Towards a Novel Approach for Smart Agriculture Predictability

Rima Grati[1] [a], Myriam Aloulou[2] [b] and Khouloud Boukadi[3] [c]
[1]Zayed University, College of Technological Innovation, Abu Dhabi, U.A.E.
[2]Liwa College of Technology, Abu Dhabi, U.A.E.
[3]Miracl Laboratory, Faculty of Economics and Management, University of Sfax, Tunisia

Keywords: IoT, Smart Agriculture, Machine Learning, Self Organization Map, Machine Learning, Deep Learning.

Abstract: The practice of growing crops and raising cattle is the traditional method of agriculture, a primary source of livelihood. The introduction of advanced technologies and tools provides solutions to predict and avoid soil erosion, over-irrigation, and bacterial infection for crops. Machine learning and Deep learning solutions are hitting high results in terms of precise farming. The most challenging factors for research society are identifying the water need, analyzing soil conditions and suggesting the best crops to cultivate, and predicting fertilizer amounts to prevent bacteria. Grouping similar features helps with accurate prediction and classification. Considering this, we introduce an integrated model Group Organize Forecast (GOF), using Machine Learning (ML) and Deep learning (DL) techniques to balance the requirements and improve automatic irrigation. GOF analyzes the irrigation requirement of a field using the sensed ground parameters such as soil moisture, temperature, weather forecast, radiation levels, the humidity of the crop field, and other environmental conditions. We use a real-time unsupervised dataset to analyze and test the model. GOP clusters the data using Self Organizing Map (SOM) organizes the classes using Cascading Forward Back Propagation (CFBP), and finally predicts the requirement for water and solution to control bacteria in the near future.

1 INTRODUCTION

The Internet of Things (IoT) is an integrated tool used to sense data from multiple devices and evaluate the internal and external state of communication. IoT applications in smart agriculture focus on crop water management, pest control, temperature adjustments, precise detection, and nutrient management with safe storage methods. According to a study, (Bakthavatchalam et al., 2022) 64% of cultivation depends on the monsoons, during which irrigation needs 85% of the water, of which 60% is wasted in the process. Analyzing the features to monitor the water-level and temperature, which prevent the bacteria from growing based on the changes in the environment, is a major concern of the research community.

Machine learning is a popular technology and a branch of artificial intelligence that allows computers to learn without explicit programming. Machine learning techniques are used to extract the features and group the required input into clusters. Deep learning techniques are suitable to analyze, classify, and predict the requirements for sustainable irrigation management. Both techniques are effective decision-support tools for precision agriculture. Traditional framing includes manual decisions on controlled water management, crop selection, weather forecasting, and analyzing soil conditions. This can be enhanced and improved based on the needs of the crops using an integrated techniques (Pathan et al., 2020).

[a] https://orcid.org/0000-0002-6995-465X
[b] https://orcid.org/0000-0001-8404-7487
[c] https://orcid.org/0000-0002-6744-711X

1.1 Motivation and Contribution

The lack of proper measures for taking the right decision in smart irrigation is a big challenge for the smart agriculture industry (Ben Abdallah et al., 2023). Prediction of the suitable crop yield, analysis of the water requirement, and equal distribution of soil nutrition are the major requirements to be considered in the development of an intelligent IoT irrigation model. Available research models are designed for a specific field structure and depend on the sensor features (Ayaz et al., 2019). (Sarker, 2021) highlighted some of the deep learning techniques and their importance to handling the challenges faced by IoT agricul-

Grati, R., Aloulou, M. and Boukadi, K.
Towards a Novel Approach for Smart Agriculture Predictability.
DOI: 10.5220/0012082400003538
In *Proceedings of the 18th International Conference on Software Technologies (ICSOFT 2023)*, pages 96-105
ISBN: 978-989-758-665-1; ISSN: 2184-2833

ture. It is highly recommended to develop a generalized model for predicting the requirements for water, food, air, temperature, and other necessities for the plant. We show a novel prediction model to protect and develop an IoT agriculture sector with the following contributions.

- Feature extraction: GOF analyzes the required features and clusters them into similar groups using the unsupervised clustering technique Self Optimization Map (SOM). This helps to predict water requirements, the presence of bacteria, and the suitable crop.
- Classification and prediction model: GOF categorizes the classes using Cascading Forward Back Propagation Neural Network (CFBPNN) model. CFBPNN identifies the pattern of the variable and the requirements for irrigation.
- Decision Tree and correlation Analysis: GOF decides the prediction factors using the decision tree algorithm and visualizes them using the correlation plot technique. The optimal cluster helps in prediction using decision rules. Finally, SOM predicts the water requirement and the factors influencing the spread of bacteria.

1.2 Organization of the Paper

The rest of the paper is organized as follows. Section 2 reviews some machine learning and deep learning-based models and their techniques for the smart irrigation system. Section 3 Focus on the SOM clustering method to divide similar groups and classify the target data using the CFBPNN classification model. An evaluation metric with state-of-the-art comparison with GOF is given in Section 4. Finally, a logical conclusion is drawn in Section 5.

2 RELATED WORK

In agriculture, Machine Learning (ML) is used to forecast soil parameters like organic carbon and moisture content; it also predicts diseases and weeds in crops and identifies species. Remote monitoring of ambient and soil characteristics is used for agronomic applications to predict crop health. A sensor-based network is being used to forecast the watering schedule for agricultural fields. A wireless sensor network collects data from external variables such as pressure, humidity, temperature, soil moisture, salinity, and conductivity. Agricultural applications can be made incredibly simple and efficient using three stages of machine learning as data acquisition, model development, and generalization. Deep Learning (DL) methods over traditional machine learning is enhanced by adding additional complexity to the model and change the input with a range of functions that hierarchically allow data representation, based on the network architecture, through multiple levels of abstraction (Jimenez et al., 2021). Selecting appropriate input variables as temperature, humidity, moisture of soil, and wind ratio helps identify the water requirement for the plant. DL techniques are used to trace the optimistic features and predict the requirements.

Compared to ML, DL models are popular in estimation of image and sound processing. Some of the popular DL models used for optimizing irrigation decision are Artificial Neural Network (ANN), Recurrent Neural Network (RNN) with loop connectivity, Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN). Precision irrigation is the solution to deliver bigger, better, and more profitable yields with fewer resources. Some of the applications using the above methods are to identify the best crop, detect factors that destroy the crops, analyze the presence of diseases to obtain insights about crop growth and help in decision making. Some of the smart agricultural frameworks proposed in recent years are given in Table 1.

The main aim of the study to perform unsupervised clustering method, to group set of variables and identify the classes having impact over the change. Some of the research models using cluster techniques are discussed in what follows:

Application of Clustering Smart Agriculture. Clustering is aimed to identify a distinct group, based on the similarity of a given dataset, while the arrangement of data into clusters results in low inter-cluster similarity and high inter-cluster similarity (Swamynathan, 2019). An integrated study using fuzzy time series techniques and clustering to manage wireless sensor nodes plotted on the agricultural field is proposed by (Prabhu et al., 2014). The study concluded that the proposed model improved the energy efficiency of the sensors with real-time monitoring of the farm. It has been observed that the low yields are caused due to attacks from pests resulting in inadequate irrigation. This has been investigated using the K-Means clustering technique with image data by enabling the Wireless Sensor Network (WSN) for smart irrigation by (Nisha and Megala, 2014). The image data of plant leaves are segmented into cluster groups based on the feature similarities, the model gave improved results compared to other WSN-based irrigation systems with clustering models for pest detec-

Table 1: Smart agriculture models using ML and DL techniques.

Reference	Model	Results
(Mehra et al., 2018)	Deep Neural Network (DNN), Artificial Neural Network (ANN)	Classify and Control actions.
(Varghese and Sharma, 2018)	Support Vector Machine(SVM)	Reduce manual labour, regular alerts for complete field control.
(Goap et al., 2018)	Support Vector Regression (SVR) and K-means	Soil Moisture Differences (SMD)
(Lavanya et al., 2020)	Fuzzy logic	Detect deficiency of nutrients from the sensed data
(Rezk et al., 2021)	wrapper feature selection, and PART classification technique	suitable for crops: Bajra, Soybean, Jowar, and Sugarcane.
(Reddy et al., 2020)	Decision Tree(DT) algorithm	Mail alert based on DT results regarding water supply in advance.
(Kashyap et al., 2021)	Long Short-Term Memory network (LSTM)	Predict the volumetric soil moisture, and spatial distribution of water required to feed
(Akhter and Sofi, 2022)	ML Linear regression model	Categorised as safe an unsafe area based on temperature and wetting duration.
(Bakthavatchalam et al., 2022)	MLP, Decision table,JRip	Accurate prediction and implementation of precision agriculture
(Geetha Lekshmy et al., 2022)	LSTM and random forest	Water dispensed, pest detection with images of field object detection technique to avoid pests and animals.

tion. Other research by Ohana-Levi et al. (Ohana-Levi et al., 2021) uses fuzzy K-means clustering, using the hierarchical method to identify the litigation management zones in a citrus field, to determine in-field variation and adjust site-specific irrigation management. Variables like: crop water stress, Normalized Difference Vegetation Index (NDVI), digital surface model, slope, aspect, and elevation were used for the experiment. The study concludes that the infield spatial variability is not constant among the variables and within the orchard.

3 METHODOLOGY

The proposed methodology includes four stages given in Figure 1. After collecting the live data from the sensors, each feature is represented as a separate element in the module /controller. The input features are traced for a period and stored in a CSV file in the first stage. Secondly as part of the pre-processing technique, data normalization maintain the balance in the data values, and the missing information is eliminated to avoid misleads in the experiment. Finally the raw data is compressed and divided into clusters using the SOM Neural Network technique. And finally, the input and class variable(target) are trained and tested on the proposed CFBPNN prediction technique. The training is repeated till the expected results are evaluated.

3.1 Data Collection

In the PRECIMED project (PRIMA, 2023), we collected real-time data provided by SENTEK technologies using the TEROS 12 soil Moisture and Electrical Conductivity (EC) and Temperature Sensor. These tools help collect Volumetric Water Content (VWC) and monitor electrical conductivity and other resources, such as temperature and the toxicity levels of soil substances. VWC works with frequency-domain technology, connected with sensors on 70MHz frequency, to minimize salinity and textural effects. TEROS 12 is the primary tool for measuring temperature and electrical conductivity with a stainless-steel electrode array and is accurate in mineral soils.

Data provided by the above sensors are:

- Volumetric Water Content (VWC) measurement
- Soil/substrate water balance
- Irrigation management

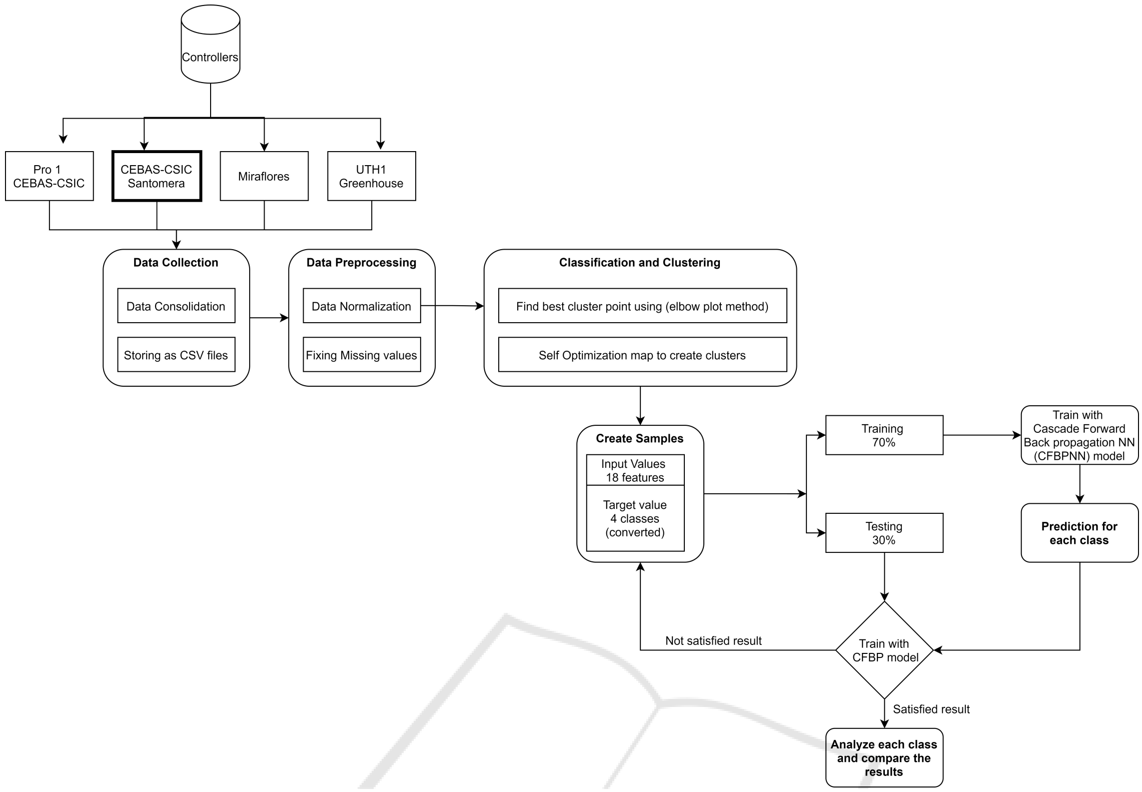

Figure 1: Procedure for proposed GOF integrated model.

- Soil Electrical Conductivity (EC) measurement
- Soil/substrate temperature measurement
- Solute/fertilizer movement

The live data is categorized with four controllers, each controller supports with input features acquired from various sensing devices.

The study gather the input from Finca experimental CEBAS-CSIC Santomera, (referred as controller 2). This provides 18 input features given in Table 2 with 1125 records collected from 28th February to 7th March 2022 for training and 9th March 2022 to 15th march 2022 for testing.

3.2 Data Pre-Processing

As the data collected from the sensors are not classified under any of the categories: to develop a prediction model, the classification of variables under certain classes is mandatory. The raw data collected from the sensors are not labeled, we are using unsupervised learning methods as pre-processing technique, to cluster the data and generate a target groups. Such groups are considered clusters that help in analyzing the relation between variables and help in tracing the pattern of behavior.

Table 2: Input features provided by Controller 2.

Feature	Details
Nivel de Bacteria	level of Bacteria present
t-uC	Electric supply
Radiation Solar	Energy received from solar panel
Precipitation Rayos	Possibility of Rainfall
Direction Viento	wind direction
Temperature	Humid and temperature in ratio
Precision-Vapor	level of water vapour
xOrientation	Direction to X axis
yOrientation	Direction to Y axis
s1-0-counts4WVC	sensor 1 Water Volumetric content (WVC)
s1-0-temp	Sensor 1 temperature ratio
s1-0-ec	Sensor 1 electrical conductivity (EC) measurement
s2-0-counts4WVC	sensor 2 Water Volumetric content
s2-0-temp	Sensor 2 temperature count
s2-0-ec	Sensor 2 electrical conductivity (EC) measurement
Humid Relative	Ratio of relative Humidity
HST	Value of HST

3.2.1 Clustering

As part of dimensional reduction, and conversion of data from unsupervised learning to supervised learning method we are using a neural network-based clustering technique. Self Organizing Map (SOM), a deep learning feature reduction technique introduced by T. Kohonen (Kohonen, 1990). The SOM training phase inputs the elements and the mapping is used to classify the new input vector. A trained map classifies a vector from the input space with the closest weighted node. In the process of training "Euclidean Distance" is computed for all weight vectors and compared with other neurons in the space. The neurons with similar input are considered the Best Matching Unit (BMU) representing a lighter color in a visual display. Other neurons closer to this are adjusted with similar weights in the input vector.

1. Self Organizing Map (SOM) is used as a preprossessing step for supervised learning.
2. SOM represents clusters by grouping similar data. This reduces data dimensions and displays similarities among data.
3. The reduction of dimensionality and grid clustering makes it easy to observe similarities in the data for prediction.
4. SOMs factor in all the data in the input to generate these clusters and can be altered such that certain pieces of data have more/less of an effect on where input is placed.
5. Unlike other learning techniques in neural networks, training a SOM requires no target vector. A SOM learns to classify the training data without any external supervision.

Use of Clustering in the Study. In our study to predict the water requirement, variables like humidity, wind rate, VWC, and temperature form a cluster. And features like Bacteria level, precipitation of rainfall, temperature, wind direction, humidity, water control, and soil electrical conductivity (EC) as another cluster.

The above-mentioned group of features has an impact on predicting the water level or crop suitable for the land. This is an excellent indicator to trace the nutrient availability and loss, soil texture, and available water capacity. As the selected dataset does not have any classification variable, we have planned to implement the clustering technique to create multiple clusters which classify the variables into related groups using the SOM technique.

This helps in tracing the features which are closely related and identifying the dependency for tracing the target class.

SOM identifies a winning neuron i^* and updates the weights of all other neurons with a certain distance $n_{i^*}(d)$ using the Kohonen rule.

To implement the SOM we need to find out the best suitable cluster point, calculated from sums of points to central distance. We have implemented the K-means elbow point technique and pointed out that x:4, indicating four clusters are suitable for the selected dataset.

We have selected HEX-TOP as the topology function and LINK-DIST for the calculation of distance between the variables with 0.9 ordering phase learning rate and 1000 as ordering phase steps with 0.02 tuning phase learning rate and 1.0 neighborhood distance and (2X2) map dimension and with random weight and bias initialization. The model took 4 : 45 minutes to train the network, The SOM input planes for four clusters and 18 input features are displayed in Figure 2. The SOM technique divides the data into

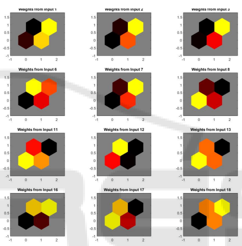

Figure 2: SOM input planes with cluster available.

cluster data points based on similarity. The dataset is represented by a special point, constructed by a learning method in the process of iteration. These points are mapped on a 2D grid to represent a topology of a honeycomb. The four clusters formed with similar points for our study are given below:

- cluster 1 -214
- cluster 2-113
- cluster 3-626
- cluster 4-171

This method indicates the movement of the neighborhood around the special point. Then these are arranged to form a cluster, of cells that are associated with the same macroscopic phenomenon. From Figure 2 variables having similar cluster points fall under the category of 2,4 clusters represented with black shade. A list of variable that fall under cluster 2,4 are given below:

- Input 4 - Precipitation of rain
- input 9 -x Orientation (wind movement)
- input 12 - s10temperature (sensor 1)
- input 15 - s20temperature (sensor 2)

From the above cluster grouping, we have grouped variables that have close association: the wind movement and the temperatures of both the sensors are clusters to show the dependency on precipitation variable. This cluster is used to represent the rainfall based on features 9,12,15 given above.

3.3 Classification

SOM trained the sample data and divided four similar groups. As the next step classification using Cascade Forward Back Propagation Neural Network(CFBPNN) a time series-based technique is trained and tested. The training and testing phase helps in tracing a behavioral pattern and analyzing the variation and then classifying the results. For example, as discussed in the above section 3.2.1, each cluster represents a target class for the prediction of rainfall, water requirement, or bacteria identification.

Examining the data and tracing the results based on time and checking if it is categorized under the same class is implemented using the CFBPNN classification technique.

Cascade Forward Back Propagation Neural Network (CFBPNN) is the combination of feed-forward with recurrent model. The proposed CFBPNN model begins with a single input and adds multiple connected layers, one by one in the process. Perceptions are added one by one in this correlation; it starts with a small number and ends up with a bigger size. Additional connections improve the speed and learning rate. We show the mathematical expression of the CFBPNN network in Equation 1.

$$y = \sum_{i=1}^{n} f^i w_i^i x_i + f^0 \left(\sum_{j=1}^{k} w_j^0 f_j^h \left(\sum_{i=1}^{n} w_{ji}^h x_i \right) \right). \quad (1)$$

In Equation 1, y represents the output layer, $\sum_{i=1}^{n}$ is used to calculate the sum of weights and bias of each layer. The special feature of this network is to carry forward the calculated weights and bias by establishing a direct relationship between the input and hidden layers using $f^i w_i^i x_i + f^0$. We use an activation function to train the complex patterns and take decisions for passing the values for the next layers. The internal working procedure of this model is explained in the Algorithm 1.

We provide the list of parameters used for training with the validation inputs below.

- Data Division: Random (70 training, 30 testing)
- Number of input layer: 18
- Number of hidden layers: 5
- Number of output layers: 4

Algorithm 1: CFBPNN operation.

1: Create a simple connected network with a input and output unit and Initialize y_i^j
2: Add the bias by increasing one by one unit $\sum_{i=1}^{n} \sum_{x}^{w} y_i^j$
3: Initialize R
4: Calculate: $f^0 \left(w^b + \sum_{j=1}^{k} w_j^0 f^h \left(w_j^b + \sum_{i=1}^{n} w_{ji}^h x_i \right) \right)$
5: **while** ($MSE == MSE_{threshold}$ OR the hidden unit is more than the given value) **do**
6: Add the linearly independent units to the network one by one.
7: Select the input unit and calculate the sum of weight $\sum_{i=1}^{n}$ from the beginning.
8: Add the bias $\sum_{x}^{w} y_i^j$ for each node, and calculate the weights by connecting each node.
9: Addition of hidden units till $(i - k < -N)$ and generate the target units.
10: Calculate MSE $\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y})^2$
11: **end while**
12: Calculate the weights $\sum_{x}^{w} y_i^j$ of output layer using back substitution method.

- Number of neurons in each layer: 10
- Training Function: Trainlm
- Maximum number of epochs: 1000

We use the Levenberg-Marquardt training function for all the network models. This is the fastest back-propagation algorithm that updates weight and bias. Comparatively, this method requires more memory.

4 RESULTS AND COMPARISON

4.1 Evaluation Metrics

A confusion matrix is the most appropriate technique to analyze the performance of the classification model. This helps to indicate the true and false classified cases in the tested model. We have considered false rate and overall accuracy for evaluating the model performance given below with mathematical expressions.

False Negative Rate (FNR): Miss classified to a selected class, calculated with the Equation 2.

$$FNR = \frac{FN}{TP + FN}. \quad (2)$$

Accuracy (A): The ratio of correctness for classified samples to right class represented with the Equation 3.

$$Accuracy(A) = \frac{TP + TN}{TP + FP + FN + TN}. \quad (3)$$

4.2 Results of Prediction Model

Our proposed CFBPNN model is used for mapping the patterns between input and target values. Various compositions of threshold functions are used in the layers with multiple combinations. We have trained and tested the model for the controller 2 data and projected the results in Figure 3.

To find the efficiency of the model and identify the water level and bacteria effect, we have analyzed the performance of the proposed model using a confusion matrix. This classifies and denotes the accuracy ratio and the false rate for each cluster. A multi-class testing is used to identify the relation of each variable with the cluster; This analysis shows that our model provides a 98.9% high accuracy rate and very minimal false rates with 1.2% for each test cluster given in Figure 3.

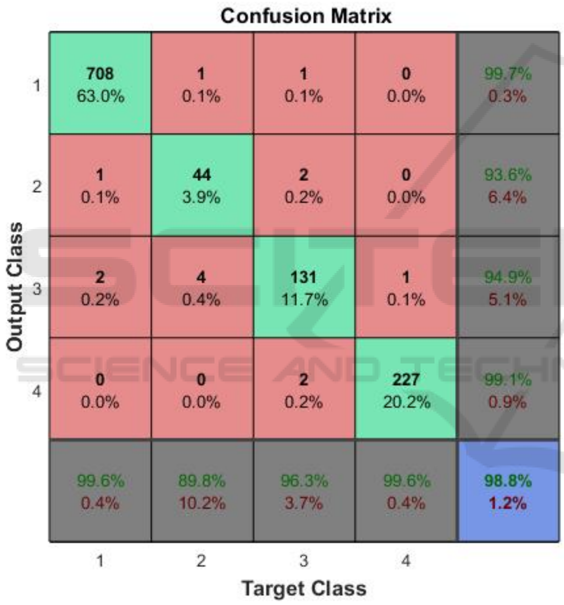

Figure 3: Confusion matrix of CFBPNN prediction model.

4.3 Variables Effecting Bacteria Levels

To identify the effect of bacteria, on the plant we have considered four input variables of cluster 1 which have a high impact on the existence of bacteria content. A co-relation plot is used to display the relationship between all the variables of cluster 1 and the bacteria as the target variable given in Figure 4.

According to Figure 4, it is to be observed that variable tuC(electricity level) solar Radiation and Temperature have a high impact on the creation and spreading of the bacteria. A notable point solar radiation with a 0.68 ratio, has been a major cause of bacteria development in the plant.

4.4 Variables Effecting Water Levels

Analysing the water requirement provide limited water, to preserve water wastage is the huge task. Considering this we have examined a cluster which influence more to predict the water requirement is given in Figure 5. Variable WVC give the count of Volumetric Water Content which helps in testing the soil moister, requirement of water level , and the atmosphere condition. This has a high impact with electric control variable with 0.88 ratio.

4.5 Comparative Analysis

Focusing the crucial requirement of IoT agriculture, as to predict the water requirement ratio, best irrigation period, suitable temperature, soil moisture levels, and best suitable crop and many more. Some of the latest proposals using machine and deep learning models, for various levels of prediction are discussed in Table 3. Crop prediction by (Varman et al., 2017) gained minimum loss with 2.135 with 57.65s a minimum training time compared to all the proposed models. Soil moisture prediction for providing best nutrition for the plant by (Goap et al., 2018), using SVR and K-means technique resulted with 96% R-square value. Weather prediction with MLR gained 99.05 accuracy in identifying the apt temperature need to grow the plant (Parashar, 2019). The combination of multiple parameters as soil-moisture, air-humidity, air-temperature using deep-learning model resulted with least error rates (Dahane et al., 2020). Other prediction model trained using LSTM resulted with RMSE 3.0 for temperature prediction and 14.55 for humidity prediction to select a plant for irrigation (Jin et al., 2020).

Capturing live image and detect the disease, using SVR, and also predict the weather to trigger action for pest control by (Sasi Supritha Devi et al., 2020). The study focused on image processing with K-Means clustering approach and resulting with android application for monitoring the irrigation process. Classification and quantitative predictions for various parameters as soil type, crop type and amount of irrigation required using SVM (Support Vector Machine), SVR (Support Vector Regression) and Random forest 81.6% accuracy (Vij et al., 2020). Prediction of Volumetric soil moisture by (Kashyap et al., 2021) gained 0.012 RMSE for Sutlej Basin plant-Ludhiana(Punjab). The model proves to be more accurate as the RMSE value is close to zero. They also explored various method of prediction amount of water saved and irrigation period by controlling the functionality of the irrigation scheduler. Crop prediction

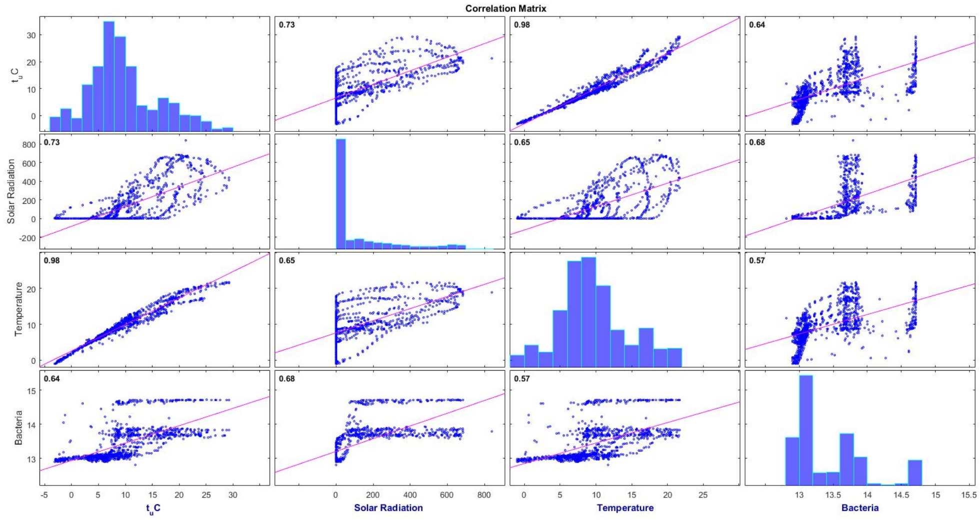

Figure 4: Variables effecting bacteria levels.

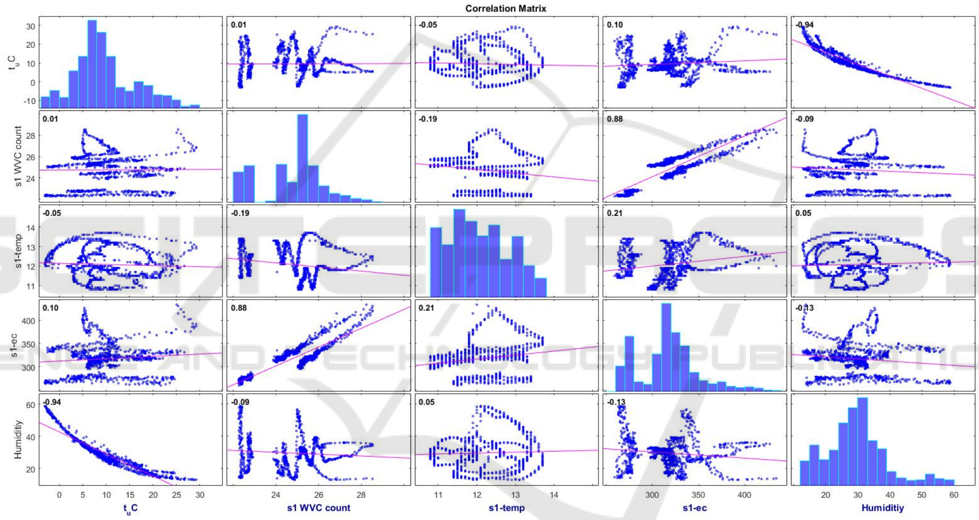

Figure 5: Variables effecting water levels.

based on temperature and rainfall using multi-layer perceptron a rules-based classifier and JRip decision table classifier by (Bakthavatchalam et al., 2022). The model resulted with 98.2% accuracy in prediction with 8.05s training time.

Out of all the above state of models, proposed model gave highest accuracy comparatively. The model has a unique technique for clustering and classifying the resulted clusters, to predict the feature set. All the given models are only used to predict a specific category, our model divide set of features to create a universal clusters and predict parameters based on requirements.

5 CONCLUSION

Research committee all over is striving hard to enhance agriculture productivity by deploying services provide by IoT technologies. In this paper, we have discussed a platform and topology which helps to access sensor-based data and facilitates farmers to take the right decision to enhance crop productivity. In this study we have proposed unsupervised clustering method SOM to form similar variable groups.

In addition, this paper provides an novel technique CFBPNN to predict the water requirement and resulted with 98.8% accuracy in classification. An overview on current technologies, models and frameworks contributed towards smart irrigation is

Table 3: Comparative analysis.

Reference	Objective	Model	Results
(Varman et al., 2017)	Predict best-suited crop	Long Short Term Memory (LSTM)	Validation Loss= 2.1354
(Goap et al., 2018)	Predict soil moisture	Support Vector Regression (SVR) ,K-means	R-squared=96%
(Parashar, 2019)	Predict the weather	Multiple Linear Regression (MLR)	Accuracy=99.05%
(Dahane et al., 2020)	Multi parameter prediction	LSTM, GRU-based model	MSE=0.02
(Jin et al., 2020)	Prediction of temperature and humidity	Gated Recurrent Unit (GRU), LSTM	RMSE=3.00
(Vij et al., 2020)	Multi parameter prediction	SVM (Support Vector Machine), SVR (Support Vector Regression)and Random forest	Accuracy=81.6%
(Kashyap et al., 2021)	Prediction of soil moisture content	long short-term memory network	RMSE= 0.012
(Rezk et al., 2021)	Crop productivity and drought predictions	wrapper feature selection, and PART classification technique	Accuracy: 98.15%.
(Bakthavatchalam et al., 2022)	Crop prediction	DL techniques	Accuracy= 98.2%
Proposed model	Clustering, classification and prediction	SOM and CFBPNN	Accuracy= 98.8%

provided. This research explore various challenges and and requirements for the better understanding of smart farming. Furthermore, features playing important role in predictions are visualized using corelation map. Finally, this study provide challenges faced by agriculture industry and prose smart method to handle it.

ACKNOWLEDGEMENTS

The work is carried out in the frame of the PRECIMED project that is funded under the PRIMA Programme. PRIMA is an Art.185 initiative supported and co-funded under Horizon 2020, the European Union's Programme for Research and Innovation. (project application number: 155331/I4/19.09.18).

REFERENCES

Akhter, R. and Sofi, S. A. (2022). Precision agriculture using iot data analytics and machine learning. *Journal of King Saud University-Computer and Information Sciences*, 34(8):5602–5618.

Ayaz, M., Ammad-Uddin, M., Sharif, Z., Mansour, A., and Aggoune, E.-H. M. (2019). Internet-of-things (iot)-based smart agriculture: Toward making the fields talk. *IEEE access*, 7:129551–129583.

Bakthavatchalam, K., Karthik, B., Thiruvengadam, V., Muthal, S., Jose, D., Kotecha, K., and Varadarajan, V. (2022). Iot framework for measurement and precision agriculture: predicting the crop using machine learning algorithms. *Technologies*, 10(1):13.

Ben Abdallah, E., Grati, R., and Boukadi, K. (2023). Towards an explainable irrigation scheduling approach by predicting soil moisture and evapotranspiration via multi-target regression. *Journal of Ambient Intelligence and Smart Environments*, (Preprint):1–22.

Dahane, A., Benameur, R., Kechar, B., and Benyamina, A. (2020). An iot based smart farming system using machine learning. In *2020 International symposium on networks, computers and communications (ISNCC)*, pages 1–6. IEEE.

Geetha Lekshmy, V., Vishnu, P., and Harikrishnan, P. (2022). Adaptive iot system for precision agriculture. In *Inventive Computation and Information Technologies*, pages 39–49. Springer.

Goap, A., Sharma, D., Shukla, A. K., and Krishna, C. R. (2018). An iot based smart irrigation management system using machine learning and open source technologies. *Computers and electronics in agriculture*, 155:41–49.

Jimenez, A.-F., Ortiz, B. V., Bondesan, L., Morata, G., and Damianidis, D. (2021). Long short-term memory neural network for irrigation management: a case study from southern alabama, usa. *Precision Agriculture*, 22(2):475–492.

Jin, X.-B., Yu, X.-H., Wang, X.-Y., Bai, Y.-T., Su, T.-L., and Kong, J.-L. (2020). Deep learning predictor for sustainable precision agriculture based on internet of things system. *Sustainability*, 12(4):1433.

Kashyap, P. K., Kumar, S., Jaiswal, A., Prasad, M., and Gandomi, A. H. (2021). Towards precision agriculture: Iot-enabled intelligent irrigation systems using deep learning neural network. *IEEE Sensors Journal*, 21(16):17479–17491.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480.

Lavanya, G., Rani, C., and GaneshKumar, P. (2020). An automated low cost iot based fertilizer intimation system for smart agriculture. *Sustainable Computing: Informatics and Systems*, 28:100300.

Mehra, M., Saxena, S., Sankaranarayanan, S., Tom, R. J., and Veeramanikandan, M. (2018). Iot based hydroponics system using deep neural networks. *Computers and electronics in agriculture*, 155:473–486.

Nisha, G. and Megala, J. (2014). Wireless sensor network based automated irrigation and crop field monitoring system. In *2014 Sixth international conference on advanced computing (IcoAC)*, pages 189–194. IEEE.

Ohana-Levi, N., Ben-Gal, A., Peeters, A., Termin, D., Linker, R., Baram, S., Raveh, E., and Paz-Kagan, T. (2021). A comparison between spatial clustering models for determining n-fertilization management zones in orchards. *Precision Agriculture*, 22(1):99–123.

Parashar, A. (2019). Iot based automated weather report generation and prediction using machine learning. In *2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT)*, pages 339–344. IEEE.

Pathan, M., Patel, N., Yagnik, H., and Shah, M. (2020). Artificial cognition for applications in smart agriculture: A comprehensive review. *Artificial Intelligence in Agriculture*, 4:81–95.

Prabhu, B., Sophia, S., and Mathew, A. (2014). A review of efficient information delivery and clustering for drip irrigation management using wsn. *A., A Review of Efficient Information Delivery and Clustering for Drip Irrigation Management Using WSN*.

PRIMA (2023). PRECIMED - Precision Irrigation Management to Improve Water and Nutrient Use Efficiency in the Mediterranean Region. https://mel.cgiar.org/projects/precimed. (Accessed on 02/26/2023).

Reddy, K. S. P., Roopa, Y. M., LN, K. R., and Nandan, N. S. (2020). Iot based smart agriculture using machine learning. In *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, pages 130–134. IEEE.

Rezk, N. G., Hemdan, E. E.-D., Attia, A.-F., El-Sayed, A., and El-Rashidy, M. A. (2021). An efficient iot based smart farming system using machine learning algorithms. *Multimedia Tools and Applications*, 80:773–797.

Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, 2(6):420.

Sasi Supritha Devi, Y., Kesava Durga Prasad, T., Saladi, K., and Nandan, D. (2020). Analysis of precision agriculture technique by using machine learning and iot. In *Soft Computing: Theories and Applications: Proceedings of SoCTA 2019*, pages 859–867. Springer.

Swamynathan, M. (2019). *Mastering machine learning with python in six steps: A practical implementation guide to predictive data analytics using python*. Apress.

Varghese, R. and Sharma, S. (2018). Affordable smart farming using iot and machine learning. In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 645–650. IEEE.

Varman, S. A. M., Baskaran, A. R., Aravindh, S., and Prabhu, E. (2017). Deep learning and iot for smart agriculture using wsn. In *2017 ieee international conference on computational intelligence and computing research (ICCIC)*, pages 1–6. IEEE.

Vij, A., Vijendra, S., Jain, A., Bajaj, S., Bassi, A., and Sharma, A. (2020). Iot and machine learning approaches for automation of farm irrigation system. *Procedia Computer Science*, 167:1250–1257.