Sustainable-Global Development (INDIACom). 2016, pp. 1913–1917.
6. N. S. Ishak, A. H. Awang, N. N. S. Bahri, and A. M. M.Zaimi. "GSM activated watering system prototype". In:2015 IEEE International RF and Microwave Conference (RFM). 2015, pp. 252–256.DOI: 10.1109/RFM.2015.7587756.
7. D. Divani, P. Patil, and S. K. Punjabi. "Automated plantWatering system". In:2016 International Conferenceon Computation of Power, Energy Information andCommuincation (ICCPEIC). 2016, pp. 180–182.DOI:10.1109/ICCPEIC.2016.7557245.
8. I. Primisima, S. A. Sudiro, and B. A. Wardijono."Automatic plant watering controller component usingFPGA device". In:2015 International Conference onAdvanced Computer Science and Information Systems(ICACSIS). 2015, pp. 43–49.DOI: 10.1109/ICACSIS.2015.7415167.
9. B. N. Getu and H. A. Attia. "Automatic control ofagricultural pumps based on soil moisture sensing".In:AFRICON 2015. 2015, pp. 1–5.DOI: 10 . 1109 /AFRCON.2015.7332052.
10. M. H. Sai Kumar K. Krishna Kishore and M. B. S.Murthy. "Automatic Plant Monitoring System". In:In-ternational Conference on Trends in Electronics andInformatics (ICEI). 2017, pp. 516–521.DOI: 10.1109/ICCPEIC.2017.8290420.
11. M. O. Sharma and P. M. Sonwane. "Remote monitoringand control for liquid fertilizer and water irrigation".In:2017 International Conference on Computation ofPower, Energy Information and Commuincation (IC-CPEIC). 2017, pp. 516–521.DOI: 10.1109/ICCPEIC.2017.8290420.
12. P. B. Sowmiya, B. K. Nagaswetha, and D. Priyad-harshini. "Design of Automatic Nutrition Supply Sys-tem Using IoT Technique in Modern Cities". In:2017International Conference on Technical Advancementsin Computers and Communications (ICTACC). 2017,pp. 109–111.DOI: 10.1109/ICTACC.2017.36.
13. Y. Jiang, J. Zhao, J. Duan, H. Hao, and M. Wang."Automatic control system of water and fertilizer basedon fuzzy control algorithm and simulation". In:2017IEEE 2nd Information Technology, Networking, Elec-tronic and Automation Control Conference (ITNEC).2017, pp. 1671–1674.DOI: 10 . 1109 / ITNEC . 2017 .8285079.
14. P. Zhang et al. "The Construction of the Integrationof Water and Fertilizer Smart Water Saving IrrigationSystem Based on Big Data". In:2017 IEEE Inter-national Conference on Computational Science andEngineering (CSE) and IEEE International Conferenceon Embedded and Ubiquitous Computing (EUC). Vol. 2.2017, pp. 392–397.DOI: 10.1109/CSE-EUC.2017.258.

15. K. K. Namala, K. K. Prabhu A V, A. Math, A. Ku-mari, and S. Kulkarni. "Smart irrigation with embeddedsystem". In:2016 IEEE Bombay Section Symposium(IBSS). 2016, pp. 1–5.DOI: 10 . 1109 / IBSS . 2016 .7940199
16. K. Lekjaroen et al. "IoT Planting: Watering systemusing mobile application for the elderly". In:2016International Computer Science and Engineering Con-ference (ICSEC). 2016, pp. 1–6.DOI: 10.1109/ICSEC.2016.7859873.
17. I. Al-Bahadly and J. Thompson. "Garden watering sys-tem based on moisture sensing". In:2015 9th Interna-tional Conference on Sensing Technology (ICST). 2015,pp. 263–268.DOI: 10.1109/ICSensT.2015.7438404