

Classification and Yield Prediction in Smart Agriculture System Using IoT

Akanksha Gupta (✉ agupta.du@yahoo.com)

University of Delhi <https://orcid.org/0000-0001-7253-3648>

Priyank Nahar

Venkateshwara Institute of Management

Research Article

Keywords: IoT, Sensors, Agriculture, Machine learning, Classification, Crop yield prediction.

Posted Date: February 21st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-830669/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Journal of Ambient Intelligence and Humanized Computing on January 5th, 2022. See the published version at <https://doi.org/10.1007/s12652-021-03685-w>.

Classification and Yield Prediction in Smart Agriculture System Using IoT

Ms. Akanksha Gupta*, Department of computer science, University of Delhi
Email: akankshagupta@ss.du.ac.in

Dr. Priyank Nahar, Shri Venkateshwara University, Gajraula, India.
Email : priyank.nahar@gmail.com

Abstract

The Modern agriculture industry is data-centred, precise and smarter than ever. Advanced development of Internet-of-Things (IoT) based systems redesigned “smart agriculture”. This emergence in innovative farming systems is gradually enhancing the crop yield, reduces irrigation wastages and making it more profitable. Machine learning (ML) methods achieve the requirement of scaling the learning performance of the model. This paper introduces a hybrid ML model with IoT for yield prediction. This work involves three phases : pre-processing, feature selection(FS) and classification. Initially, the dataset is pre-processed and FS is done on the basis of Correlation based FS (CBFS) and the Variance Inflation Factor algorithm (VIF). Finally, a two-tier ML model is proposed for IoT based smart agriculture system. In the first tier, the Adaptive k-Nearest Centroid Neighbour Classifier (aKNCN) model is proposed to estimate the soil quality and classify the soil samples into different classes based on the input soil properties. In the second tier, the crop yield is predicted using the Extreme Learning Machine algorithm (ELM). In the optimized strategy, the weights are updated using modified Butterfly Optimization algorithm (mBOA) to improve the performance accuracy of ELM with minimum error values. PYTHON is the implementation tool for evaluating the proposed system. Soil dataset is utilized for performance

evaluation of the proposed prediction model. Various metrics are considered for the performance evaluation such as accuracy, RMSE, R2, MSE, MedAE, MAE, MSLE, MAPE and Explained Variance Score (EVS).

Keywords: *IoT, Sensors, Agriculture, Machine learning, Classification, Crop yield prediction.*

Declarations:

Funding - Not applicable.

Conflicts of interest/Competing interests (include appropriate disclosures) - Not applicable

Availability of data and material - Available at soilhealth.dac.gov.in

Code availability - Custom code

Ethics approval - Not applicable

Consent to participate - The participant has consented to the submission of the case report to the journal

Consent for publication - I give my consent for the publication of identifiable details, which can include photograph(s) and/or videos and/or case history and/or details within the text to be published in the Journal.

1. Introduction

IoT is an advanced technology for monitoring and controlling devices anywhere in the world. In many fields, it creates a remarkable mark due to its easy accessibility [1]. Few IoT developed technologies such as remote sensors, drones and robots made human's life simpler and beneficial. Moreover, these technologies have experimented on fundamental needs such as food that is obtained from the agricultural field [2, 3]. From the recent survey of World Bank, it is approximated that more than 50% food is required to cultivate before 2050 based on present population rate [4]. However, such huge production of the crop is a challenging task because of the current climatic changes. In such cases, Smart agriculture system plays a vital role to increase the yield by monitoring and predicting the production of the crops [5]. IoT based crop yield prediction enables the farmers to enhance productivity. In general, IoT based smart farming system is deployed in an agriculture field for monitoring the crop field with the help of sensors namely DHT11 (temperature and humidity sensor), TOC (Total Organic Carbon) and nitrogen, phosphorus, and potassium (NPK) sensors [6]. Using this setup, farmers can monitor the field conditions from anywhere. Gateways are responsible for receiving data from the crop area and forward them to the storage unit. The prediction engine is used to predict the results and sends information to the notification server [7]. Therefore, various IoT based techniques can assist the farmers to produce huge crops and the techniques used are sensors, hydroponic farming, advanced tractors and drones [8]. These devices assist agriculturists and researchers in analyzing and processing

the data remotely for decision making. The main advantage is that these techniques are low cost; hence it is affordable for all farmers [9].

Agricultural supervision, particularly crop yield observation is essential for examining the food security in a region. Due to several difficult aspects, predicting the crop yield manually is a challenging task [10]. Based on the water quality as well as availability, pest infestations, genotype, landscape, soil quality, climatic condition, etc., the yield of the crop may vary. The strategies and the processes are non-linear in nature, intricate and varied with time because of external aspects and correlated factors [11]. Recently, several studies illustrate that ML approaches such as support vector regression, multilayer perceptron (MLP), etc. have comparatively more enhanced potential than the traditional techniques. These approaches have the ability to predict linear and non-linear agricultural architecture. From the learning process, these methods were obtained in ML agricultural framework [12-14].

On observing the most outstanding frameworks in agriculture, artificial and deep neural networks are the commonly utilized models [15]. Artificial neural network (ANN) is a network model that generates approximation by bias as well as weight optimization for a node-link structure including input, hidden and output layers [16]. Deep learning (DL) is a subgroup of ML used to predict the crop yield based on the varying arrangement of raw data via the intensive learning process in a deep network [17]. Moreover, these DL algorithms have the capability to design a probability model using field

data. Along with this benefit, the data about the crop performance under different climatic changes are provided by DL approaches [18]. For example, Reinforcement learning is one of the major areas of artificial intelligence. It is the preparation of ML models for decision-making sequences and is the significant class of algorithm that is used to streamline logic for dynamic programming [19]. Besides, extreme learning machine (ELM) is also a ML approach that has the capability to empower neural network training for predicting the crop yield. It accelerates the learning process and provided better outcomes. However, these approaches have several disadvantages like less sustainability, computationally expensive, high complexity and false prediction [20]. To overcome these challenges, a more efficient ML based crop yield prediction model is proposed in this work.

Motivation:

Agriculture is the major economic resource of India. To overcome the issues of high cost and complex management of conventional agricultural planting, IoT is applied for realizing real time detection, crop growth intelligent management and changing the conventional agricultural planting mode. Various mathematical and empirical yield approaches have been evaluated for several crops. These models need huge knowledge about soil and crops which make it hard for implementing for various localities. Many satellite based remote sensing methods were also developed in yield modelling. But these approaches are not able to provide enough spatial details of small farms for optimizing crops. Recent developments in

machine learning models have become popular that enable researchers for solving and understanding complex predictions. Many ML approaches have been employed for crop prediction like decision trees, ANN and Support vector machine (SVM) and in this research work ELM and aKNCN are used for crop yield prediction thereby achieving better results.

Contributions:

The proposed work's major contributions are:

- The proposed work provides an IoT based farming system that ensures the deployment of effective crop yield prediction model. This work involves pre-processing, FS and classification. The data is pre-processed and features are selected by FS algorithms.
- Then IoT based smart agriculture system using two-tier ML model is proposed for the better prediction of crop yield.
- ML based classification is performed to classify the soil samples for different classes by considering the properties of soil dataset.
- ELM model is proposed for predicting the crop yield and the weights of the model are updated using mBOA algorithm for the performance improvement of the prediction system with less error values.
- The systematic evaluation of the proposed crop yield prediction system is implemented based on machine learning performances through various experiment consequences.

Paper outline: Section 1 presents the introduction and highlights smart agriculture in brief. Recent related works during (2019-2021) are discussed in section 2. Section 3 focuses on the problems in the IoT based crop yield prediction. Section 4 provides the detailed description about IoT devices used in a smart agriculture system. Section 5 presents the proposed framework along with important measurables. Section 6 discusses the experimental analysis and results. Section 7 concludes the presented work.

2. Related works

Abbas, *et al.* [21] predicted the crop yield via proximal sensing and the ML algorithms. The objective was to extract significant data that are responsible for controlling the yield of crop. The properties of potato tuber crop and the data of soil have been gathered by proximal sensing. A large dataset was utilized for the prediction performance. Support vector regression (SVR), k-nearest neighbour (K-NN), linear regression and elastic net ML algorithms were utilized for the classification and prediction of crop yield. R^2 , MAE and RMSE were determined for the performance evaluation. The performance achieved for KNN was poor in crop yield prediction because of higher number of features.

Rezk, *et al.* [22] presented an IoT based smart agriculture system using ML algorithm. The drought and the crop productivity were predicted by WPART and it was a combination of wrapper and PART techniques. Feature selection and classification were the two important phases in the prediction process. Wrapper feature selection technique selected the

optimal features for further classification. PART was a partial decision tree approach used for classification and prediction. Accuracy, precision, sensitivity and F1 score were considered for the experiment of WPART. The crops taken for experiment were Sugarcane, Jowar, Bajra and Soybean. Some samples in the dataset were misleadingly labelled, thus the false prediction rate was high. In future, some parameters like soil nutrients, agricultural inputs, irrigated area and soil quality will be considered for forecasting the crop yield.

Bu, *et al.* [23] developed deep reinforcement learning based ML technique for smart farming IoT system. Cloud computing and artificial intelligence were combined for the classification and prediction of crop yield. The key goal of this research was to minimize resource consumption and maximize the food production. A hierarchical Bayesian based multi-task reinforcement learning method has been utilized for modelling the Markov decision process. Then, the Q-value regression function was examined using policy distillation. However, computational complexity was considered as one of the major drawback of this approach. Also, human-level performance was not achieved in complex task solution and in adaptation to dynamic environments. In future, this research aims to design an incremental model and transfer learning approaches for the enhancement of performance efficiency.

Nevavuori, *et al.* [24] proposed a deep learning technique for crop yield prediction. The key objectives of this research were crop yield prediction, biomass evaluation, crop and weed

detection. Convolutional Neural Network (CNN) was modelled for extracting the features, training, hyperparameter tuning and regularization to predict the yield of wheat and barley crops. MAE and MAPE are the evaluation metrics used for simulation analysis. But, the presented CNN does not perform well for the large dataset. Also, the performance efficiency of this method was not good. The future scope of this research is to train the developed model for a large set of features like soil and climate with time series image data.

Dos Santos, *et al.* [25] introduced AgriPrediction model for IoT based smart agriculture system. It was an end-to-end model that predicted agricultural crops. It was the integration of prediction as well as short and medium wireless network range system. The components of AgriPrediction model have been designed according to the ARIMA prediction model and LoRa IoT technology. Initially, the data were gathered using sensors, then the discrete moving average-based prediction has been performed. If the predicted crop goes wrong, then the notification was given to the farmer's mobile phone. This model was computationally expensive and less sustainable. Moreover, the accuracy of this AgriPrediction model was not evaluated in this research. The future scope of this research is to generate a mobile application for crop's real-time monitoring.

Saranya, C. P., and N. Nagarajan [26] presented a neural network with population based incremental learning (NN-PBIL) method for the prediction performance enhancement. The neural network was used to classify and predict the crop yield. The weight of the neural

network was updated by the PBIL approach. Hadoop framework has been utilized for the prediction performance. Neural network along with ANN and multiple linear regressions (MLR) were implemented for the crop yield prediction. Low convergence and getting stuck within local minimum were the major drawbacks in this presented model. The future scope of this research is to use the optimization approaches for the crop yield forecasting.

Filippi, *et al.* [27] proposed the empirical modelling scheme for forecasting the yield of barley, wheat and canola crops. In this research, several fields are considered for the prediction performance instead of single field in isolation. Random forest models and publically available data with temporal and spatial data collected on-farm were combined for the yield prediction of canola, barley and wheat. The experimental results showed that the accuracy obtained by this predictive model was low. In future, this research will be extended by exploring enhanced feature extraction calibration and more publically available data sources for forecasting the yield of the crop.

Sun, Jie, *et al.* [28] predicted the yield of both in-season and end-of-season soybean using deep CNN-LSTM based on remote sensing data. The training data such as MODIS surface Reflectance (SR) data, MODIS land surface temperature (LST) data and weather data were correlated and transformed to histogram based tensors according to the Google earth engine (GEE). The performance of crop yield prediction at large scale was not evaluated in this research. Time and computational complexity was high and fed the raw remote sensing data into DL mode was a

complex task. The future scope of this research is to include more features for the yield prediction and the performance will be performed at large scale.

Sinwar, *et al.* [29] discussed about different methods of Artificial Intelligence (AI) for smart irrigation and crop yield prediction system. From the research it was proved that AI based system offered adequate data regarding the crop yields at an early stage. Several ML techniques such as Support vector machine (SVM), Linear Discriminant Analysis (LDA), ANN, Generative Adversarial network (GAN), Deep Boltzmann machine (DBN), K nearest neighbour (KNN), Backpropagation neural network (BNN) and Deep neural network (DNN) performed better in crop yield prediction. However, some limitations are there in these approaches like computational complexity, high cost, dependencies between target and input variables, proper model representation and accuracy affected by data quality.

3. Problem Statement

The method in [21] attained better performance but in some cases it achieves poor result due to the small datasets. The method [22] outperformed conventional model for 5 datasets for classification of drought and productivity of crop but it did not focus on time series analysis. Deep reinforcement [23] showed better growth in model design but it takes more time for training. CNN [24] achieves better accuracy on yield prediction but does not support large features like soil and climate. The method [25] achieves better result in

online prediction model in agri prediction but it has high computational complexity. ANN [26] achieves good prediction in crop yield but handling noise in the images has to be investigated. The method [27] achieved higher accuracy but it has to be focussed on time series prediction. Deep CNN-LSTM [28] model is highly efficient in error performance at any time node but the processing time is large due to its complex structure. AI [29] minimizes the effort of human and improves agricultural practices but the system cost is too high. Due to these drawbacks the proposed model introduces 2 classification models with optimal weight selection for crop yield prediction.

4. IoT Devices

The IoT based smart agriculture model is designed to construct the whole crop prediction system. The key goal of this model is to predict the crop yield based on the gathered data using IoT devices. It assists the crop yield prediction system for better decision making on the yield of crop according to the gathered data from monitoring camera analysis. The data obtained from camera is integrated with environment data from IoT devices. Availability of environment data as well as its timely delivery is significant for preserving the crops and property at the time of disasters. The important information considered for an effective predictive model are Statistical agriculture data and contemporary IoT sensing data. The parameters available in statistical and contemporary IoT sensing data are rainfall, temperature, pH level, soil nutrients and fertilizers. Figure 1 represents the structure of IoT services for agriculture.

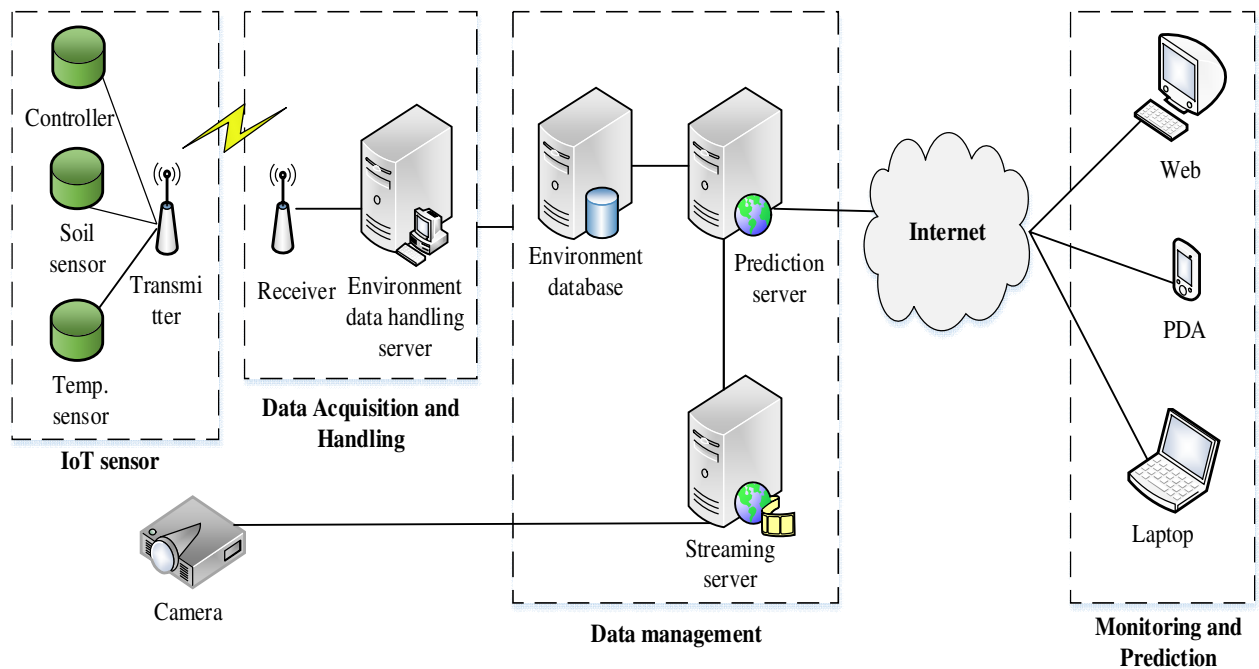


Figure 1: Service of IoT for Agriculture

The major components in IoT based agriculture model are sensors, data acquisition, data management devices, monitoring and prediction devices. These devices enhanced the performance of the crop yield prediction. Smart farm

embedded with IoT system can improve the production of crop by predicting the soil type, climatic changes, soil quality and several other factors. Figure 2 illustrates the example picture representation of IoT based smart agriculture system.

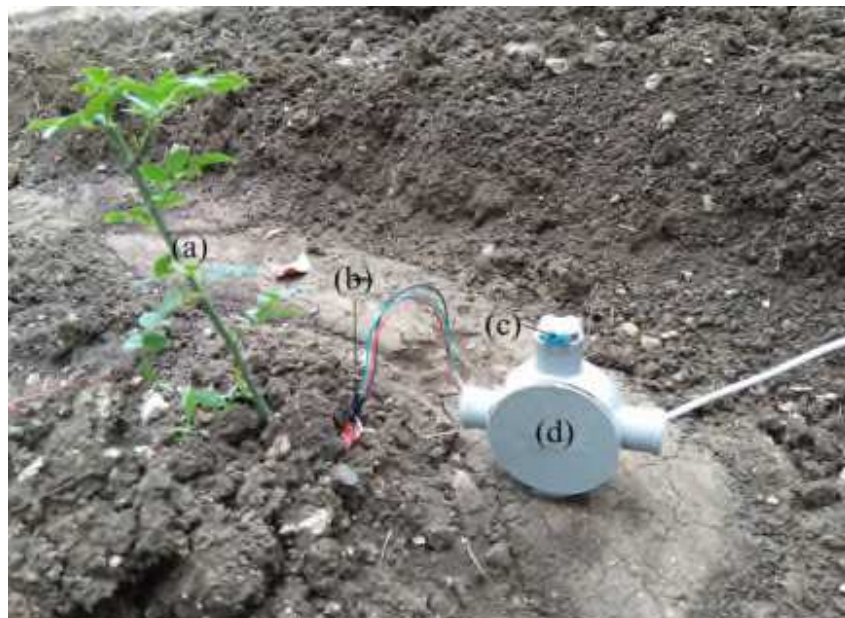


Figure 2: Example pictorial representation of IoT based smart agriculture system:

(a) plant (b) moisture sensor (c) temperature and humidity sensor and (d) USP

Sensors:

Sensors in IoT are used to track the particular soil and water pH value, temperature and humidity value and fertilizer control for the growth of crop [32]. The moisture, temperature and humidity and pH sensors are discussed in the following subsections.

(i) Soil moisture sensor:

It is a sensor that determines the water content available in soil. Several sensors are there in a probe, which save the water and manage irrigation system economically and successfully. Also, the quality and yield of the crop is improved using these sensors. It provides the accurate outcomes immediately and is less expensive. Figure 3 shows the pictorial representation of a soil moisture sensor.



Figure 3: Soil moisture sensor

Using this sensor, the soil's water content can be examined and it averages the water content over the complete span of the sensor. The main purpose of this sensor is to monitor the moisture of the soil for irrigation management, examine the loss of humidity over time because of plant uptake and evaporation and to find the optimum soil moisture stuffing. The sensor readings are then transferred to the transmitter and the threshold value is set in the application.

The water motor in the field is operated based on the threshold value, i.e., if the threshold value is greater than the water level, then the needed water is supplied. Once the water level is attained, the sensor reading provides the indication and the irrigation is stopped.

(ii) Temperature and Humidity sensor:

This is used to measure pending rain, temperature changes, rain fallen at a particular period and water consumption. It is comprised of sense of wet NTC temperature measuring devices and resistive element. This sensor provides fast response, high cost performance, incredible quality and anti-interference capability. Figure 4 illustrates the pictorial representation of a temperature and humidity sensor.

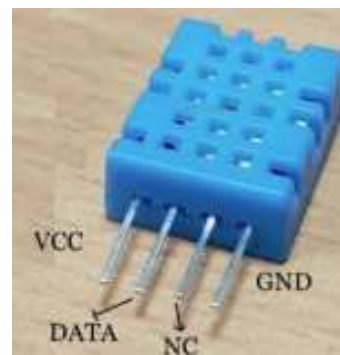


Figure 4: Temperature and humidity sensor

(iii) pH sensor:

pH is an indicator of a solution's acidity and alkalinity. Normally, the range of pH level is considered between 0 and 14. Moreover, it represents the concentration in certain hydrogen +ions solutions. Figure 5 depicts the pictorial representation of a pH sensor. The potential difference between two electrodes such as hydrogen-sensitive glass electrode and reference electrode is detected by this sensor. It can

be used with the microcontroller like Arduino.



Figure 5: pH sensor

Controller:

Arduino controller is used to monitor and control the farm's environmental conditions. The inputs provided to the controller are location from the farm and the pH value. Besides that, the percentage of nutrients such as Calcium, Zinc, Manganese, Organic matters, Boron, Magnesium, Iron, Sulphur, Nitrogen, Potassium and Phosphorous. Based on the pretrained network or real time values or both, the decisions regarding the prediction are made by the controller. Figure 6 indicates the arduino controller.



Figure 6: Arduino controller

Data acquisition:

The initial process in data acquisition is to register each device deployed in the farm into Mobius. The registration is done using &cube device and it acts as an interface between the deployed devices and the Mobius. Then, the virtual representation is generated for each device based on the resource type. The environmental data gathered by sensors are transmitted to &cube and the &cube forwards the data to Mobius [33]. Finally, the available virtual representations of the devices are accessed by the farmers or end-users to monitor and manage their connected farm with the help of IoT applications such as tablets, laptops, smartphone, etc.

(i) Mobius:

It is an open IoT service platform that performs operations based on one machine-to-machine (M2M) standard rules. The generation of physical IoT devices' virtual representations is assisted by this Mobius. It is designed to conform with one M2M specifications and helps general M2M/IoT service functions like security, device registration, subscription as well as notification and data repository and management. Moreover, the data resources preserved in Mobius are accessed and the IoT devices are controlled by REST APIs.

(ii) &cube

The &cube is a device software platform that is installed into IoT gateways. Through the standard REST APIs, the gathered data from physical devices is transmitted to Mobius. Several protocols are assisted by &cube such as CoAP, MQTT and HTTP. Raspberry-Pi is utilized in this work for the connected farm. It is a single-board Linux installed computer.



Figure 7: Raspberry-Pi

Data Management:

In data management service, the required external data are gathered from the sensors and the servers preserve the collected data for further process. Also, the data gathered from camera is stored to predict the yield of the crop. Storage, maintenance and management of data for maintenance of corresponding services are provided in this data management system.

5. Proposed Framework

The ingredients of soil like Phosphorous, Potassium and Nitrogen, crop rotation and

atmospheric temperature etc. play a vital role in cultivation. ML methods are an essential decision support device for the prediction of crop yield like supporting decision on what crops to grow. Many ML algorithms are employed to support the prediction of crop yield. In the proposed smart agricultural framework, pre-processing, FS and a two-tier model is implemented for crop yield prediction. Here, akNCN is proposed to be deployed which is an improved version of KNCN. Generally ELM is an influential model with more fast learning methods, higher performance and less training error when compared with other algorithms. Therefore these two classification algorithms are proposed in this work to improve the accuracy of the system and provide better results than the existing models.

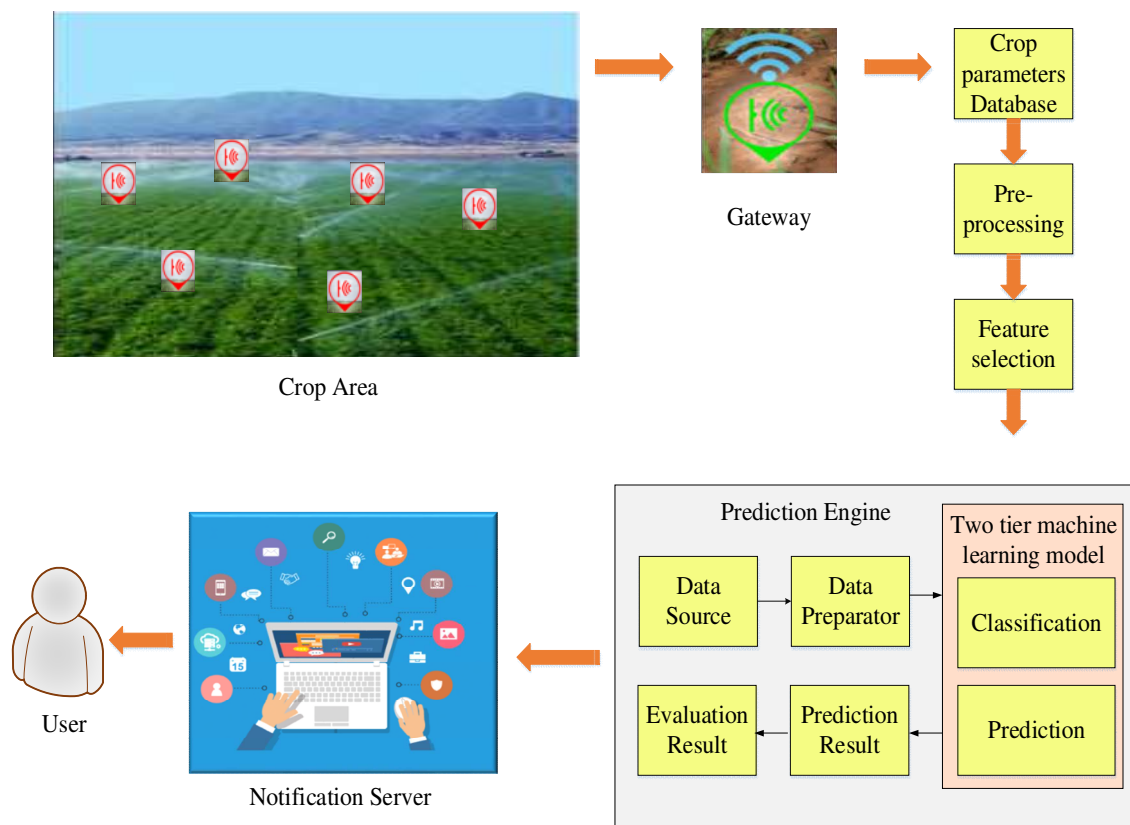


Figure 8: Architecture of the proposed methodology

Figure 8 illustrates the architecture of the proposed crop yield prediction model. Initially, Pre-processing is done to remove the noise in data and the features are selected on the basis of features selection methods like CBFA and VIF. Finally, the classification uses two tier systems. In the first tier, the proposed aKNCN model is used to classify the soil quality based on IoT system collected soil nutrients. Then in the second tier, ELM-mBOA is utilized for crop yield prediction and the accuracy is improved by optimal weight selection using mBOA. This model improves the accuracy of the system with minimum error values.

5.1 Pre-processing

The data is gathered from various sources and pre-processing is done. Pre-processing is a necessary phase in ML since it can't handle noisy data. Noisy data means it has errors and outliers. Before applying the data to classification it has to be pre-processed for inserting missing values, eliminate unwanted data, functionality extraction and maintain the appropriate data range. In this work isnull() approach is used to check the null values then the label encoder() is used to convert categorical data (string format) into numerical data (numeric format). Since Python does not handle categorical data it must be converted into numeric format. Once the data is converted to numeric format, it is applied for feature selection.

5.2 Feature selection(FS)

ML is a computational learning model that works on prediction from statistical value. FS model is applied to identify necessary features which are powerful in correlation with crop production. The main reason to employ FS is that it enables the ML algorithm to train faster, minimizes the model complexity and makes it easy to interpret. It also increases the system accuracy when the proper subset is selected and reduces overfitting. The computation time of the algorithm is less necessary than its classification for normal size feature sets. But the feature selection is necessary for large datasets. Various statistical approaches can be employed in FS like filter, embedded and wrapper methods. Filter methods choose the intrinsic characteristics of the features computed by univariate statistics instead of performance of cross-validation. These methods are faster and less computationally expensive than wrapper methods. When dealing with high-dimensional data, it is computationally cheaper to use filter methods. Hence, in this work filter based FS methods like CBFA algorithm and VIF algorithm are used. CBFA chooses the best feature set which is mainly correlated with yield. VIF verifies the multicollinearity among independent features. Therefore, it eliminates all multicollinear features.

5.2.1 Correlation based Feature selection Algorithm (CBFA)

CBFS orders feature subset based on the correlation heuristic evaluation function. This function is towards a subset that has features which have high correlation

amongst class and uncorrelated with each other. The features which are not relevant must be removed since they have less correlation amongst class and high correlation with other features. Feature acceptance is based on the extent to which it identifies classes in areas which are already not identified by other features. The CBFS is computed as:

$$M = \frac{Nr\bar{c}}{\sqrt{N + N(N-1)r\bar{f}}} \quad (1)$$

Where N is the total number of features, $r\bar{c}$ is the average correlation, $r\bar{f}$ is average feature to pair wise correlation.

5.2.2 Variance Inflation Factor algorithm (VIF)

VIF computes the strength of the multi colinearity in the analysis of least squares regression. It gives an index that computes how much the variance of an evaluated regression coefficient is enhanced due to colinearity. VIF model is employed for removing correlated independent features. This method is fast and it exploits one pass search to the predictor. In addition, this method is computationally efficient in testing every predictor to the model and it avoids the overfitting issue. It is achieved by regressing each independent variable, let Y on the remaining independent variables (W and Z) and checking how much of it (of Y) is explained by these variables. VIF is measured by

$$V = \frac{1}{1 - R^2} \quad (2)$$

From the expression it is shown that the higher the VIF, higher the R^2 which means the variable X is collinear with Y and Z variables. If all the variables are

completely orthogonal, R^2 will be 0 resulting in VIF of 1.

5.3 Tier 1-Classification

In this work, aKNCN [34] is used to classify the soil classes from the different parameters. The proposed aKNCN overcomes the challenges of conventional KNCN and enhanced the performance of KNCN classification. KNCN is a non-parametric classifier that depends on the centroid distance. This states the nearest neighbours of the test samples should satisfy the following criteria - it should be close to test samples and the nearest neighbour distribution should be symmetrical in test samples. But it is complex to determine neighbours in a feature that satisfy these properties. Though KNCN achieves good accuracy it lags in classification time. Hence aKNCN is developed for improving the classification time by adaptively adjusting the nearest centroid neighbour for every input sample to enhance the classification accuracy. Two properties of aKNCN are given as follows:

Property 1:

The aKNCN method satisfies a stable searching phase only when j^{th} distance of nearest centroid is more than pre-defined limit which is multiplier product of k_i and the first nearest centroid $z_{ncn,1}$ to the test sample $d(y, z_{ncn,1})$. The size of neighbourhood is represented as

$$d(y, z_i^s) > k_1 \times d(y, z_{ncn,1}) \quad (3)$$

Where $d(y, z_i^s)$ is the nearest centroid distance among test samples, z_i^s and y . The multiplier product is higher or equal to 1 and the 1st centroid distance is $d(y, z_{ncn,1})$.

Property 2:

The aKNCN method satisfies searching phase only when the entire sample class, M_i , is found amongst j nearest neighbour and the whole samples per class to compete class is lesser than $M_i - 1$. Then the property is defined as

$$(\vee MCM_i (\exists V_{wi}) \wedge (\exists V_{wi}')) \quad (4)$$

$(\vee MCM_i(x) = \{mcm_x^y \mid mcm_x^y \in V\}_{y=1}^j$ is j nearest neighbour of $V_{wi} = \{y_y^i \mid y_y^i \in V\}_{y=1}^{M_i}$ is a subset of V from wi with training samples, M_i and $V_{wi}' = \{y_y^i \mid y_y^i \in V\}_{y=1}^{M_i'}$ is a subset of V' from wi with training samples, M_i . Finally, the soil quality is classified and the crop yield is predicted based on the classified properties for different classes. The yield prediction is performed using ELM, which is discussed in the next subsection.

5.4 Tier 2-Prediction

In this phase, ELM [36] is proposed to predict the crop yield based on the classified soil properties of different classes along with different parameters such as rainfall and temperature. In ELM, a new metaheuristic algorithm called mBOA is hybridized to tune the optimal set of ELM parameters such as thresholds and weights that enhance the performance accuracy with fast convergence. mBOA is a novel approach, which solves the

convergence problems and provides robustness ELM has learning speed and has a better generalization because there is no need to tune the initial parameters of the hidden layer. The hidden layer Feed Forward Network (FFN) is converted into the linear equation by minimum norm least squares. The aim of the ELM is to reduce the output norm weight and training error at the same time. For the samples $\{(Z_i, T_i) \mid X_i \in S^m, T_i \in S^n, i = 1, 2, \dots, N\}$, the P neurons hidden layer with the output function is:

$$f_p(Z) = \sum_{i=1}^P \beta_i r(Z) = r(Z) \beta \quad (5)$$

Where $\beta = [\beta_1, \beta_2, \dots, \beta_P]$ is the output weight vector between output neuron and P . The hidden layer output vector to the input X is given by

$$r(Z) = [r_1(Z), r_2(Z), \dots, r_P(Z)] \quad (6)$$

For enhancing the generalization and to reduce the training error of neural networks, at the same time both output weight and the training error must be minimized.

$$\min : \| h\beta - T \|, \| \beta \| \quad (7)$$

According to (Karush–Kuhn–Tucker) the equation (19) can be written as

$$\beta = h^T \left(\frac{1}{R} + hh^T \right)^{-1} T \quad (8)$$

Where h is the output matrix of the hidden layer, R is the coefficient of the reflection and T is the expected samples and the ELM algorithm output function is

$$f(Z) = g(Z)h^T \left(\frac{1}{R} + hh^T \right)^{-1} T \quad (9)$$

When the feature mapping function $r(Z)$ is unknown, ELM kernel matrix on the basis of Mercer's condition is given by

$$M = hh^T, : m_{ij} = r(Z_i)r(Z_j) = L(Z_i, Z_j) \quad (10)$$

The output function $g(Z)$ on the basis of KOELM is given by

$$f(Z) = L(Z, Z_1).....L(Z, Z_n) \left[\left(\frac{1}{R} + M \right)^{-1} T \right] \quad (11)$$

Where $L(Z, Z_1)$ and $M = hh^T$ are the hidden neurons kernel function of single hidden layer FFN networks. The functions like polynomial kernel, exponential kernel, linear kernel and Gaussian kernel will satisfy the Mercer condition.

BOA [35] is a nature based metaheuristic approach which influences the behaviour of mating and foraging of butterfly. One of the major properties of BOA varies from other optimization approaches that are every butterfly has its separate scent. The fragrance is expressed as:

$$f_r = sI^b \quad (12)$$

Where f_r represents the identified magnitude of fragrance, s is modality of sensor and I^b represents stimulus intensity with absorption of fragrance.

The value of s ranges from $[0, \infty]$ but the value is identified by a particularity of the optimization issues in the BOA iterative

procedure. The s in the optimal solution of the method is expressed as

$$s_{t+1} = s_t + \left[\frac{0.025}{c_t T_{\max}} \right] \quad (13)$$

Where T_{\max} is the maximum iteration and initial value of s is 0.01.

Further, there are two stages in the process, global search and local search space. The mathematical calculation of global search is calculated as

$$x_j^{t+1} = x_j^t + (r_i^2 \times g_b - x_j^t) f_r \quad (14)$$

Here x_j^t is the solution vector x_j of the j^{th} butterfly in iteration t and r_i is a random number and ranges from $[0,1]$. Then g_b is the present best solution identified among every stage in the present stage. Then the local search space is expressed as

$$x_j^{t+1} = x_j^t + (r_i^2 \times x_i^m - x_j^t) f_r \quad (15)$$

Here x_i^m and x_j^t are the m^{th} and j^{th} butterflies selected randomly and when x_i^m and x_j^t is considered under same iteration, that means butterflies becomes a local random walk. When x_i^m and x_j^t is not considered under same iteration random walk may diversify the solution.

Both local search and global search for mating and food partner via the butterflies in nature can happen. Hence a switch probability is considered to transform the intensive local search and normal global search.

It is seen from Equation (14) and (15) that choosing randomly local and global search

will affect BOA is trapped by local optima. Further the parameter r_i capacity for adjusting local and global is limited. Therefore some modification is needed. Hence the new optimal solution is obtained by the following equations.

Therefore the new equation for global search is calculated as

$$x_j^{t+1} = w \cdot x_j^t + (r_i^2 \times g_b - x_j^t) f_r \quad (16)$$

The new equation for local search space is expressed as

$$x_j^{t+1} = (w-1) x_j^t + (r_i^2 \times x_i^m - x_j^t) f_r \quad (17)$$

Where w is a weighting coefficient. Comparing with Equation (14) and (15), the updated equation (16) and (17) has the features like the weighting coefficient is able to adjust among local and global search when compared to the original BOA. The best solution is updated either by equation (16) or equation (17). These two equation provides better convergence speed because of the weighting coefficient. Further this model has better convergence speed and avoid local optima. Therefore this mBOA provides better results due to the optimal value. The Pseudo-code modified BOA is algorithm 1.

Algorithm 1: Pseudo-code of mBOA

```

Initialize butterflies population
Initialize stimulus intensity  $I^b$ 
Initialize switch probability and modality of sensor  $s$ 
While termination criteria not met do
    Compute fragrance by Equation (12)

```

```

End for
Find the best fragrance
For every search agent
    Initialize the random number
    If  $r_i < p$  then
        Move towards best position using Equation (16)
    Else
        Move randomly by Equation (17)
    End if
End for
Update the power exponent  $b$ 
End while
Return the best fitness solution

```

6. Experimental results and discussion

This section gives the performance analysis and discussion about the developed scheme. The entire implementation has been processed on a system with 8 GB RAM and Intel Core i5 CPU with 3.0 GHz speed. To implement the proposed scheme, PYTHON 3.8 is utilized. The dataset taken in this paper for the experimentation is soil dataset. The developed approach performance is implemented with metrics like RMSE, R^2 , MSE, MedAE, MAE, MSLE, MAPE and EVS, error measures and accuracy are utilized for the performance evaluation.

6.1 Evaluation metrics

The error measures are estimated at every iteration and the mean of each metric is

evaluated after the imputation of all the values to obtain their overall performance. The mathematical expression for each metric is defined as:

MSE computes the mean of the squares of the errors. That means mean squared variation among actual and estimated values. MSE is represented as

$$MSE = \frac{1}{n} \sum_{i=1}^n (p_i - \hat{p}_i)^2 \quad (18)$$

MSLE is a ratio among the actual and predicted values. Mean squared logarithmic error is as the name suggests, it is a variation of the MSE.

$$MSLE = \frac{1}{n} \sum_{i=1}^n \left(\ln(1 + p_i) - \ln(1 + \hat{p}_i) \right)^2 \quad (19)$$

RMSE is the differences between values predicted using an estimator and the observed values and it is represented as

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_i - \hat{p}_i)^2} \quad (20)$$

EVS is the name suggested, it is metric used to calculate the ratio among error variance and true values variance and it is represented as

$$EVS = \frac{1}{n} \sum_{i=1}^n \left| \frac{p_i - \hat{p}_i}{p_i} \right| \quad (21)$$

MedAE is a robust computation of the variability of a univariate sample of quantitative data and it is expressed as

$$MedAE = median\left(\left|p_i - \hat{p}_i\right|, \dots, \left|p_n - \hat{p}_n\right|\right) \quad (22)$$

MAPE also known as mean absolute percentage deviation (MAPD), it is a calculation of prediction accuracy of a forecasting approach in statistics and it is computed as

$$MAPE = \frac{1}{n} \sum_{i=1}^n p_i \quad (23)$$

MAE is the absolute difference between an observed value of a quantity and the true value. That is the difference between true and measured length and it is expressed as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |p_i - \hat{p}_i| \quad (24)$$

R^2 is a statistical measure of fit that represents how much difference of a dependent variable is expressed by the independent variable in a regression model and it is expressed as

$$R^2 = \left(1 - \frac{\sum_{i=1}^n (p_i - \hat{p}_i)^2}{\sum_{i=1}^n p_i} \right)^{1/2} \quad (25)$$

Accuracy is defined as the corrected prediction to the total number of prediction and it is calculated by the below expression

$$Accuracy = \frac{\text{Correct prediction}}{\text{Total number of prediction}} \quad (26)$$

Where the expected outcome is denoted as p_i , the predicted outcome for data is represented as \hat{p}_i , where $i = 1, 2, \dots, n$.

6.2 Comparative analysis of aKNCN - ELM-mBOA and Existing approaches

The performance of proposed aKNCN-ELM-mBOA is compared with the state of

the art techniques like ELM, artificial neural network (ANN), support vector machine (SVM), gradient boost (GB) and random forest (RF). The simulation is performed on aKNCN -ELM-mBOA with these existing methods using the error metrics to determine the prediction efficiency of each method.

Table 1: Error measures of various approaches

Approaches	Error measures							
	MAE	MSE	RMSE	MSLE	R ²	EVS	MedAE	MAPE
aKNCN -ELM-mBOA	0.064	0.091	0.301	0.011	0.817	0.818	0.0504	3.932
aKNCN- ELM-BOA	0.067	0.095	0.309	0.011	0.806	0.808	0.0529	3.871
aKNCN -ELM	0.097	0.134	0.366	0.016	0.730	0.731	0.078	5.565
aKNC _ANN	0.130	0.181	0.426	0.022	0.636	0.636	0.105	7.809
aKNC-SVM	0.165	0.230	0.480	0.028	0.538	0.538	0.132	9.685
aKNC-GB	0.231	0.320	0.566	0.039	0.359	0.359	0.187	13.71
aKNC-RF	0.293	0.413	0.642	0.050	0.174	0.17	0.233	17.4

Table 1 provides the analysis of error values attained by various approaches in terms of RMSE, MAE, MSLE, MAPE, MedAE, EVS and MSE. MSE is the mean of sum of the squared errors. MSLE is the squared logarithmic error's predicted

value. From the table, the MAE, MSE, RMSE, MSLE, MedAE and MAPE value obtained by the proposed aKNCN -ELM-mBOA is lower than the existing methods. While the R² and EVS values attained by the proposed aKNCN -ELM-mBOA are higher than other strategies.

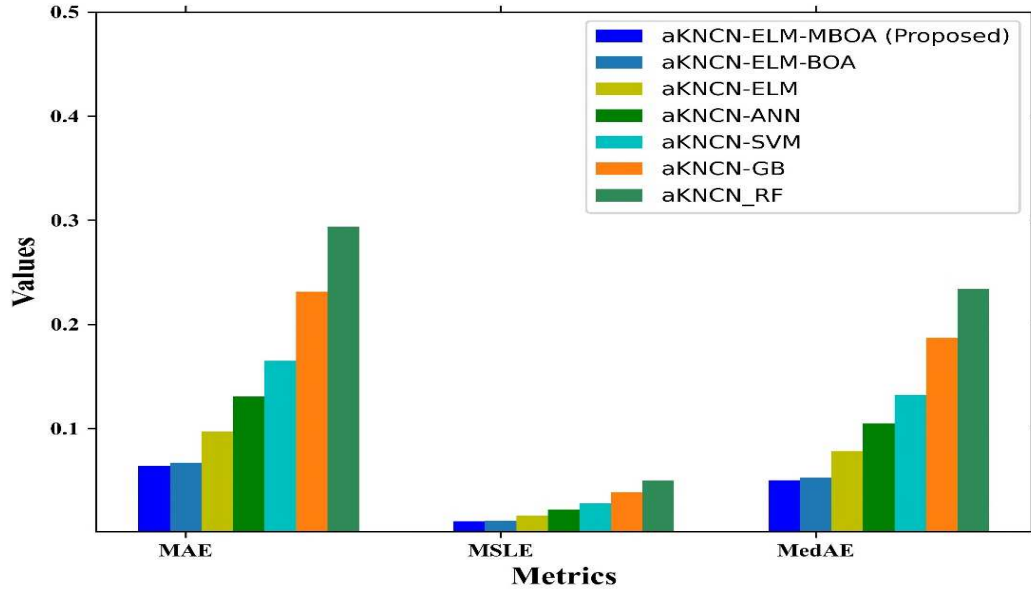


Figure 9: Comparative analysis on MAE, MSLE and MedAE

Figure 9 illustrates the resultant graph of MAE, MSLE and MedAE for various approaches. From the graphical representation, the error values for MAE, MSLE and MedAE obtained by the proposed method is lower than the existing methods. The MAE of the aKNCN-ELM-mBOA is 0.064 and the MAE of the aKNCN -ELM-BOA, aKNCN -ELM, aKNC _ANN, aKNC-SVM, aKNC-GB

and aKNC-RF are 0.067, 0.097, 0.130, 0.165, 0.231 and 0.293 respectively. Similarly the proposed model achieves the better value of MSLE and MedAE. In general, if the error occurred is less in the prediction system then it is considered as an effective model. Therefore, the proposed model is efficient for crop yield prediction.

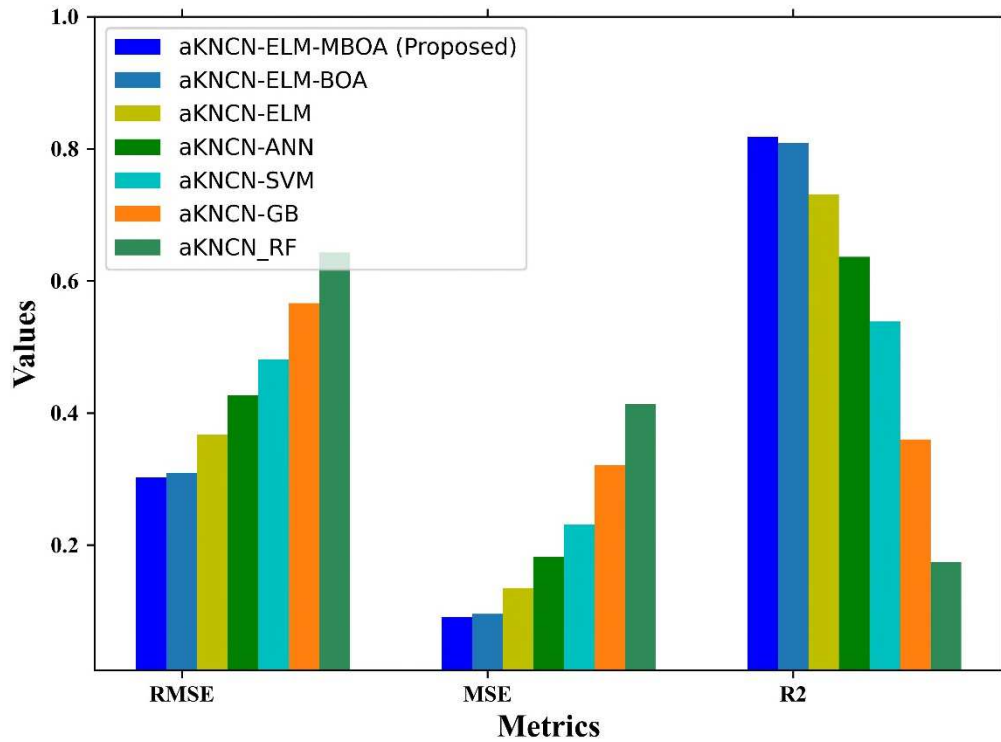


Figure 10: Comparative analysis on RMSE, MSE and R^2

Figure 10 shows the resultant graph of RMSE, MSE and R^2 for the proposed and the existing techniques. From the graphical representation, the proposed model obtained less error values for MSE as well as RMSE than other methods. However, R^2 value attained by the proposed approach is higher than the existing techniques. The R^2 value of the proposed model is 0.817, where the R^2 value of

aKNCN-ELM-BOA, aKNCN-ELM, aKNC-ANN, aKNC-SVM, aKNC-GB and aKNC-RF are 0.806, 0.730, 0.636, 0.538, 0.359 and 0.174. In addition the RMSE value of the proposed model is 0.301 which is less error rate than the other classification methods. Thus the proposed model proved its performance in all the cases.

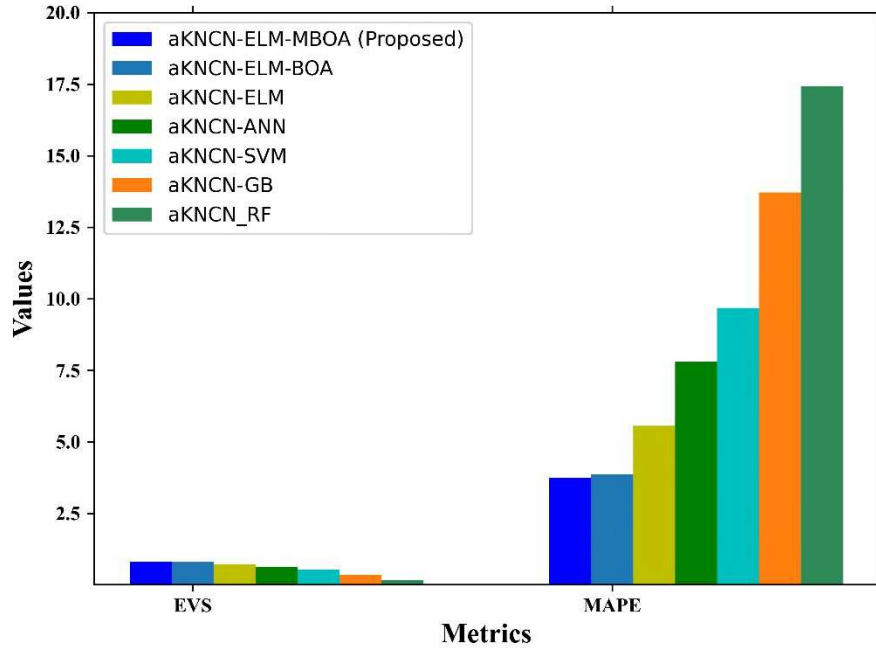


Figure 11: Comparative analysis on EVS and MAPE

Figure 11 represents the resultant graph of EVS and MAPE for the proposed and the existing techniques. From the graphical representation, the MAPE value obtained by the proposed method is lesser than any other methods. However, EVS value attained by the proposed approach is higher than the existing techniques. That is, MAPE value of aKNCN -ELM-mBOA and aKNCN -ELM-mBOA are 3.932 and 3.871 where the MAPE value of aKNCN-

ELM, aKNCN -ANN, aKNCN-SVM, aKNCN-GB and aKNCN-RF are 5.565, 7.809, 9.685, 13.71 and 17.1 respectively. In all the metrics comparison, the proposed model has generated better outcomes and other methods achieve less accuracy due to the computational complexity and overfitting problem. Our proposed model attains higher outcomes due to the optimal selection by mBOA.

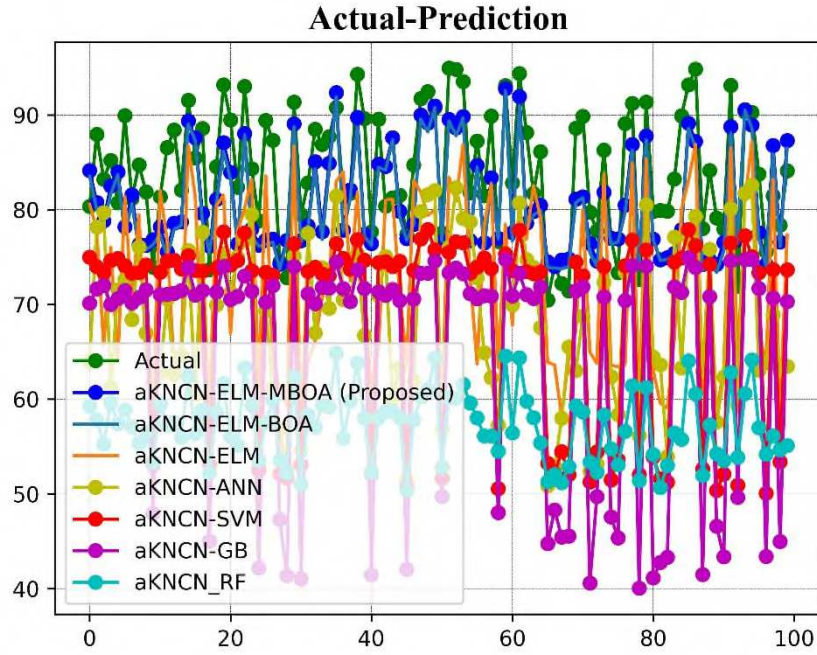


Figure 12: Accuracy measure of proposed aKNCN -ELM-mBOA and existing methods

Figure 12 represents the accuracy measure for the proposed aKNCN-ELM-mBOA and the existing approaches. While considering the accuracy measure, the actual and the predicted data are nearly same then the system is said to be efficient for the crop yield prediction. It depicts the actual data and the predicted data of different techniques. From the graph, the proposed aKNCN-ELM-mBOA predicted the result more accurately than the existing methods. The proposed aKNCN-ELM-mBOA reached near to the actual data whereas the other techniques did not attain a better accuracy. AKNCN-GB achieved very low accuracy that other strategies. Therefore, the proposed aKNCN-ELM-mBOA is effective than the existing techniques.

7. Conclusion

This work focuses on predicting the yield of the crop based on two-tier ML approach named aKNCN and ELM-mBOA. In the first tier, the proposed aKNCN model is used to estimate the soil quality based on IoT system collected soil nutrients. In the second tier, the soil quality score along with other crop yield related parameters like temperature and rainfall are taken as the input of ELM model to predict the crop yield. The hyper parameter tuning of ELM prediction model is achieved by mBOA to enhance the prediction performance of ELM. PYTHON tool is used for the implementation of proposed system. Soil dataset is utilized for performance evaluation of the proposed prediction model. The proposed scheme attains better results than the other classification models on the basis of accuracy, RMSE, R^2 , MSE, MedAE, MAE, MSLE, MAPE and EVS. The RMSE and MAE of the aKNCN-

ELM-mBOA is found to be 0.301 and 0.064 respectively. In future, analysis based on time-series will be done to predict the future values. The use of different parameters like soil nutrients, soil quality, irrigated area and agricultural points can be used to extend the scope of the research as well as improve the accuracy of the system. In addition, deep learning based smart agriculture can be used with the IoT system in order to enhance the production quality.

References

- [1] Muangprathub, Jirapond, Nathaphon Boonnam, Siriwan Kajornkasirat, Narongsak Lekbangpong, Apirat Wanichsombat, and Pichetwut Nillaor. "IoT and agriculture data analysis for smart farm." *Computers and electronics in agriculture* 156 (2019): 467-474.
- [2] Mekala, Mahammad Shareef, and P. Viswanathan. "(t, n): Sensor Stipulation with THAM index for smart agriculture decision-making IoT system." *Wireless Personal Communications* 111, no. 3 (2020): 1909-1940.
- [3] Ayaz, Muhammad, Mohammad Ammad-Uddin, Zubair Sharif, Ali Mansour, and El-Hadi M. Aggoune. "Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk." *IEEE Access* 7 (2019): 129551-129583.
- [4] Savchenko, Olesya M., Maik Kecinski, Tongzhe Li, and Kent D. Messer. "Reclaimed water and food production: Cautionary tales from consumer research." *Environmental research* 170 (2019): 320-331.
- [5] Terence, Sebastian, and Geethanjali Purushothaman. "Systematic review of Internet of Things in smart farming." *Transactions on Emerging Telecommunications Technologies* 31, no. 6 (2020): e3958.
- [6] Kalimuthu, M., P. Vaishnavi, and M. Kishore. "Crop Prediction using Machine Learning." In *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 926-932. IEEE, 2020.
- [7] Farooq, Muhammad Shoaib, Shamyala Riaz, Adnan Abid, Kamran Abid, and Muhammad Azhar Naeem. "A Survey on the Role of IoT in Agriculture for the Implementation of Smart Farming." *IEEE Access* 7 (2019): 156237-156271.
- [8] Reddy, Kasara Sai Pratyush, Y. Mohana Roopa, and Narra Sai Nandan. "IoT based Smart Agriculture using Machine Learning." In *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 130-134. IEEE, 2020.

- [9] Miranda, Jhonattan, Pedro Ponce, Arturo Molina, and Paul Wright. "Sensing, smart and sustainable technologies for Agri-Food 4.0." *Computers in Industry* 108 (2019): 21-36.
- [10] Shastry, K. Aditya, and H. A. Sanjay. "Hybrid prediction strategy to predict agricultural information." *Applied Soft Computing* 98 (2021): 106811.
- [11] Elavarasan, Dhivya, and PM Durairaj Vincent. "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications." *IEEE Access* 8 (2020): 86886-86901.
- [12] Dang, Chaoya, Ying Liu, Hui Yue, JiaXin Qian, and Rong Zhu. "Autumn Crop Yield Prediction using Data-Driven Approaches:-Support Vector Machines, Random Forest, and Deep Neural Network Methods." *Canadian Journal of Remote Sensing* (2020): 1-20.
- [13] van Klompenburg, Thomas, Ayalew Kassahun, and Cagatay Catal. "Crop yield prediction using machine learning: A systematic literature review." *Computers and Electronics in Agriculture* 177 (2020): 105709.
- [14] Bhojani, Shital H., and Nirav Bhatt. "Wheat crop yield prediction using new activation functions in neural network." *Neural Computing and Applications* (2020): 1-11.
- [15] Gopal, PS Maya, and R. Bhargavi. "A novel approach for efficient crop yield prediction." *Computers and Electronics in Agriculture* 165 (2019): 104968.
- [16] PS, Maya Gopal. "Performance evaluation of best feature subsets for crop yield prediction using machine learning algorithms." *Applied Artificial Intelligence* 33, no. 7 (2019): 621-642.
- [17] Shook, Johnathon, Tryambak Gangopadhyay, Linjiang Wu, Baskar Ganapathysubramanian, Soumik Sarkar, and Asheesh K. Singh. "Crop yield prediction integrating genotype and weather variables using deep learning." *arXiv preprint arXiv:2006.13847* (2020).
- [18] Nevavuori, Petteri, Nathaniel Narra, Petri Linna, and Tarmo Lipping. "Crop Yield Prediction Using Multitemporal UAV Data and Spatio-Temporal Deep Learning Models." *Remote Sensing* 12, no. 23 (2020): 4000.
- [19] Elavarasan, Dhivya, and PM Durai Raj Vincent. "A reinforced random forest model for enhanced crop yield prediction by integrating agrarian parameters." *Journal of Ambient*

[20] Suchithra, M. S., and Maya L. Pai. "Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters." *Information processing in Agriculture* 7, no. 1 (2020): 72-82.

[21] Abbas, Farhat, Hassan Afzaal, Aitazaz A. Farooque, and Skylar Tang. "Crop yield prediction through proximal sensing and ML algorithms." *Agronomy* 10, no. 7 (2020): 1046.

[22] Rezk, Nermeen Gamal, Ezz El-Din Hemdan, Abdel-Fattah Attia, Ayman El-Sayed, and Mohamed A. El-Rashidy. "An efficient IoT based smart farming system using ML algorithms." *Multimedia Tools and Applications* 80, no. 1 (2021): 773-797.

[23] Bu, Fanyu, and Xin Wang. "A smart agriculture IoT system based on deep reinforcement learning." *Future Generation Computer Systems* 99 (2019): 500-507.

[24] Nevavuori, Petteri, Nathaniel Narra, and Tarmo Lipping. "Crop yield prediction with deep convolutional neural networks." *Computers and electronics in agriculture* 163 (2019): 104859.

[25] dos Santos, U.J.L., Pessin, G., da Costa, C.A. and da Rosa Righi, R., 2019. AgriPrediction: A proactive internet of things model to anticipate problems and improve production in agricultural crops. *Computers and electronics in agriculture*, 161, pp.202-213.

[26] Saranya, C. P., and N. Nagarajan. "Efficient agricultural yield prediction using metaheuristic optimized artificial neural network using Hadoop framework." *Soft Computing* 24, no. 16 (2020): 12659-12669.

[27] Filippi, Patrick, Edward J. Jones, Niranjana S. Wimalathunge, Pallegedara DSN Somarathna, Liana E. Pozza, Sabastine U. Ugbaje, Thomas G. Jephcott, Stacey E. Paterson, Brett M. Whelan, and Thomas FA Bishop. "An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning." *Precision Agriculture* 20, no. 5 (2019): 1015-1029.

[28] Sun, Jie, Liping Di, Ziheng Sun, Yonglin Shen, and Zulong Lai. "County-level soybean yield prediction using deep CNN-LSTM model." *Sensors* 19, no. 20 (2019): 4363.

[29] Sinwar, Deepak, Vijaypal Singh Dhaka, Manoj Kumar Sharma, and Geeta Rani. "AI-based yield prediction and smart irrigation." In *Internet of Things and*

Analytics for Agriculture, Volume 2, pp. 155-180. Springer, Singapore, 2020.

[30] Cao, Juan, Zhao Zhang, Fulu Tao, Liangliang Zhang, Yuchuan Luo, Jing Zhang, Jichong Han, and Jun Xie. "Integrating Multi-Source Data for Rice Yield Prediction across China using Machine Learning and Deep Learning Approaches." *Agricultural and Forest Meteorology* 297 (2021): 108275.

[31] Guo, Yahui, Yongshuo Fu, Fanghua Hao, Xuan Zhang, Wenxiang Wu, Xiuliang Jin, Christopher Robin Bryant, and J. Senthilnath. "Integrated phenology and climate in rice yields prediction using machine learning methods." *Ecological Indicators* 120 (2021): 106935.

[32] Wakhare, Prashant B., S. Neduncheliyan, and Gaurav S. Sonawane. "Automatic Irrigation System Based on Internet of Things for Crop Yield Prediction." In 2020 *International Conference on Emerging Smart Computing and Informatics (ESCI)*, pp. 129-132. IEEE, 2020.

[33] Patil, Suhas M., and R. Sakkaravarthi. "Internet of things based smart agriculture system using predictive analytics." *Asian J. Pharm. Clin. Res* 10 (2017): 148-152.

[34] Rosdi, B.A., Mukahar, N. and Han, N.T., 2021. Finger Vein Recognition

Using Principle Component Analysis and Adaptive k-Nearest Centroid Neighbor Classifier. *International Journal of Integrated Engineering*, 13(1), pp.177-187.."

[35] Arora, Sankalap, and Satvir Singh. "Butterfly optimization algorithm: a novel approach for global optimization." *Soft Computing* 23, no. 3 (2019): 715-734.

[36] Li, B., Rong, X. and Li, Y., 2014. An improved kernel based extreme learning machine for robot execution failures. *The Scientific World Journal*, 2014.

AUTHORS PROFILE



Ms. Akanksha Gupta is working as an assistant professor in the department of Computer Science, University of Delhi and currently pursuing Ph.D in the area of Internet of Things, from Shri Venkateshwara University. She completed B.Sc honors in Computer science from University of Delhi in 2009 and MCA from Bharatiya Vidyapeeth college, Indraprastha University, Delhi in 2012. She has also qualified UGC NET in computer science in 2013. She has a teaching experience of over 7 years and has earlier published 2 research papers in reputed international journals and 2 chapters in edited books and also participated in many conferences.



Dr. Priyank Nahar pursued his UG degree in Computer

Applications(BCA). He followed it with MCA, M.Phil(CS), Ph.D(CS). He has over 14 years of teaching experience. He worked as a corporate trainer in companies like Hewlett Packard and Integer software, Bangalore. He has attended more than 5 FDPs/seminars and conferences. He has written 5 research papers in international and national journals. He has also authored 13 books on paper solution series along with a book on VB.Net for BCA students. Currently, he is associated with MCA department as an Associate Professor.