International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I4P210

Volume 70 Issue 4, 122-134, April 2022

Original Article

Crop Rotation based Crop Recommendation System with Soil Deficiency Analysis through Extreme Learning Machine

G. Murugesan[1], B. Radha[2]

[1]*Sree Saraswathi Thyagaraja College, Pollachi, TamilNadu, India.*
[2]*Sri Krishna Arts and Science College, Coimbatore, TamilNadu, India.*
[1]g29.murugesan@gmail.com

Received: 21 February 2022 Revised: 30 March 2022 Accepted: 04 April 2022 Published: 26 April 2022

Abstract - *Agriculture is the pillar of the country's economy. Climatic change, soil fertility level, temperature and moisture level, pH value and the crop predecessors often impact crop yield in agriculture. The prediction of the right crop at the right place at the right time will be extremely helpful in increasing the crop yield, which also results in economic proliferation. Machine learning is an emerging technique in the field of agriculture in various ways, including soil classification, soil nutrient analysis, crop prediction or suggestion. This paper presents the crop recommendation system by considering various significant factors, including soil fertility and condition, season and crops predecessor, to recommend appropriate crops for improvising the cultivation based on precision agriculture. Based on the given input, the model applies an extreme learning machine, a fast learning classifier algorithm, for suggesting the appropriate crop to its users. The model also includes the deficiency analysis to identify the deficiency of nutrients in the soil with current crop requirements. The experimental analysis shows that the proposed model provides better accuracy of about 96.5% with a minimum false rate of 3.5% in predicting suitable crops and detecting the deficiency in the soil.*

Keywords - *Crop predecessor, Crop recommendation, Crop rotation, Extreme learning machine, Soil deficiency, Soil fertility.*

1. Introduction

Agriculture is the foremost and most significant source of national income for most evolving countries like India. The development of the agricultural sector subsidies to a marketable surplus. Most of the steps taken in the field of agriculture are aimed at increasing crop productivity. However, the primary requirement is maintaining soil health through enriching its nutrients. Cultivating the same crop in the same land deteriorates the soil nutrients [1]. The crop growth and yield hinge not only on the soil nutrients but also on the acidic or alkaline level of the soil, the climatic change and season, as well as the temperature of the soil [2][3]. Cultivation of crops without knowledge about these parameters will lead to a downfall in crop yield and production.

Owing to the technological development in this digital era, e-agriculture came into existence which is the integration of technology and digital assistance into the agricultural process for achieving effective output in farming. Indian growers will need to reach an impressive level of food production to help in feeding a growing world population. Several studies have been proposed to predict the crops to be cultivated in a specific piece of land based on historical data [4]. However, knowing the level of potassium, nitrogen, phosphorous, and other micronutrients in the soil as well as the season, predicting the suitable crops for cultivation is still a challenging task [5].

Many modern technologies such as machine learning combined with data mining and deep learning help the researcher to achieve the solution for the problems related to the agriculture field [6]. It includes the intensity of rainwater prediction [7], climate prediction using time series analysis [8], disease prediction and classification [9], [10] and water management [11]. However, many solutions related to this problem utilize machine learning algorithms. However, they focus on predicting the type of soil based on the acid or alkaline level in the soil and nutrients present in the soil [12], [13]. Also, in suggesting the crop for cultivation, most of the methods concentrate on the primary nutrients, secondary nutrients, micronutrients present in the soil, pH level, organic carbon and climatic conditions [14], and some methods identify the pests for better yield [15].

Controversially, in ancient times, the soil nutrients were managed by cultivating the crops on a rotational basis based on the crop characteristics. Some crop intake specific nutrients and returns other nutrients to the soil. Thus, cultivating the crops in a specific rotational manner helps the land maintain its nutrients [16]. However, existing models focus only on soil nutrients and the season for recommending the crops, which do not consider the crop predecessor, which is mandatory for maintaining the soil nutrients [17]. Thus, a new algorithm that predicts the crop to be cultivated considering maximum parameters is necessary for the farmers in achieving the maximum result in their field.

This paper presents the crop recommendation system based on a machine learning approach by considering significant factors, including soil fertility and its physical properties, season and crop predecessor, in order to recommend appropriate crops for enhancing crop productivity. The main novelty of the work is that it utilizes crop predecessor along with soil nutrients for crop prediction using an extreme learning machine which offers better results in crop prediction than other traditional algorithms. The model uses the Internet of Things (IoT) to extract various inputs from the farmer, such as global positioning system (GPS) to extract the current location, IoT sensors to read the soil fertility level, moisture and temperature level using sensors and obtains crop predecessor and other details from the farmers. The obtained input is passed over to the extreme learning machine classifier for predicting the crops that fulfil the requirement from the given parameters for better yielding. The model also includes the deficiency analysis to identify the deficiency of nutrients in the soil based on the crop requirements.

The organization of the paper is as follows. Section 2 presents the literature survey related to the field of the study. Section 3 describes the proposed crop recommendation and deficiency analysis system along with an overall architecture, with a subsection explaining the crop recommendation phase and soil deficiency analysis phase through the algorithm pseudocode and a detailed workflow. Section 4 presents the experimental analysis with the dataset used and performance evaluation, followed by the conclusion in section 5.

2. Related Work

Many works were suggested in the literature for classifying soil fertility and predicting the crops for its user. Most of the approaches make use of machine learning algorithms such as Artificial neural networks (ANNs), Naive Bayes (NN), Decision Tree (DT), Random Forests (RF), Support Vector Machine (SVM), AdaBoost, Neigherest Neighbor (NN) and Logistic Regression (LR), Bayesian network, with ensemble concepts of combining results from more models such as boosting, bagging, stacking and more for soil classification and crop prediction [18], [19].

An improved sigmoid kernel SVM classifier was suggested with the soil fertilizer recommendation system for paddy fields by adjusting the cost and gamma parameters. The selection of optimized parameters is also achieved by using the Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) [19]. The author extended their work by classifying soil features based on the available fertility index. This study reduces the unwanted expenditure on consuming fertilizer and thereby increasing profitability [20]. They also proved that the Extreme Learning Machine (ELM) classifier produces better results for the soil dataset. However, they didn't consider the other factors for predicting crops.

Analysis of various regression models in predicting the soil fertility index in the villages in Maharashtra was surveyed. The study concludes that the extremely randomized trees achieve better results [21]. Soil yield prediction [22], disease predictions [23], fertilizer predictions [24], [25] and cultivation patterns [11] are other common systems proposed for agricultural applications.

A yield potential prediction model was suggested specifically for winter wheat by estimating the soil parameters using online soil spectroscopy with a prototype sensor [26]. The model applies supervised self-organizing maps such as counter-propagation based artificial neural networks, XY-fused networks and supervised Kohonen networks. Though the method offers effective results, however, analysing the soil parameters will not be sufficient for the crop yield. Crop prediction depends on various parameters, and thus feature selection plays a significant role. Accordingly, a comparative study of various feature selection methods was made for predicting crops using different classifiers with soil characteristics and environmental factors, including season, rainfall, temperature, and texture [5]. The results show that the recursive feature elimination with an adaptive bagging classifier outperforms other methods used in the comparison.

Another machine learning approach was suggested to analyse the various soil parameters and suggest the crop for cultivation [27]. It utilizes the k nearest neighbour classifier for improving the performance of the recommendation system by mapping the soil and crop data to predict appropriate crops for the specific land. A knowledge-oriented approach for soil classification was proposed that utilizes the random forest for predicting suitable crops. The model aims at improving yield production by analysing the soil in Pakistan in an effective way [2], [28].

Another model was proposed that analyses the various level of soil nutrients such as nitrogen, potassium and phosphorous present in the soil to suggest the crops that can be cultivated in the future. The model utilizes a naïve ratio classifier to predict the crops [16]. Similarly, a naive Bayes algorithm was suggested for crop suggestion that takes the temperature, humidity and moisture content of the soil as a critical factor [17]. However, these models need improvement in prediction accuracy.

To process the uncertainties and to predict the crops, hybridizing rough set on fuzzy approximation space and neural network was proposed. The rough set reduces the computational procedure in data reduction, and the neural network is extremely useful in the prediction process. The model has been analysed based on the agriculture data of the Vellore District in Tamil Nadu [29]. An IoT based gradient descent approach for suggesting crops using a deep learning algorithm such as Artificial Neural Network (ANN) was suggested. However, the model utilizes a few parameters such as soil moisture level, humidity, temperature and pH that are collected through sensors for analysis [30].

A recommendation of a cropping system was proposed for primary crops such as rice, ragi, gram, potato, and onion using a regression-based ensemble model that uses various machine learning algorithms using environmental factors along with yield and prediction [31]. A similar model that makes use of various environmental factors for crop suggestion using an ensemble model that makes use of random forest and XGBoost was suggested [32]. This model lacks result and comparison analysis. A classification algorithm was proposed that utilizes attribute group rank with filter-based instance selection for the effective classification of soil data [33].

Apart from various classifier models suggested for agriculture data, various other classification algorithms were also suggested that are suitable for various other applications, such as an ensemble learning classifier for credit scoring [34] and an instance-based classifier for imbalanced class [35] and more. Though these algorithms were proposed for other applications, the performance of the model concerning the crop prediction application was analysed and compared with the classifier employed in the proposed model. The notable works related to the field of research on various prediction models are summarized and are presented in Table 1.

Table 1. Theoretical analysis of existing crop recommendation models

Model	Methods Used	Advantage	Disadvantage
Soil classification based on crop relevancy [2]	Random forest classifier	Helps in identifying suitable crops	Do not consider other environmental factors
Crop prediction model [5]	Recursive feature elimination technique with the adaptive bagging	Considers soil and environmental factors	Do not consider the crop predecessor
Crop rotation and yield analysis [16]	Naïve ratio algorithm	Predicts suitable crops than naïve Bayes	No detailed study on crop rotation is available
Crop prediction model [17]	Naïve bayes classifer	Uses environmental factors	No detailed result analysis is given
Soil fertilizer recommendation system for paddy fields [19]	Improved SVM classifier with parameter optimization using GA and PSO	Improved accuracy among other variations of SVM	No comparison of results with other classifiers
Prediction model soil fertility indices and pH values [20]	Extreme learning machine	Improved classification result	No comparison of results with other classifiers
Wheat yield potential prediction model [26]	Supervised self-organizing maps	provides better results	Do not consider other factors such as climate for yield prediction
Crop recommendation system [27]	K nearest neighbour classifier	Displays soil deficiency	Only soil properties are used for predicting crops
Crop suitability prediction [29]	Rough set on fuzzy approximation space	Improved classification accuracy	Do not consider crop predecessors
Precision crop suggestion [30]	Deep neural network	Insist on good prediction accuracy	No comparison with other classifier models
Environmental factor-based recommendation crop system [31]	Regression-based ensemble	Most suitable for primary crops	No comparison with other classifier models
Environmental factor-based crop recommendation system [32]	Ensemble classifier	Offers better prediction obtained from random forest and XGBoost	Lacks in comparison with other classifiers

Thus, several models exist for predicting the crops using various parameters. Most of the works concentrate on various soil parameter, and only a few of them uses climate as a parameter. Though crop rotation was a major factor that affects the prediction of suitable crops, unfortunately, it is not focused on by most of the researchers. This creates a need for improving the performance of the crop recommendation system that better suits the soil parameters and climatic conditions.

3. Proposed Crop Recommendation and Deficiency Analysis System

The proposed model utilizes various parameters such as soil fertility, including the number of major nutrients such as Nitrogen (N), Phosphorus (P), Potassium (K) and minor nutrients like sulphur (S), zinc (Zn), iron (Fe), copper (Cu), Manganese (Mn), boron (B) present in the soil, the physical properties of soil such as pH, organic carbon (OC), electric conductivity (EC), the season in which the crop has to be seeded, the crop predecessor for predicting the suitable crop at the right place and at the right time for better crop yield and increased production. The input soil fertility parameters can be obtained based on the IoT sensors at a particular location, whereas the values for the season and the crop predecessors can be obtained from the user based on which appropriate crops suitable for cultivation can be recommended. Additionally, with the details obtained regarding soil fertility, the deficiency that exists in the soil for the current seeded crop can also be analysed. The overall framework of the proposed crop rotation based crop recommendation with a soil deficiency analysis model is shown in Fig. 1.

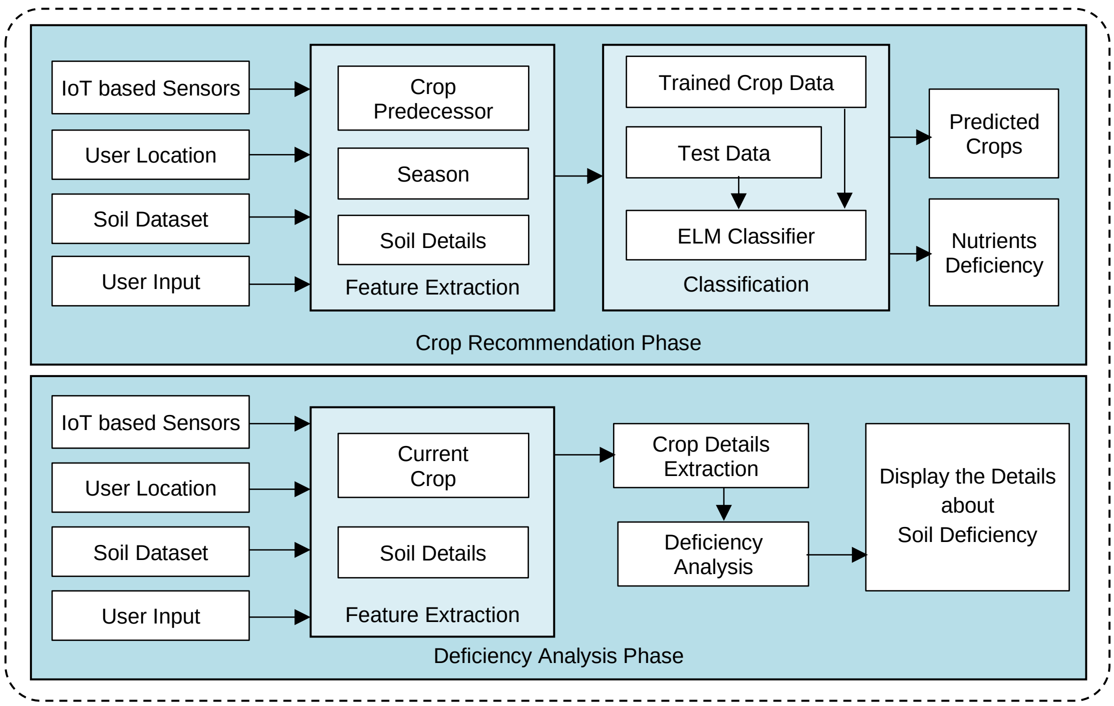

Fig. 1 Overview of the proposed crop recommendation system

The proposed framework is divided into two primary phases, namely crop recommendation and soil deficiency analysis, in which the first phase suggests the crops that can be cultivated in future, and the second phase focuses on improving the yield of the crop currently seeded through soil deficiency analysis. The model utilizes two datasets, namely soil and crop datasets. The soil dataset includes various soil parameters such as soil fertility consisting of the amount of major and minor nutrients present in the soil, and physical parameters such as organic carbon, electric conductivity and pH value of the soil at a particular location identified by its name of the location along with latitude and longitude details. The crop dataset contains the various nutrients required for the particular crop and acceptable crop predecessors for the specific crop. These datasets are trained using extreme learning machines (ELM) for predicting the appropriate crop suitable with respect to the specific factors. An ELM classifier is used in the study as it provides improved accuracy with the soil dataset [20].

3.1 Crop Rotation based Crop Recommendation Phase

The first phase performs crop rotation based crop recommendation analysis for predicting suitable crops. Several existing crop recommendation systems concentrate on climatic conditions and soil fertility to recommend the crops for better yielding. Out of which, only a few systems concentrate on soil physical parameters such as pH value of soil [27]. However, most of the systems are not focusing on the crop predecessor cultivated previously in the specific field. However, in ancient days, before the improvement in technologies, agriculture in India mainly chose the crop to be cultivated in the field based on the two main parameters such as seasons and the crop predecessor. Despite using various organic manures to improve soil fertility, farmers choose suitable crops on a rotational basis to maintain soil stability. Thus, the proposed model primarily focuses on the season, the crop predecessor and the soil nutrients in recommending the appropriate crops. It gets various parameters such as soil fertility, physical characteristics, season and crop predecessor from the user, and the extracted features are analysed and classified using an ELM classifier.

3.1.1 Soil Feature Extraction

The proposed model supports the use of IoT and the cloud for the instant implementation of smart farming and precision agriculture. The model has various input sensors to read the values for various soil fertility parameters and physical parameters such as Nitrogen (N), Phosphorus (P), Potassium (K) and minor nutrients sulfur (S), zinc (Zn), iron (Fe), copper (Cu), Manganese (Mn), boron (B) present in the soil, the physical properties of soil such as pH, organic carbon (OC), electric conductivity (EC) at a particular location using GPS module. For the obtained latitude and longitude from the GPS module, soil parameters from various IoT sensors connected with the Raspberry Pi Microcontroller are retrieved and are stored in the cloud in the soil database.

However, if the IoT kit is not available to the user, then the user is allowed to enter the soil characteristics as a result of independent soil testing done by the user at the soil testing laboratory. Apparently, the input details concerning various soil parameters may or may not be complete. In the event of obtaining complete soil parameters from the user, the crop predecessor and the season are also obtained from the user for further processing. Besides, if the soil parameter values obtained from the user are incomplete, then the Multivariate imputation by chained equation (MICE) is applied to impute the incomplete values.

Also, in the absence of an IoT sensor kit and the independent soil parameters, the latitude and longitude details or the location of the user can be obtained from the android mobile or user. If the latitude and longitude are found in the dataset, then the parameters are fetched based on the last observation carried forward (LOCF), in which the missing values are filled with the most recent data. On having the latitude and longitude, the data for the nearest locations can be extracted, and then the appropriate values for the test data can be evaluated. In both cases, the MICE method is applied to evaluate the values. Thus, if the given latitude and longitude are not available in the soil dataset, the records corresponding to the particular location will be retrieved. The model utilizes the Haversine distance formula to find the nearest locations for the given latitude and longitude [36], and the parameters from the nearest location are extracted to compute the soil details for the test data by imputing with the MICE model. It finds the distance between two points along a great circle of the sphere giving the latitude and longitude of specific locations. The simple representation of the Haversine formula with ⊖ being the central angle of two given points in the sphere can be given as in Eq. (1).

$$\ominus = \frac{Sperical\ distance\ between\ two\ points\ (d)}{Radius\ of\ the\ sphere} \quad (1)$$

However, the Haversine formula supports the Haversine value of θ that can be evaluated using the latitude and longitude of two points (la_1, lo_1) and (la_2, lo_2) in radians as given in Eq. (2).

$$hav(\ominus) = hav(la_2 - la_1) + \cos(la_1)\cos(la_2)\,hav(lo_2 - lo_1) \quad (2)$$

Where the hav() for the difference in the latitude and longitude can be computed as in Eq. (3), the distance between two points in kilometres can be computed as in Eq. (4).

$$hav(\theta) = sin^2\left(\frac{\theta}{2}\right) \quad (3)$$

$$d = 2 \times r \times \arcsin\left(\sqrt{hav(\ominus)}\right) \quad (4)$$

Here, r is the radius of the earth as 6371km. In the proposed model, the records with a distance of less than 5 km are considered for evaluating the approximate value of the test data.

With these selected records, multivariate imputation chained equation (MICE), a statistical model is applied to fill the missing soil nutrients values for the particular location [37]. The main advantage of this imputation method is that it fills the missing values several times instead of having a single attempt that may not be appropriate [38]. Also, the imputation method supports various types of data. The primary idea is to utilize the observed values to estimate the missing values multiple times. It has three stages in which multiple copies of the dataset are created, the missing values are filled with the mean value, and then the regression is applied over the missing variable for each copy, and finally, the results are combined with having a single value. The procedure for MICE() is given in Fig. 2. The proposed model assigns n as 5, in which it executes the model for 5 cycles to compute the complete dataset.

Input: dataset with missing values d and number of cycles n
Output: a complete dataset
Procedure MICE(dataset, n)
1. Simple mean imputation is performed for the missing value of an attribute in a dataset.
2. Repeat the following steps for n times
   a. For each missing variable value
      i. Set back the variable as a missing value.
      ii. Perform logistic regression by considering the missing variable as a target-dependent variable and other variables as the independent predictor variable.
      iii. Fill in the missing values with the predicted or imputed value obtained from the regression.
3. Average the parameter estimates to obtain the missing value.
4. Return the complete dataset

End Procedure

Fig. 2 Multiple imputations to predict the missing values

3.1.2 Season based Crop Rotation

Crop rotation is an important factor earnestly used in ancient times as it provides various benefits in crop yield. The basic idea is to plant different crops sequentially in a particular field over various seasons [39]. The main advantages of using crop rotation are to manage the soil nutrients, diseases, weeds as well as pests. Different crops require different nutrients and are susceptible to different diseases and pests. For example, the corn consumes more amount of nitrogen from the soil, and upon harvesting corn in a specific field, beans can be planted in the next season. Generally, beans return nitrogen to the soil, and therefore the soil fertility will be maintained. On the other hand, planting the same crop repeatedly will deplete the same nutrients present in the soil, as a result of which pests and diseases will infect the soil due to the absence of specific nutrients [40]. The soil health will also be improved due to the increase in the biomass from the various root structure of different crops.

In general, the crops can be grouped by their biological family in which the crops belonging to the same family have similar characteristics, and thus they need similar nutrient requirements and are vulnerable to identical diseases as well as insect pests and can be treated with the same fertilizers and pesticides. The group of crops that belong to the same family are presented in Table 2.

Table 2. Crop families and various crops at each family

Crop Family	**Various Crops**
Solanaceae/ Nightshade	Eggplants, Bell Peppers, Potatoes, Tobacco, Tomatoes.
Brassicaceae/ Brassica	Cabbages, Broccoli, Watercress, Turnips, Radish, Mustard, Kale, cauliflower, Brussels Sprouts
Cucurbitaceae/ Cucumber/Squash	Pumpkins, Squash, Cucumbers, Melons, Gourds, Cantaloupe.
Fabaceae/ Legume	Beans, Peas, Lentils, Peanuts, Soy, Fava Beans.
Poaceae/Grass	Corn, Rice, Wheat, Barley, Oats, Rye, Sorghum, Millet.
Liliaceae/Onion	Garlic, Asparagus, Chives, Shallots, Onions, Leeks.
Umbelliferae/Carrot	Dill, Anise, Garden Angelica, Carrots, Caraway, Celery, Chervil, Cilantro, Cumin, Fennel, Parsnips, Parsley, Beetroot
Asteraceae/ Aster	Absinthe, Artichokes, Chamomile, Cardoons, Chicory, Tarragon, Lettuce, Dandelions, Salsify, Sunflower, Marigold, Zinnia.
Chenopodiaceae/ Spinach	Swiss Chard, Beets, Spinach.

Crop rotation improves soil fertility and manages diseases to increase the yield. The crop groups that can be planted sequentially in a specific field are presented in Fig. 3. To acquire the complete benefit of the crop rotation, the planting can be done sequentially based on the rotational group of crops belonging to the same family, and the crops of the same family should not follow each other.

Choosing the suitable time to seed the plant is also as important as choosing the suitable crop to increase the yield. The proper germination of the seed requires a specific temperature, and thus, the climate or the season plays a significant role [41]. For example, tomato seeds require a temperature of about 20-30 ºC for proper germination, so they can be seeded during the spring season, whereas the carrot and beetroot require 10-30 ºC for seed germination and can be sowed during the end of the monsoon season. Thus, the model also makes use of season to predict the suitable crop to be planted.

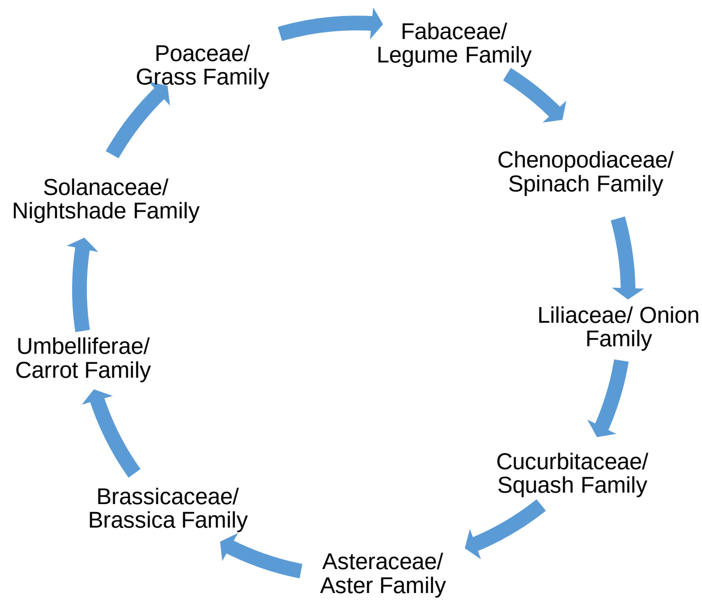

Fig. 3 Crop rotation based on crop family

3.3.3 Extreme Learning Machine Classifier

The proposed model utilizes extreme learning machines for the effective classification of crops to be recommended for the given input parameters. Extreme Learning Machine (ELM) is a regularized feed-forward neural network specifically used for classification and regression with a single layer of hidden nodes. The key feature of the model is that it does not need parameter tuning and thus provides the result with minimum execution time [42]. The model can be trained in a single step which is extremely fast.

ELM is supposed to solve the generalized linear problem denoted as in Eq. (5)

$$H\beta = Y \tag{5}$$

Here H is the matrix of the hidden output layer, β is the weight vector that connects the hidden layer with that of the output, and Y is the target variable to be classified.

The problem can be solved by evaluating the value of β using the pseudo inverse of H. This is similar to finding the value of β attaining the minimum value in the least square problem as in Eq. (6).

$$\beta = \min_{\beta} \|H\beta - T\| \tag{6}$$

ELM procedure can be easily understood with some simple steps. Initially, the model starts by assigning random weights w_i and the bias b_i for the input layer where I varies from 1 to L and L is the number of hidden nodes. The next step is to compute the hidden layer output matrix $g(w_j \cdot x_i + b_j)$ with an initial matrix defined as in Eq. (7).

$$H_{ini} = X.W^t \tag{7}$$

Here g(x)is an activation function. Finally, compute the output weight matrix as in Eq. (8).

$$\beta = H^+T \tag{8}$$

H^+Is the inverse matric of H. Finally, the new target can be predicted by using the output weight matrix β as in Eq. (9).

$$Y = H\beta \tag{9}$$

Thus the proposed model suggests the crops by classifying the given input parameters.

Additionally, it also presents the details concerning the seeding or sowing of the suggested crops for its user. The various details include the sowing depth of seeds, the distance between two seeds vertically and horizontally and the number of days to cultivate the crops and so on. The detailed workflow of the crop rotation based crop recommendation phase is shown in Fig. 4. The algorithm pseudocode for the crop rotation based crop recommendation phase is given in Fig. 5.

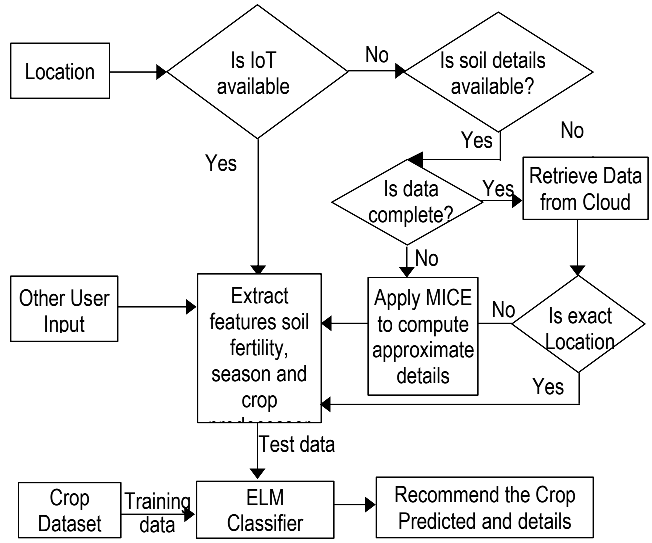

Fig. 4 Detailed workflow of crop recommendation phase

3.2 Soil Deficiency Analysis Phase

This phase is responsible for analysing the soil nutrients required for the specific crops by comparing the given input parameters with that of the requirement stored in the crop dataset. The soil input parameters can be acquired through the IoT sensors if it is available. In the absence of IoT devices, the user can provide the soil fertility details through the results obtained from independent soil testing. On the other hand, if neither way works, then the parameters are fetched through the latitude and longitude of the user location based on the last observation carried forward (LOCF), in which the missing values are filled with the most recent data. However, if the record for the latitude or longitude is not available, the records of the nearest location can be obtained using the Haversine distance formula and on which the incomplete data can be imputed using the multivariate imputation by chained equation (MICE).

Upon receiving the soil details, the crop for which the deficiency has to be analysed is obtained from the user. The details for the particular crop are compared with that of the one stored in the crop database. Each soil parameter extracted is compared with that of the requirement of the crop (from the crop database), and if the extracted value is minimum, then it signifies there is a deficiency in soil nutrients, and the details are displayed to the user for which the action can be carried out to treat the deficiency. The detailed workflow of the soil deficiency analysis phase is shown in Fig.6. The algorithm pseudocode for the soil deficiency phase is given in Fig. 7.

```
Input: soil dataset, crop dataset
Output: recommendation of crop
Procedure crop_recommend()
   Get crop predecessor, season from the user
   //Extract the soil database (db) stored in the cloud
   If IoT_deveice is available then
       Extract location and soil parameters & update soil db
   Else if soil details are available from the user then
       If soil data is incomplete then
            Call MICE(dataset,n) & Extract soil parameters
       End if
   Else search the location of Soil data in the cloud
       If the exact location is available then
            Extract LOCF(soil parameters)
       Else apply Hav. formula to find k nearest locations
            Call MICE(dataset, n) & Extract soil parameters
        End If
    End If
    //Classification using Extreme Learning Machines
    Train the crop dataset using an ELM classifier
        Assign random weights wi and bias bi for the input layer
        Compute hidden layer matrix H with activation fn
        Evaluate the output weight matrix β.
    Predict target class with weight matrix as in Eq. (9)
    // Display the details
    Display the predicted plant groups
    List seeding details (sowing depth, dist., maturity time)
End Procedure
```

Fig. 5 Algorithm pseudocode for crop recommendation phase

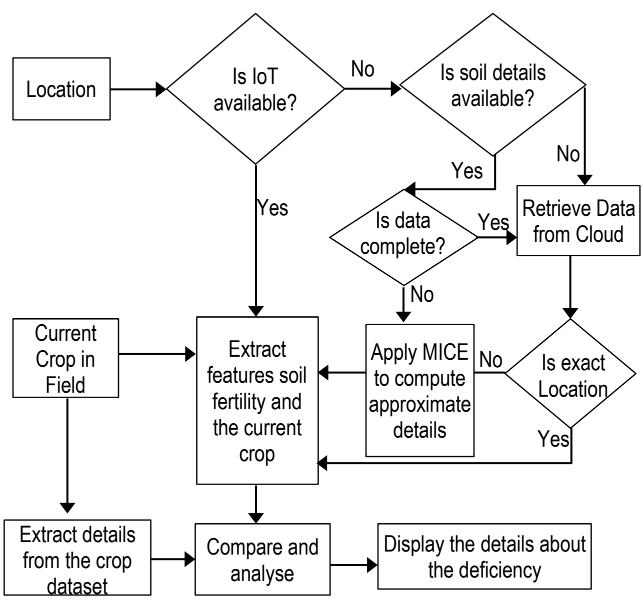

Fig. 6 Detailed workflow of the deficiency analysis phase

```
Input: soil dataset, crop dataset
Output: recommendation of crop
Procedure soil_deficiency()
   Get crop details from the user
   //Extract the soil database (db) stored in the cloud
   If IoT_deveice is available then
       Extract location and soil parameters and update db
   Else if soil details are available from the user then
       If soil data is incomplete then
            Call MICE(dataset,n) & Extract soil parameters
       End if
   Else search the location of Soil data in the cloud
       If the exact location is available then
            Extract LOCF(soil parameters)
       Else apply Hav. formula to find k nearest locations
            Call MICE(dataset,n) & Extract soil parameters
        End If
    End If
    //Compare the soil nutrients with the required data
    For each soil_parameter in the test record
        If soil_parameter < required (crop dataset) then
            // Display the soil deficiency details
         End If
     End For
End Procedure
```

Fig. 7 Algorithm code for soil deficiency analysis phase

4. Experimental Analysis

For evaluating the performance of the proposed study, the soil database has been created by utilizing the soil database at the Pollachi region available at the Department of Agriculture, Cooperation and Farmers Welfare under the Ministry of Agriculture and Farmers Welfare, Government of India at various location [43]. The dataset includes 6718 soil data samples with 3452 samples from Pollachi north surrounding with 49 villages and 3266 samples from Pollachi south region surrounding with 29 villages by eliminating the samples having missing values. The target class trait has 7 distinct values based on the appropriateness and fertility of the soil, such as very low, low, moderately low, moderate, moderately high, high and very high. The database contains 15 attributes that include the location details and primary nutrients, secondary nutrients, micronutrients nutrients and physical properties of the soil at each particular location. The list of attributes is presented in Table 3.

Table 3. List of attributes in soil dataset

Attribute Groups	Attributes
Location	Latitude (LA)
	Longitude (LO)
	Location(LOC)
Soil Primary Nutrients	Nitrogen (N)
	Phosphorus (P)
	Potassium (K)
Soil Secondary Nutrients	Sulfur (S)
Soil Micronutrients Nutrients	Zinc (Zn)
	Iron (Fe)

	Copper (Cu)
	Manganese (Mn)
	Boron (B)
Soil Physical Properties	The potential of Hydrogen (pH)
	Organic Carbon (OC)
	Electrical Conductivity (EC)

The obtained values of the soil characteristics are groped and are rated with 5 categorical values very low, low, medium, high and very high [44]. The soil ratings for various attributes, such as nutrients present in the soil including N, P, K, S, Zn, Fe, Cu, Mn, B, and physical properties including OC and EC, are presented in Table 4.

Table 4. List of attributes in soil dataset

Attribute	Class				
	Very Low	Low	Medium	High	Very High
OC (%)	<0.3	0.3-0.5	0.51-0.8	0.81-1.0	>1
EC (dS/m)	<0.2	0.2-0.5	0.51-0.75	0.76-2.25	>2.26
N (Kg/ha)	<140	141-280	281-440	441-560	>560
P (Kg/ha)	<7	7.1-14	14.1-28	28.1-35	>35
K (Kg/ha)	<100	100-160	161-250	251-300	>300
S (mg/kg)	<6	6-8	8.1-14	14.1-18	>18
Zn (mg/kg)	<0.25	0.25-0.4	0.41-0.7	0.71-1.0	>1.0
Fe (mg/kg)	<1	1-2.5	2.6-3.5	3.6-4.5	>4.5
Cu (mg/kg)	<0.1	0.1-0.3	0.31-0.4	0.41-0.5	>0.5
Mn (mg/kg)	<0.2	0.2-0.5	0.51-0.8	0.81-1.0	>1.0
B (mg/kg)	<0.2	0.2-0.5	0.51-0.8	0.81-1.0	>1.5

Similarly, soil pH value also influences plant growth. Generally, acid soil decreases the soil nutrients and plant growth by increasing the toxic elements in the soil. On the other hand, alkaline soil lowers soil resistivity. For the plant growth to be effective, the soil must be neutral or mildly alkaline. The various rating for the soil pH values and the number of records at each rating is given in Table 5 [45].

Table 5. Soil pH values and ratings

Range of pH value	Ratings	No. of Records
<4.5	Extremely Acidic	2
4.5-5.0	Strongly Acidic	2
5.1-5.5	Highly Acidic	2
5.6-6.0	Moderately Acidic	42
6.1-6.5	Slightly Acidic	217
6.6-7.3	Neutral	446
7.4-7.8	Mildly Alkaline	612
7.9-8.4	Moderately Alkaline	4589
8.5-9.0	Strongly Alkaline	562
>9.0	Very strongly alkaline	244

However, for achieving better results using classification algorithms, the categorical ratings are converted to numerical scale values with the lowest rating is assigned as 1, and each higher rating is assigned with successive integers. The crop dataset containing the soil nutrients requirement for 50 crops is also created in which the crop can have more than one record due to its support towards one or more attribute values for a particular attribute, and the list of attributes in the crop dataset is presented in Table 6. The attributes include the requirement of primary nutrients, secondary nutrients, micronutrients, physical properties, environmental factors such as rainfall, moisture, temperature, and season in which it can be sowed, and the acceptable crop predecessor based on crop rotation as given in Table 2.

Table 6. Attributes of crop dataset

Attributes ID	Attributes
1	Nitrogen (N)
2	Phosphorus (P)
3	Potassium (K)
4	Sulfur (S)
5	Zinc (Zn)
6	Iron (Fe)
7	Copper (Cu)
8	Manganese (Mn)
9	Boron (B)
10	Potential of Hydrogen (pH)
11	Organic Carbon
12	Electrical Conductivity (EC)
13	Season
14	Crop Predecessors

To analyse the performance of the model, the ELM for various activation functions has been evaluated [46]. Then accuracy and the false rate of classification with the soil dataset is evaluated for 10 activation function, and the obtained results are presented in Table 7.

Table 7. Accuracy computation for the various elm activation function

Activation Function	Abbr.	Acc.	False Rate
Sigmoid	elm_sig	87.11	12.891
Swish	elm_swi	88.00	11.998
Exponential Linear Squashing	elm_els	91.43	8.574
Hyperbolic Tangent	elm_htan	92.96	7.041
Hard Hyperbolic	elm_hhb	92.13	7.874
Rectified Linear Unit	elm_rlu	89.12	10.881
TanhRe	elm_tanhr	96.39	3.612
Exponential Linear Units	elm_elu	93.91	6.088
Softplus Function	elm_sftp	92.83	7.175
Leaky ReLU	elm_lrlu	92.27	7.726

From the results obtained, it is clear that the ELM classifier with TanhRe as the activation function offers better results than other activation functions. Thus, the proposed model utilizes an ELM classifier with TanhRe activation for further analysis.

The proposed model is also compared with various standard and existing classification algorithms to analyse the performance. The ELM classifier is compared with that of the Naïve Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Logistic Boost (LB), AdaBoost (AB), Bagging (BAG), Decision Table (DT), RIPPER (JRip), Zero R Classifier (Xero), Random Forest (RF), attribute group rank with a filter based instance selection model (AGRFIS) [33] using the crop dataset for the test data. Various standard evaluation metrics such as accuracy, false rate, sensitivity, fall out, and specificity for assessing the classification performance are utilized in the analysis. The obtained results for the classifiers are given in Table 8.

Table 8. Performance analysis for different classifiers

Classifiers	Performance Metrics				
	A.	False Rate	Sen.	Fall out	Spe.
Naïve Bayes	80.2	19.8	0.802	0.199	0.801
Logistic Regression	90.2	09.8	0.902	0.111	0.889
Support Vector Machine	88.0	12.0	0.880	0.132	0.868
K-Nearest Neighbour	84.0	16.0	0.83	0.183	0.817
Logit Boost	73.0	27.0	0.730	0.272	0.728
Adaboost	69.8	30.2	0.698	0.703	0.697
Bagging	87.2	12.8	0.872	0.141	0.859
Decision Table	88.2	11.8	0.882	0.128	0.872
Ripper	79.5	20.5	0.794	0.212	0.788
ZeroR	49.4	50.6	0.494	0.507	0.493
Random Forest	91.6	08.4	0.916	0.084	0.916
RF-MWMV	91.2	08.8	0.912	0.099	0.901
AWPS	91.8	08.2	0.918	0.092	0.908
AGRFIS	93.6	06.4	0.936	0.069	0.931
ELM (proposed)	96.5	03.5	0.968	0.051	0.969

Among several algorithms, the ELM classifier provides higher accuracy of about 96.5% with a minimum false rate of 3.5% though the AGRFIS model offers better performance with 93.6% accuracy and 6.4% of false rate among other algorithms under comparison. Since the number of samples in the crop dataset is less, the classification rate of different classifiers varies radically from one to another. Similarly, the classifiers such as Random Forest, RF-MWMV and AWPS acquire 91% of accuracy and an 8% of false rate approximately. The accuracy and false rate obtained for various models are presented as a graph in Fig. 8.

The classifiers used for the comparison are also evaluated using various statistical analyses such as precision, f-measure, Area under RoC curve (AUC), Geometric Mean (GM) and Adjusted Geometric Mean (AGM) [47]. The obtained values for the various statistical measure are given in Table 9.

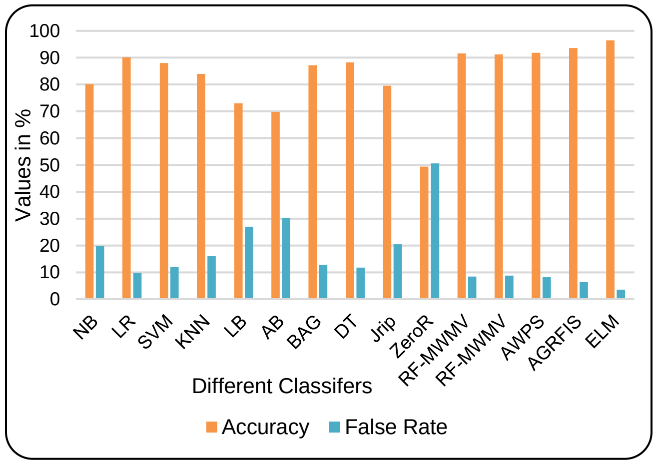

Fig. 8 Classification performance comparison

Table 9. Performance analysis for different classifiers

Models	Statistical Analysis				
	Prec.	F-meas.	AUC	GM	AGM
Naïve Bayes	0.802	0.809	0.801	0.798	0.801
Logistic Regression	0.902	0.891	0.896	0.902	0.889
Support Vector Machine	0.880	0.859	0.874	0.879	0.868
K-Nearest Neighbour	0.840	0.833	0.829	0.842	0.817
Logit Boost	0.730	0.722	0.729	0.733	0.728
Adaboost	0.698	0.692	0.698	0.702	0.697
Bagging	0.872	0.849	0.865	0.872	0.859
Decision Table	0.882	0.857	0.877	0.882	0.872
Ripper	0.795	0.769	0.791	0.794	0.788
ZeroR	0.494	0.488	0.494	0.498	0.493
Random Forest	0.916	0.920	0.916	0.916	0.916
RF-MWMV	0.912	0.901	0.906	0.912	0.900
AWPS	0.918	0.903	0.913	0.918	0.908
AGRFIS	0.936	0.920	0.934	0.936	0.931
ELM (proposed)	0.969	0.968	0.964	0.966	0.963

From the analysis, the AGRFIS model produces better results with a precision of 93.6%, f-measure of 0.92, AUC as 0.93 and 93.57% of Geometric Mean (GM) and 93.14% of Adjusted Geometric Mean (AGM) that are lower than the ELM classifier with 96.9% of precision, 0.968 of F-measure, 0.964 of AuC, 96.6% of GM and 96.3% of AGM. Other classifiers such as Logistic Regression, Random Forest, RF-MWMV and AWPS also produce good results with slightly lower values than AGRFIS and ELM classifiers.

Also, to evaluate the performance of the ELM classifier in predicting suitable crops for the specified region, classification has been made using cross-validation and data split by varying its values. Table 10 shows the performance of the ELM classifier by varying the n value in n-fold cross-validation from 10 to 90. The performance of the ELM classifier is measured using metrics such as accuracy, kappa statistics, precision, recall, specificity, F1 score, mean absolute error (MAE) and log loss.

Table 10. Performance analysis for different classifiers

Various Metrics	ELM Classifier using n-fold Validation								
	10	20	30	40	50	60	70	80	90
Accuracy	95.9	94.5	95.1	95.2	94.2	95.1	94.3	94.3	94.3
Kappa	94.8	94.3	94.3	94.1	93.4	92.9	93.2	93.7	93.3
Precision	97.2	97.1	96.9	96.8	96.6	96.1	95.2	95.9	96.3
Recall	96.5	96.2	96.1	96.0	95.6	95.9	95.2	96.1	96.1
Specificity	99.3	99.1	98.9	98.8	98.5	98.6	98.3	97.9	97.1
F1 Score	97.7	97.1	97.0	96.6	96.8	96.6	96.2	96.1	97.1
MAE	0.7	0.9	1.0	1.0	1.0	1.0	0.8	0.8	0.8
Log Loss	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

The results show that the 10-fold cross-validation offers better results for all the performance metrics with an accuracy of 95.9%, kappa of 94.8%, precision of 97.2%, recall of 96.5%, specificity of 99.3% and f1 score of 97.7%. It not only increases the accuracy but also reduces the error with an MAE of 0.7. On the other hand, the 40 fold offers better results for recall, and the 60 fold offers good results for precision and recall. However, the number of folds is increased to 90, and the performance of the model seems to be very low than others.

Table 11 shows the performance of the ELM classifier by varying split percentages of training and test dataset. The performance of the ELM classifier with different split percentages is measured using metrics such as accuracy, kappa statistics, precision, recall, specificity, F1 score, mean absolute error (MAE) and log loss for the crop dataset in predicting the appropriate crop for the particular agriculture field with its characteristics.

Table 11. Performance analysis for different classifiers

Data Split %	Performance Metrics							
	Acc.	Kappa	Prec.	Rec.	Spe.	F1	MAE	Log Loss
25–75	80.5	83.3	86.9	84.7	88.3	84.3	2.1	1.45
30–70	84.7	86.4	88.2	88.2	91.4	88.8	2.0	0.82
35–65	86.4	88.5	90.7	90.8	93.5	90.5	1.9	0.46
40–60	89.2	90.5	92.5	92.7	95.6	92.7	1.2	0.34
45–55	90.2	91.5	93.7	93.4	96.9	93.2	1.1	0.25
50–50	91.5	92.4	94.8	94.2	97.5	94.9	1.0	0.12
55–45	93.2	93.2	95.9	95.1	98.1	96	0.9	0.06
60–40	93.6	93.7	96.1	95.2	98.4	96.2	0.9	0.03
65–35	94.9	93.7	96.8	96.2	98.3	96.5	0.8	0.02
70–30	95.8	94.9	97.8	96.2	99.1	97.2	0.7	0.02
75–25	95.2	94.8	97.3	96.1	98.7	97	0.8	0.02

The results show that the data splitting percentage of 70% training set and 30% test set offers better results for all the performance metrics. The obtained results for accuracy by varying the data split percentage for training and test set are depicted in Fig 9. The split percentage of 25-75 offers minimum accuracy. The graph shows that the increase in the percentage of training set also increases the accuracy of the model. On the contrary, more percentage of the training set with a minimum test set of 75-25 also degrades the performance.

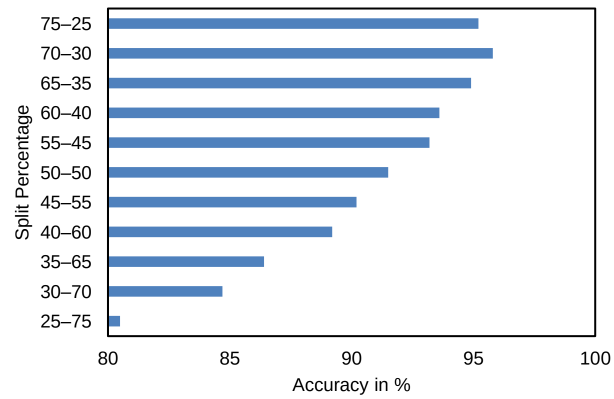

Fig. 9 Accuracy analysis with the varied split percentage

The extensive result analysis performed for the proposed model is compared with various other standard classifiers which are utilized by the various existing crop recommendation models. Also, an analysis was made for the proposed model that makes use of various factors such as soil nutrients, soil physical properties and crop predecessor. The model using soil nutrients offers an average accuracy of 92.38%, and soil nutrients along with the soil physical properties offer an average accuracy of 93.98%, and the factors including soil nutrients, physical properties, as well as crop predecessors offer accuracy of 96.5% using ELM classifier.

Thus, once the crop is classified, the group of crops belonging to the same family can be suggested to the user. Also, upon providing the details about the crop sowed currently, the model compares the current soil nutrients with that of the required soil nutrients stored in the crop dataset, based on which the deficiency in the soil can be listed to the user. However, based on the deficiency, the fertilizers to be used to treat the soil deficiency can also be suggested.

5. Conclusion

Machine learning and data mining have become vital techniques for solving problems in various fields, including agriculture which is the backbone of any country. Like analysing the soil characteristic, and nutrients present in the soil, suggesting the crops to be cultivated by considering various aspects of soil properties and nutrients, temperature, and the season are also crucial as it highly influences the crop yield and economy. In this paper, a crop recommendation system has been introduced that analyse the climatic change, soil fertility level, temperature and moisture level, pH value of the soil and the crop predecessors for improvising the cultivation. The model uses various inputs based on which the crop to be cultivated is classified using an extreme learning machine. It also includes the deficiency analysis to identify the deficiency of nutrients in the soil based on the requirement of the crops. For the experimental analysis, the soil dataset has been used that belongs to the Pollachi region. The results show that the ELM classifier provides better accuracy of about 96.5%, with a minimum false rate of 3.5% and 96.9% of precision. The results show that the model yields better crop prediction and soil deficiency analysis effectively. The future work focuses on recommending the crops based on the demand analysis.

References

[1] A. J. Bennett, G. D. Bending, D. Chandler, S. Hilton, and P. Mills, Meeting the Demand for Crop Production: the Challenge of yield Decline in Crops Grown in Short Rotations, Biological reviews. 87(1) (2012) 52-71.

[2] W. Haider, A. Rehman, M. N. Durrani, and S. Rehman, Knowledge-Based Soil Classification towards Relevant Crop Production, International Journal of Advanced Computer Science and Applications. 10(12) 2019 488-501.

[3] V. Bhatnagar, and R. Chandra, IoT-based Soil health Monitoring and Recommendation System. In the Internet of Things and Analytics for Agriculture, Springer, Singapore. 2 (2020) 1-21.

[4] M. A. Al Maruf, and S. Shatabda, iRSpot-SF: Prediction of Recombination Hotspots by Incorporating Sequence-Based Features into Chou's Pseudo Components, Genomics. 111(4) (2019) 966-972.

[5] A. Suruliandi, G. Mariammal, and S. P. Raja, Crop prediction Based on Soil and Environmental Characteristics using Feature Selection Techniques, Mathematical and Computer Modelling of Dynamical Systems. 27(1) (2021) 117-140.

[6] M. A. Hossen. Mechanization in Bangladesh: Way of Modernization in Agriculture International Journal of Engineering Trends and Technology, 67(9) (2019) 69-77.

[7] S. Sakthicvel, and G. Thailambal, Ada Naïve Bayesian Algorithm for Predicting the Intensity of Rain to Improve the Accuracy. International, Journal of Engineering Trends and Technology, 70(2) (2022) 24-31.

[8] J. Jagannathan, and C. Divya, Time Series Analyzation and Prediction of Climate using Enhanced Multivariate Prophet, International Journal of Engineering Trends and Technology, 69(10) (2021) 89-96.

[9] M. Thanjaivadivel, and R. Suguna, Leaf Disease Prediction Using Fast Enhanced Learning Method, International Journal of Engineering Trends and Technology, 69(9) (2021) 34-44.

[10] T. Jayaraj, and J. Abdul Samath, Disease Forecasting and Severity Prediction Model for COVID-19 using Correlated Feature Extraction and Feed-Forward Artificial Neural Networks, International Journal of Engineering Trends and Technology, 69(8) (2021)126-137.

[11] B. Vinoth, and N. M. Elango, Data Mining of Paddy Cultivation Patterns And Water Resource Management In Late Samba Season of Tamilnadu, International Journal of Engineering Trends and Technology, 69(1) (2021) 152-165.

[12] M. Anita, and S. Shakila, Predictive Analytics in Soil for Agriculture Using Kendall Normalized Feature Selection Based Jaccarized Rocchio Boyer-Moore Bootstrap Aggregative Mapreduce Classifier for Predictive Analytics with Big data, International Journal of Engineering Trends and Technology, 69(9) (2021) 80-91.

[13] S. J. Reshma, and A. S. Pillai, Edaphic Factors and Crop growth using Machine Learning A Review, in International Conference on Intelligent Sustainable Systems (ICISS), IEEE. (2017) 270-274.

[14] H. Teng, R. A. Viscarra, Z. Shi, and T. Behrens, Updating a National Soil Classification with Spectroscopic Predictions and Digital soil Mapping, Catena. 164 (2018) 125–134.

[15] J. Lacasta, F. J. Lopez-Pellicer, B. Espejo-García, J. Nogueras-Iso, and F. J. Zarazaga-Soria, Agricultural Recommendation System for Crop Protection, Computers and Electronics in Agriculture. 152 (2018) 82-89.

[16] M. K. S. Preetha, P. K. Priya, K. DivyaPrabha, and S. Dharanipriya, Crop Rotation and Yield Analysis using Naive Ratio Classification, International Journal of Scientific & Engineering Research. 8(5) 2017 29-34.

[17] M. Kalimuthu, P. Vaishnavi, and M. Kishore, Crop Prediction using Machine Learning, In Third International Conference on Smart Systems and Inventive Technology (ICSSIT, IEEE. (2020) 926-932.

[18] M. S. Suchithra, and M. L. Pai, Improving the Performance of Sigmoid Kernels in Multiclass SVM using Optimization Techniques for an Agricultural Fertilizer Recommendation System, In International Conference on Soft Computing Systems, Springer, Singapore. (2018) 857-868.

[19] A. Chougule, V. K. Jha, and D. Mukhopadhyay, Crop Suitability and Fertilizers Recommendation using data Mining Techniques, in Progress in Advanced Computing and Intelligent Engineering, Springer, Singapore (2019) 205-213.

[20] M. S. Suchithra, and M. L. Pai, Improving the Prediction Accuracy of Soil Nutrient Classification by Optimizing Extreme Learning Machine parameters, Information processing in Agriculture, 7(1) (2020) 72-82.

[21] M. S. Sirsat, E. Cernadas, M. Fernández-Delgado, and S. Barro, Automatic Prediction of Village-wise soil Fertility for Several Nutrients in India using a wide range of Regression Methods, Computers and Electronics in Agriculture, 154 (2018) 120-133.

[22] M. Rekha Sundari, G. Siva Rama Krishna, V. Sai Naveen, and G. Bharathi, Crop Recommendation system using k-nearest Neighbors Algorithm, in Proc. International Conference on Recent Trends in Computing, Springer, Singapore. (2021) 581-589.

[23] S. Gupta, G. Garg, P. Mishra, and R. C. Joshi, CDMD: An Efficient crop Disease Detection and Pesticide Recommendation System using Mobile Vision and Deep Learning, in Proc. International Conference on Big Data, Machine Learning and their Applications, Springer, Singapore. (2021) 295-305.

[24] R. Guan, H. Pan, W. He, M. Sun, H. Wang, X. Cui, Y. Lou, Y. Zhuge, Fertilizer Recommendation for Foxtail Millet Based on Yield Response and Nutrient Accumulation, Journal of Plant Nutrition. 45(3) (2022) 332-345.

[25] J. Rurinda, S. Zingore, J. M. Jibrin, T. Balemi, K. Masuki, J. A. Andersson, M. F. Pampolino, I. Mohammed, J. Mutegi, A. Y. Kamara, and B. Vanlauwe, Science-based Decision Support for Formulating crop Fertilizer Recommendations in Sub-Saharan Africa, Agricultural Systems. 180 (2020) 102790.

[26] X. E. Pantazi, D. Moshou, T. Alexandridis, R. L. Whetton, and A. M. Mouazen, Wheat yield Prediction using Machine Learning and Advanced Sensing Techniques, Computers and Electronics in Agriculture. 121 (2016) 57-65.

[27] A. K. Mariappan, C. Madhumitha, P. Nishitha, and S. Nivedhitha, Crop Recommendation System Through Soil Analysis using Classification in Machine Learning, International Journal of Advanced Science and Technology. 29(3) (2020) 12738 – 12747.

[28] Y. J. N. Kumar, V. Spandana, V. S. Vaishnavi, K. Neha, and V. G. R. R. Devi, Supervised Machine Learning Approach for Crop yield Prediction in the Agriculture sector, In International Conference on Communication and Electronics Systems (ICCES), IEEE. (2020) 736-741.

[29] A. Anitha, and D. P. Acharya, Crop Suitability Prediction in Vellore District using Rough set on fuzzy Approximation Space and Neural network, Neural Computing and Applications, 30(12) (2018) 3633-3650.

[30] P. K. Priya, and N. Yuvaraj, An IoT Based Gradient Descent Approach for Precision Crop Suggestion using MLP, In Journal of Physics: Conference Series, IOP Publishing. 1362(1) (2019) 012038.

[31] M. Garanayak, G. Sahu, S. N. Mohanty, and A. K. Jagadev, Agricultural Recommendation System for Crops using Different Machine Learning Regression Methods, International Journal of Agricultural and Environmental Information Systems (IJAEIS), 12(1) (2021) 1-20.

[32] T. Ashok, and P. Suresh Varma, Crop Prediction Based on Environmental Factors using Machine Learning Ensemble Algorithms. In Intelligent Computing and Innovation on Data Science, Springer, Singapore. (2020) 581-594.

[33] G. Murugesan, and B. Radha, Soil Data classification using Attribute group Rank with Filter-based Instance Selection model, International Journal Of Scientific & Technology Research. 9(6) (2020) 202-208.

[34] P. S. Ramila Rajaleximi, M. S. Irfan Ahmed, Ahmed Alenezi, Classification of Imbalanced Class Distribution using Random Forest with Multiple Weight-based Majority Voting for Credit Scoring, International Journal of Recent Technology and Engineering. 7(6S5) (2019) 517-526.

[35] S. Sathya Bama, and A. Saravanan, Efficient Classification using Average Weighted Pattern Score with Attribute Rank Based Feature Selection, International Journal of Intelligent Systems and Applications, 11(7) (2019) 29.

[36] E. Winarno, W. Hadikurniawati, and R. N. Rosso, Location-Based Service for Presence System using Haversine Method, in International Conference on Innovative and Creative Information Technology (ICITech), IEEE. (2017) 1-4.

[37] P. Royston, Multiple Imputations of Missing Values, The Stata Journal. 4(3) (2004) 227-241.

[38] G. Chhabra, V. Vashisht, and J. Ranjan, A Comparison of Multiple Imputation Methods for data with Missing Values, Indian Journal of Science and Technology. 10(19) (2017) 1-7.

[39] D. R. Sumner, Crop Rotation and Plant Productivity, In CRC Handbook of Agricultural Productivity, CRC Press. (2018) 273-314.

[40] P. Benincasa, G. Tosti, M. Guiducci, M. Farneselli, and F. Tei, Crop Rotation as a System Approach for soil Fertility Management in Vegetables, Advances in Research on Fertilization Management of Vegetable Crops. (2017) 115-148.

[41] D. B. Lobell, W. Schlenker, and J. Costa-Roberts, Climate Trends and Global Crop Production since 1980, Science. 333(6042) (2011) 616-620.

[42] D. Devi, S. K. Biswas, and B. Purkayastha, Early Detection of Parkinson's disease: an Intelligent Diagnostic Approach, in Research Anthology on Diagnosing and Treating Neurocognitive Disorders, IGI Global. (2021) 295-328.

[43] Soil Health Management, Soil Health Card Portal, Department of Agriculture, Cooperation and Farmers Welfare under Ministry of Agriculture and Farmers Welfare, Government of India. https://soilhealth.dac.gov.in/

[44] P. L. Patil, B. I. Bidari, M. Hebbara, J. Katti, S. Dilvaranaik, S. Vishwanatha, H. M. Geetanjali, and G. S. Dasog, Identification of Soil Fertility Constraints by GIS in Bedwatti sub-watershed Under the Northern dry zone of Karnataka for Site-Specific Recommendations, Journal of Farm Sciences, 30(2) (2017) 206-211.

[45] P. R. Chaudhari, N. H. Desai, P. P. Chaudhari, and K. V. Rabari, Status of Chemical Properties and Available Major Nutrients in Soils of Patan District of Gujarat, India, Crop Research, 53(3/4) (2018) 147-153.

[46] D. E. Ratnawati, W. Marjono, and S. Anam, Comparison of Activation Function on Extreme Learning Machine (ELM) Performance for Classifying the Active Compound, in AIP Conference Proceedings, AIP Publishing LLC. 2264(1) (2020) p. 140001.

[47] R. Batuwita, and V. Palade, Adjusted Geometric-mean: a Novel Performance Measure for Imbalanced Bioinformatics Datasets is Learning. Journal of Bioinformatics and Computational Biology, 10(4) (2012) 1250003.