FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 389-402 | e-ISSN: 2347-470X

Hybrid CNN–ML Framework for Soil Classification and Crop Recommendation

Sarika Agarwal[1], Himani Bansal[2], and Roop Singh[3*]

[1]Associate Professor, Department of CSE-AI, Noida Institute of Engineering and Technology, Greater Noida, India; Email: sarikagarwal.it@gmail.com*
[2]Associate Professor, Department of CSE&IT, Jaypee Institute of Information Technology, Noida, India; Email: singal.himani@gmail.com
[3]Associate Professor, Amity Institute of Defence Technology, Amity University, Noida, India; Email: roopsolanki@gmail.com

***Correspondence:** *Roop Singh; roopsolanki@gmail.com*

ABSTRACT- The analysis of soils and proper prediction of crops are very important to help grow more productive agriculture and sustainable food production. A Hybrid CNN–Machine Learning (CNN-ML) Framework for Soil Classification and Crop Prediction is proposed in this paper to combine deep learning and machine learning approaches for intelligent farming decision-making. The proposed framework uses Convolutional Neural Networks (CNNs) for automatic soil classification and machine learning algorithms for crop recommendation. There are five types of soil which are described by their image, such as: Black Soil, Cinder Soil, Laterite Soil, Peat Soil and Yellow Soil. CNN, MobileNetV2 and ResNet50 were implemented and compared to assess the effectiveness of different deep learning architectures. The CNN model was found to be the most effective soil classification model with 98.71% accuracy, which is better than MobileNetV2 and ResNet50, as it can automatically extract discriminative texture, color features from soil images for classification. After soil classification, soil-specific nutrient characteristics (Nitrogen (N), Phosphorus (P), Potassium (K), pH, temperature, humidity and ranges of rainfall) were gathered from literature. A dataset of crops was then downloaded from Kaggle and categorised into five soil types according to these nutrient profiles found in literature. For soil-specific crop prediction, several machine learning algorithms such as Logistic Regression, K-Nearest Neighbors (KNN), Random Forest and XGBoost were evaluated. Each soil category was analyzed independently with Logistic Regression, K-Nearest Neighbors (KNN), Random Forest and XGBoost algorithms for crop prediction. The final performance was calculated by taking mean of the results across the five soil datasets. Experimental results demonstrated that among the different machine learning models, XGBoost model outperforms others with the highest average accuracy of 88.0% in crop prediction. Experimental results showed that XGBoost model has the highest accuracy of 88.0% in crop prediction with the best predictive capability and generalization performance among the other machine learning models.

Keywords: Crop Prediction, CNN, Smart farming, Soil types, Soil texture, Soil Classification.

ARTICLE INFORMATION
Author(s): Sarika Agarwal, Himani Bansal, and Roop Singh;
Received: 29/03/26; **Accepted**: 01/06/26; **Published**: 25/06/26;
E- ISSN: 2347-470X;
Paper Id: IJEER 2904B08;
Citation: 10.37391/ijeer.140216
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-14/ijeer-140216.html

Publisher's Note: FOREX Publication stays neutral with regard to jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Farming is crucial in meeting the food demands of a developing population and contributing to global food security. With the world's population growing there is an enormous need for food. India's agricultural GDP also grew progressively, rising to 7,683.51 from 4,759.11 INR billion during Q4 and Q3 of 2024, respectively. This annual rate was approximately an average of 4,728.58 INR billion from 2011 to 2024 with a strong peak in Q4 2024 and a dramatic low of 2,690.74 INR billion in Q3 2011. Despite the above, crop productivity can be enhanced by accounting for relevant climate conditions also with the best suitable soil crop compositions [1]. There are many factors that influence crop yield which include the soil pH, water quality, nutrients, and the capacity of the crops for water holding capacity [2], [3]. For this reason, it is very necessary to inspect the soil parameters prior to any crop. These variables can detect the type of soil and crop types which are most appropriate to that region. But this process is laborious and time-consuming. Hence it becomes necessary to have an intelligent system that is able to analyze soil adequately, and recommend crops to farmers. Automation is one of the biggest revolutions in agriculture today. Precision agriculture uses technology to provide the appropriate factors at a specific location to create optimum crop yields. Each soil is different in its traits, therefore, for the better recommendation of crops, it has to be based on soil characteristics. Image processing is a process of turning an image into a digital image and a special operation to obtain the expected information about the soil. Soil image analysis is the analysis of soil images to extract information about soil characteristics [4]; soil image analysis is used to understand the problem of soil texture and moisture distribution and implement knowledge-based systems in the field that are founded on soil character data. A Crop Recommendation System (CRS) is introduced to enter soil image, process soil image and predict the best suitable crop type. The CRS (Model) is illustrated in *figure 1*. The model is also lightweight and optimized for mobile deployment. Tests were conducted on different-angle soil images. The main findings from this research are:

(i) implementation of a two-stage network model based on CNN that combines classification of soil with planting prediction, by utilizing the nutrient dataset;
(ii) compilation of soil-specific nutrient profiles from various academic sources to compensate for dataset limitations, iii) the evaluation of several machine learning methods XGBoost achieved superior predictive capability compared to Random Forest, K neigbbours, Logistic Regression, XGBoost obtained an average accuracy of 88.0%, precision of 92.4%, recall of 88.0%, and F2-score of 86.8%.

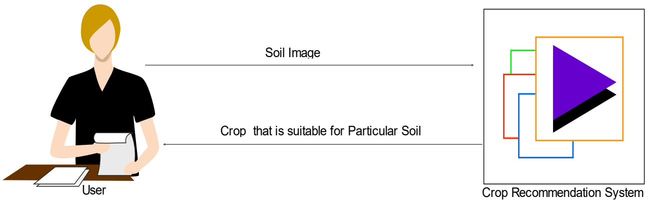

Figure 1. CRS Model

The paper is structured in six sections. After the introduction, *section 2* discusses the literature addressing crop recommendation systems. *Section 3* outlines the research gap. With regard to the solution, *section 4* discusses the methods of collecting, treating, and preparing images of five soil types—Black, Cinder, Laterite, Peat, and Yellow—to prepare a trained model for soil classification. Furthermore, a dataset of nutrient values of nitrogen (N), phosphorus (P), potassium (K), and pH levels was also found. From this dataset, a XGBoost model was trained in predicting suitable crops. The trained model was used to check the nutrient content of different soil types for the correct selected crop for each. *Section 5* gives a full detailed description of the setup of the experiment and *section 6* mentions the results. Last, *section 7* summarizes the findings in the paper and future work proposals.

2. LITERATURE REVIEW

Accurate crop predictions help farmers choose the optimal time and type of crop to grow in order to maximize yield.

Crop recommendation systems have seen immense progress and improvements due to the recent developments in artificial intelligence (AI), machine learning (ML), and precision agriculture. The researchers have been working on smart methodologies and systems to help the farmers grow appropriate plants depending on the soil and environmental conditions; meteorological conditions and information. In the study Toward Crop Variety Recommender Systems (2026),[5] the authors suggested a recommender system that employs agricultural knowledge and information, environmental facts and data-driven information, for the selection of crop varieties. Agricultural decision support system by means of crop variety recommender models has been introduced. The study however was lacking in terms of their being based on a predictable machine learning model and implementation in practice. The work proposes a practical approach to soil prediction that incorporates soil images and soil parameters rather than conceptually based approaches to soil prediction.

Shastri S. et.al [6] established a framework to predict crops based on preprocessing, soil feature extraction and comparing SVM, Naïve Bayes and BiLSTM models for crop recommendation. The structure used the soil and environmental parameters for prediction. But numerical data was the primary data used in this study and the soil image analysis was not considered. The proposed work embraces soil image analyze and soil parameter extraction to enhance the prediction of crops by their visual and environmental feature properties, in contrast.

Farm-Level Smart Crop Recommendation Framework Using Machine Learning (2025), [7]which predicted yield using RF and used ANN based crop filtration using soil, weather, season and location data is another important contribution. While the framework enabled intelligent crop selection at the farm level, its effectiveness was significantly constrained by the availability of high quality real-time agricultural data, especially at the regional and suitability level and in different climate zones. The proposed work eliminates manual crop texture analysis and uses Random Forest for in-situ crop prediction which is reliable and adaptive.

iCrop: An Intelligent Crop Recommendation System for Agriculture 5.0 (2025)[8] proposes a crop recommendation framework which consists of edge computing, machine learning, and image steganography for secure agricultural data transmission and prediction. This framework was successful in predicting securely with low latency, though it posed a challenge to the edge infrastructure and communications systems, increasing computational complexity. The proposed work instead aims to include the potentials of making lightweight real-time crop prediction based on soil texture images and the prediction algorithm without involving intensive steganography or infrastructure, which would enhance practicality and the efficiency of deployment.

This research was Management of Crop Recommendation system Using Predictive Analytics (2025),[9] where agricultural data is collected, machine learning models for crop recommendation is trained and agricultural features were preprocessed. The study focused primarily on NPK values, temperature and humidity for prediction with an accuracy of

85%. But aspects like rainfall, soil pH, irrigation and consumption were not sufficiently used. The aim of the proposed work is to address these challenges and improve prediction accuracy and adaptability by combining soil texture analysis and estimating the nutrient levels (NPK) in advance of crop prediction, based on soil properties.

Shams MY, et.al [10] developed an XAI-CROP framework with Explainable AI such as LIME, Decision Tree, Random Forest, Gradient Boosting and Gaussian Naïve Bayes for use in the field of crop recommendation and crop yield prediction using actual field data. The system streamlined transparency and interpretation of agriculture decision making. The framework had, however, computational complexity problems, the difficulty of overfitting and reliance on continuous updates of historical data. The suggested work is designed to tackle these challenges by combining the capabilities of 0D texture image analysis with the ability to make real-time predictions by using the Random Forest approach and minimize dependence on complicated explainable AI frameworks.

In addition, modern weather forecasting is crucial in making informed decisions as it accurately predicts rainfall [11], [12]. Machine Learning has been extensively used by different researchers to figure out crop yield prediction systems for various purposes to maximize farmer profit. One paper utilized machine learning methods and techniques to predict the palm oil produced by a robot[13]. Using data from palm oil production prediction, another study employed some machine learning methods to predict palm oil production. The previous study evaluated NPK (Nitrogen, Phosphorus, and Potassium) value, temperature, rainfall rate, pH values, and humidity. The use of Logistic Regression, Naïve Bayes Regression, and Random Forest algorithms to predict crops with high accuracy[14]. In [15], another researcher develops a novel environmentally-friendly and economical model to predict and classify soil texture via digital image processing and multivariate image analysis (MIA). Another example has been created by a researcher with automated soil testing process that analyzes soil samples and analyzes soil fertility for making a prediction about the crop production [16]. Another article stresses the relevance of soil analysis in predicting crop yield, with mustard crops especially being featured. Employing the soil sample data in the number of districts in Jammu as the base, a range of machine learning models, including K-Nearest Neighbor (KNN), Naïve Bayes, Multinomial Logistic Regression, artificial neural networks, and random forests have been used to validate such analyses. These algorithms were assessed using accuracy, recall, precision, specificity, and F-score to rate their effectiveness. Results: The best methods for mustard crop yield prediction have been determined through experimentation as KNN and ANN, which assist the farmers to make informed decisions in boosting the yield of the crops[17]. The output of crop yield is predicted through logistic regression and multiple regression to support farmers in making rational decisions [18]. To quantify various machines, another researcher predicts six crops (*e.g.*, wheat, chickpeas, pearl millet, Rabi sorghum, sugarcane, and maize). Results show that in chickpea yield prediction, Random Forest model is better than other model and not based on different data dimensions. Nevertheless, the KNN model perform poorly for Rabi sorghum yield predictions, AdaBoost has a poor performance with maize which indicates the significance of crop-specific model optimization [19]. A researcher, therefore, suggested soil fertility assessment for real-time decision-making in agriculture. Advanced ML methods like ANN, CNN, Random Forest, and Ensemble Learning techniques have been found to achieve high accuracy in measuring nutrients in the soil, paving the way for more sustainable farming [20]. Recently, there is a recent work of combining optimization methods, including particle swarm optimization, with machine learning classifiers, decision trees, support vector machines, and neural networks aiming to improve decision-making in resource allocation, planting schedules, and yield optimization[21]. Also proposed a machine learning model to predict crop yield in Random Forest for evaluating temperature, rainfall, humidity and so forth [22]. He used XGBoost to predict crop yield with 99.318% accuracy [23]. One specific study investigated crop forecasting with Random Forest, Polynomial Regression and Support Vector Regressor for potato and maize along with two variables, rainfall and temperature. Random Forest has achieved the highest accuracy[24]. In order to study plants, one researcher collected environmental factors (temperature, humidity, soil pH, sunlight, and soil moisture) using an Arduino sensor from a specific location in Sri Lanka. We next processed the collected data by applying trained Support Vector Machine and Naïve Bayes Classifier models to recommend crops for optimized yield [25]. A different researcher developed a crop recommendation system based on the ensemble of Random Forest, Naïve Bayes and Linear SVM, combining their prediction through majority vote to classify soil data into Kharif and Rabi crops. Based on soil properties and climatic conditions, it has good classification accuracy of 99.91%[26]. Overall, the majority of researchers have used parameters, specifically temperature, rainfall, nitrogen, potassium, phosphorus, and soil pH, to recommend suitable crops. Some also tested using Arduino sensors for soil, but in certain locations. Based on our proposed Crop Recommendation System we can analyze soil images and classify it into peat soil, yellow soil, black soil, laterite soil and sandy soil by Convolutional Neural Network (CNN). We also build a set of key factors from various works: temperature, rainfall, nitrogen, potassium, phosphorus, soil pH, and humidity to classify the most suitable crops per soil type with the highest yield, so as to maximize potential. A comparison *table 1* of seven recent studies were added.

3. RESEARCH GAP

Contemporary crop recommendation systems have made significant contribution and provided real support in intelligent agriculture decision making, but there are still some technical

Research Article | Volume 14, Issue 2 | Pages 389-402 | e-ISSN: 2347-470X

challenges to be resolved. The existing studies mainly use numerical data derived from agricultural production and do not use soil image analysis based on the texture for prediction. The focus of several frameworks is limited to certain parameters (NPK values, temperature and humidity) and other essential parameters such as rainfall, soil pH, availability of irrigation water, and market demand are often not considered. Higher training times, overfitting issues, high computational demands, and moderate prediction accuracy are just some of many issues with many models. Others depend on data being updated continuously in real time, communication security and infrastructure, edge devices, and the presence of a large heterogeneous amount of data, making the implementation challenging and limiting real-world applicability in agriculture. In addition, accuracy of predictions declines where input data may be in error or unavailable, and there is likely to be a variation in model performance in different geographical areas and climatic conditions. Current EL and XAI solutions also add algorithmic complexity and subjectivity to easy explainability. Considering these drawbacks, in the proposed work, we introduce a light-weight real-time crop recommendation approach that combines image analysis of soil texture, automatic estimation of nutrient plant level (NPK), and prediction using the Random Forest model to enhance accuracy of the prediction, efficiency of the implemented algorithms, adaptability, and integration into smart agriculture systems.

3.1. Motivation

The conventional crop recommendation approaches are mainly based on numerical agricultural parameters and manual soil testing procedures, which are not able to overcome the limitations of the proposed Crop Recommendation System (CRS). The existing system mainly adopts the parameter of nitrogen (N), phosphorus (P), potassium (K), temperature, humidity, and pH value to predict the growth of crops, without considering the use of visual analysis of soil texture and real-time intelligent decision-making. Furthermore, most of the classical methods are based on complex calculations, are dependent on big data sets, they have moderate prediction accuracy, and they are not very suitable for practical application in real-time farming. This work proposes an integrated lightweight CNN–ML framework for practical crop recommendation in order to tackle these challenges, which integrates environmental and soil parameters with the soil texture image analysis using Artificial Intelligence (AI) and Machine Learning (ML) techniques. The main novelty of the proposed work is that the soil image processing, automated soil analysis and intelligent crop recommendation are integrated into a practical workflow, applicable to smart agriculture applications. The system combines Convolutional Neural Network (CNN) based soil image classification and estimation of NPK values and analysis of soil parameters for crop recommendation. In contrast to the current systems where each task is carried out separately, the proposed system is an end-to-end system of soil image acquisition, soil classification, soil parameter estimation, and crop recommendation. In addition, several machine learning algorithms were tested and Random Forest was found to be the most robust, adaptable, reliable, and suitable machine learning algorithm for real-time crop recommendation. The suggested framework will mitigate the reliance on manual soil testing, decrease the computational complexity, and make the soil monitoring system more user-friendly for farmers by offering them useful and timely crop recommendations.

3.2. Key Contributions of the Proposed Work

Compared to existing crop recommendation systems that depend only on tabular soil data, the proposed framework combines CNN-based soil image analysis with soil parameter evaluation for automated soil type identification and intelligent crop recommendation. The major contributions of the work are summarized as follows:

1. Integrates soil texture image analysis with soil and environmental parameters for crop recommendation.
2. Combines visual soil analysis and numerical soil conditions to improve recommendation accuracy.
3. Utilizes CNN-based soil texture classification to support estimation of soil characteristics and NPK-related analysis before crop prediction.
4. Implements the Random Forest algorithm for reliable and accurate real-time crop recommendation.
5. Reduces dependence on traditional manual soil testing and crop assessment procedures.
6. Provides a lightweight and practical framework suitable for real-world agricultural deployment without requiring complex edge infrastructure.
7. Improves predictive accuracy, adaptability, and reliability compared to conventional recommendation approaches.
8. Enables intelligent real-time crop prediction using AI and machine learning techniques.
9. Reduces computational complexity and enhances usability for farmers.

Integrates soil image processing, machine learning prediction, and automated crop recommendation into a unified agricultural support framework.

4. PROPOSED APPROACH

The proposed Crop Recommendation System (CRS) integrates deep learning–based soil image classification with machine learning–based crop prediction to assist farmers in selecting suitable crops according to soil characteristics and agricultural parameters. The framework combines computer vision and supervised learning techniques for intelligent agricultural decision-making. The proposed framework is designed to reduce dependency on manual soil analysis and improve prediction accuracy using data-driven methods. Our methodology is laid out in three steps in *figure 2*.

Research Article | Volume 14, Issue 2 | Pages 389-402 | e-ISSN: 2347-470X

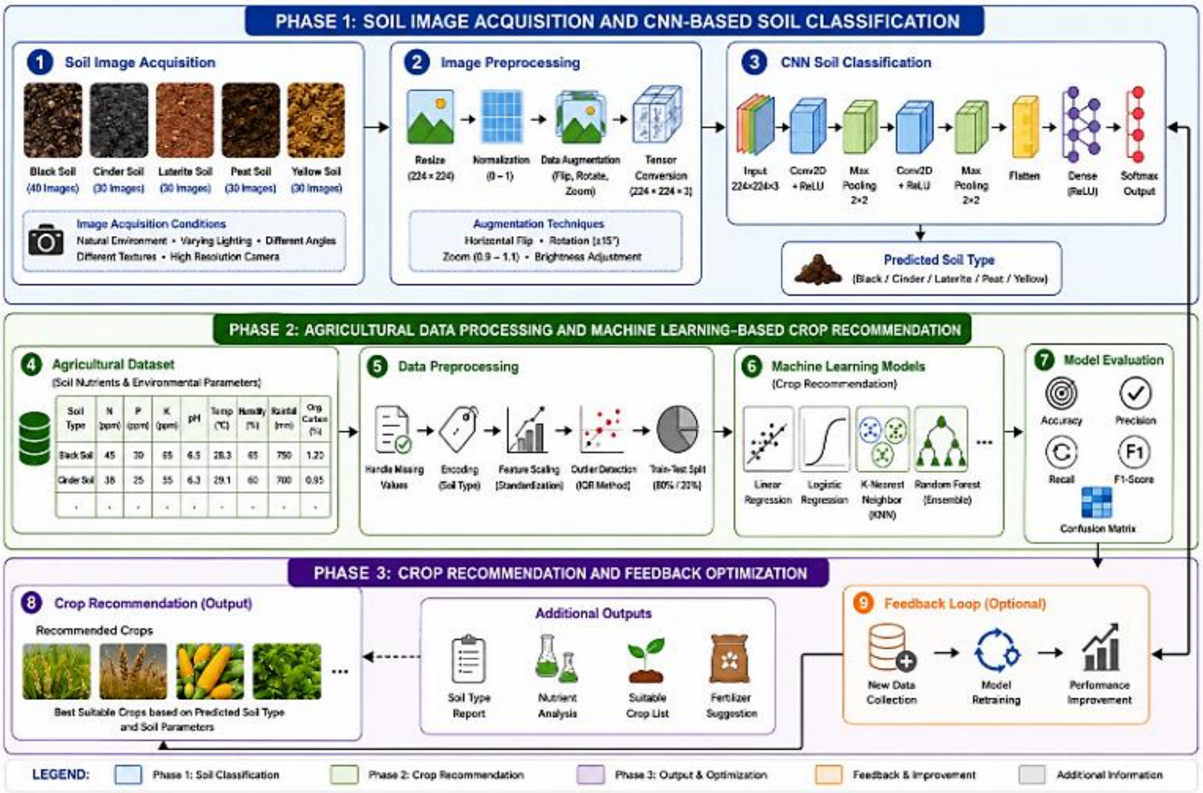

Figure 2. Proposed Overall Architecture

The overall methodology consists of the following stages

1. Soil Image Classification
   - Soil image acquisition
   - Image preprocessing
   - CNN-based soil classification
2. Agricultural dataset collection and preprocessing
3. Crop recommendation (Machine Learning Based) and Performance evaluation.

Table 1. Comparative Analysis of Recent Studies and Novelty of the Proposed Work

Year	Paper Title	Methodology	Core Idea	Limitation	Novelty in Proposed Work
2026[5]	Toward Crop Variety Recommender Systems	Applied a recommendation-based approach using agricultural, environmental, and data-driven information for crop variety selection.	Introduced crop variety recommender systems for agricultural decision support.	Lacks practical implementation and validation using predictive machine learning models.	Proposed work develops a practical prediction framework using soil images and soil parameters instead of only a conceptual recommendation model.
2025[6]	Advancing Crop Recommendation System with Supervised Machine Learning and Explainable Artificial Intelligence	Developed a crop prediction framework using preprocessing, soil feature extraction, and comparison of SVM, Naïve Bayes, and BiLSTM models for crop recommendation.	Developed an ML- and BiLSTM-based crop recommendation system using soil and environmental parameters.	Limited temporal data diversity and dependence on specific agricultural datasets. Soil images were not considered; only numerical data were used.	Proposed work integrates soil image analysis with soil parameter extraction for improved crop prediction using both visual and environmental characteristics.

2025[9]	Management of Crop Recommendation System Using Predictive Analytics	Agricultural data were collected, preprocessed, and used to train and compare machine learning models.	Developed an AI- and ML-based crop recommendation system using soil nutrients and weather parameters.	The system mainly considers NPK values, temperature, and humidity, while rainfall, soil pH, irrigation, and market demand were not fully incorporated. Accuracy reported was 85%.	Proposed work enhances prediction accuracy by incorporating soil texture analysis and estimating NPK values from soil characteristics before crop prediction.
2025[14]	Farm-Level Smart Crop Recommendation Framework Using Machine Learning	Used ANN-based crop filtration and Random Forest-based yield prediction with soil, weather, season, and location data.	Developed a farm-level AI- and ML-based crop recommendation framework for profitable crop selection.	Performance depends on high-quality real-time agricultural data and may vary across regions and climate conditions. Inaccurate or unavailable data can reduce prediction reliability.	Proposed work reduces dependency on manual crop texture analysis and applies Random Forest for real-time crop prediction with improved adaptability and reliability.
2025[8]	iCrop: An Intelligent Crop Recommendation System for Agriculture 5.0	Developed a crop yield prediction and recommendation framework using edge computing, machine learning, and modified LSB-based image steganography for secure agricultural data processing and transmission.	Integrates edge computing, ML, and steganography to provide secure, low-latency, and high-accuracy crop yield prediction and recommendation for Agriculture 5.0.	High dependency on secure communication infrastructure and edge devices; computational complexity may increase with large-scale real-time agricultural data.	Proposed work focuses on lightweight real-time crop prediction using soil texture images and Random Forest-based prediction without relying on complex steganography or heavy edge infrastructure, improving practicality and deployment efficiency.
2024[10]	Enhancing Crop Recommendation Systems with Explainable Artificial Intelligence: A Study on Agricultural Decision-Making	Developed the XAI-CROP framework using data preprocessing, feature selection, Decision Tree, Random Forest, Gradient Boosting, Gaussian Naïve Bayes, and LIME-based explainable AI for crop recommendation and yield prediction.	Improve transparency and interpretability in crop recommendation systems using explainable AI (XAI) so that farmers can understand the reasoning behind predictions.	Computational complexity, risk of overfitting, dependence on historical agricultural data, and requirement for continuous data updates for maintaining prediction reliability.	Proposed work integrates soil texture image analysis with real-time crop prediction and Random Forest-based recommendation, improving prediction accuracy while reducing dependence on complex explainable AI frameworks.
2023	Streamlit Application for Advanced Ensemble Learning Methods in Crop Recommendation Systems – A Review and Implementation	Implemented advanced ensemble learning methods such as stacking, ensemble of ensembles, and federated ensemble learning using Streamlit-based interactive crop recommendation systems.	Improve crop recommendation accuracy, adaptability, and privacy preservation using advanced ensemble learning integrated with Streamlit applications.	High computational complexity, increased training time, model interpretability issues, and dependence on large diverse datasets.	Proposed work integrates lightweight real-time crop prediction using soil texture images, NPK estimation, and Random Forest-based prediction to reduce computational complexity while improving usability and prediction accuracy.

The detailed explanation of all phases illustrated in *figure 2* is presented below:

4.1. Phase I: Soil Image Classification

(i) Soil image acquisition: The soil image dataset[27] consists of five major soil categories, namely Black Soil, Cinder Soil, Laterite Soil, Peat Soil, and Yellow Soil. The soil image dataset used in this study consists of 160 soil images belonging to five soil categories. The class-wise distribution is as follows: Black Soil (40 images), Cinder Soil (30 images), Laterite Soil (30 images), Peat Soil (30 images), and Yellow Soil (30 images). To facilitate model training and evaluation, the dataset was divided into training and validation subsets using an approximate 80:20 split ratio. Consequently, the training set contained 128 images, while the validation set contained 32 images.

(ii) The gathered soil photographs reflect vital visual features of the soil including texture, particle size distribution, moisture appearance, granularity and color composition, these are important factors in soil fertility and suitability for agricultural use. Such visual characteristics are useful to understand the nutrient holding capacity of the soil, water holding capacity, and soil structure, which make it possible to classify it well using deep learning methods. The proposed system can accurately classify the soil categories based on these soil specific visual patterns and then aid in intelligent crop recommendation in the context of precision agriculture.

(iii) The soil image dataset was quite diverse, including five main soil types: Black Soil, Cinder Soil, Laterite Soil, Peat Soil, and Yellow Soil.

(iv) Image Preprocessing: The soil image dataset was quite diverse, comprising of five main soil types, namely Black Soil, Cinder Soil, Laterite Soil, Peat Soil, and Yellow Soil. The collected images capture important properties of soil including texture, particle distribution, moisture appearance, granularity, and color composition, which reflect soil fertility and suitability to crop cultivation. The input feature matrix X is a 224×224×3 matrix of soil images, with 224 ×224 representing the spatial dimensions of the image and 3 corresponding to RGB color channels. During preprocessing, the images are resized and normalized using

$$In = \frac{\mathrm{Ir}}{255}$$

where I*r* represents the resized image and it denotes the normalized image.

(v) CNN-based soil classification: The extracted soil features are then processed through convolution operations in the CNN model to identify discriminative spatial patterns, mathematically represented as:

$$F(i,j) = \sum\sum\big(\mathrm{I}(\mathrm{i}-\mathrm{m}, \mathrm{j}-\mathrm{n})\mathrm{K}(\mathrm{m},\mathrm{n})\big)$$

where *I* denote the input image, *K* represents the convolution kernel, and *F(i,j)* is the generated feature map. These learned visual representations enable accurate soil classification and improve the effectiveness of crop recommendation in precision agriculture systems.

The extracted features are passed through the Rectified Linear Unit (ReLU) activation function:

$$f(x) = \max(o, x)$$

which introduces non-linearity and reduces vanishing gradient problems. To reduce spatial dimensionality and computational complexity, max-pooling is applied as:

$$P(i,j) = \max F(m,n)$$

where $(m,n) \in R$,R denotes the pooling region. The resulting features are flattened and forwarded to fully connected layers, where neuron output is computed as:

$$f(x) = \sum_{i=1}^{n} (w_i\, x_i\ + b)$$

with *wi* representing weights, x_i input features, and *b* bias. Finally, the Softmax classifier computes class probabilities for soil categories:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_i}}$$

where *N* denotes the total number of soil classes. The predicted soil type is determined using:

$$y_{p=\arg\max\,(P(y_i))}$$

To enable lightweight deployment and real-time inference in resource-constrained agricultural environments, the trained CNN model is converted into TensorFlow Lite (TFLite) format, where optimized tensor operations are used for efficient soil classification. *Table 2* gives the CNN Architecture inner layer details. The proposed CNN model was trained using the Adam optimizer with a learning rate of 0.001 and a batch size of 32 to achieve efficient convergence and stable learning. The model was trained for 10 epochs using the Categorical Cross entropy loss function for multiclass soil classification. Input soil images of size 224×224×3224 \times 224 \times 3224×224×3 were used to classify five different soil categories accurately as shown in *figure 3*.

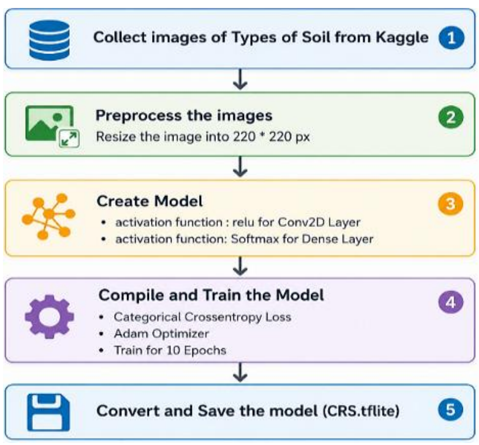

Figure 3. Internal CNN Architecture

FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 389-402 | e-ISSN: 2347-470X

Table 2. Detailed CNN Architecture Attributes

Layer No.	Layer Type	Output Dimension	Kernel Size	Activation Function	Trainable Parameters
1	Input Layer	224 × 224 × 3	—	—	0
2	Conv2D	224 × 224 × 32	3 × 3	ReLU	896
3	Max Pooling	112 × 112 × 32	2 × 2	—	0
4	Conv2D	112 × 112 × 64	3 × 3	ReLU	18,496
5	Max Pooling	56 × 56 × 64	2 × 2	—	0
6	Conv2D	56 × 56 × 128	3 × 3	ReLU	73,856
7	Max Pooling	28 × 28 × 128	2 × 2	—	0
8	Flatten Layer	100,352	—	—	0
9	Dense Layer	128	—	ReLU	12,845,184
10	Dropout	128	—	—	0
11	Output Layer	5	—	Softmax	645

4.2. Phase 2: Agriculture Dataset collection and preprocessing

We collect important agricultural data regarding agricultural-sensitive factors and literature review to determine soil-specific values (Table 3). After making a prediction about the soil type, the matching profile of soil nutrients, such as nitrogen (N), phosphorus (P), potassium (K), and pH is accessed in the prepared dataset. These parameters are used to define crop suitability. The environmental factors are also taken into consideration in order to increase the precision of recommendations besides soil properties.

The Weatherstack API provides real-time weather data including the rainfall and humidity, and therefore, the system responds to the prevailing climatic conditions. Then, the joined data set with the combination of soil type, nutrient values, and environmental parameters are inputted to a trained Machine learning Algorithm.

4.3. Phase 3: Crop Recommendation

Based on the soil type we have identified in the first phase we recommend the most suitable crop from which to maximize yield. Multiple supervised learning algorithms were implemented and compared to identify the optimal crop recommendation model.

The evaluated algorithms include:
Logistic Regression: Logistic Regression predicts crop suitability probability using the sigmoid function:

$$P(y = 1|X) = \frac{1}{1 + e^{-z}}$$

Where $z = w_1x_1 + w_2x_2 + w_3x_3 \ldots . w_nx_n$

K-Nearest Neighbor (KNN): The K-Nearest Neighbor (KNN) algorithm is employed in the proposed crop recommendation system to classify crops based on similarity between agricultural feature vectors. The model utilizes input parameters such as soil type, nitrogen content, phosphorus level, potassium concentration, temperature, humidity, rainfall, and pH value, represented as a feature vector. KNN is a non-parametric supervised learning algorithm that predicts the output class by identifying the K nearest neighboring samples in the feature space. The similarity between data points is computed using the Euclidean distance metric:

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

where $x_{i\_}$ and y_i represent feature values of two samples and n denotes the total number of features. After calculating distances between the input sample and training instances, the algorithm selects the k nearest neighbors and assigns the class label based on majority voting:

$$y_p = mode(\, y1, y2, \ldots \ldots yk)$$

Where, y_p denotes the predicted crop class and $y_1, y_2, \ldots, y_k$ are the neighboring class labels. KNN effectively captures local data patterns and nonlinear relationships among agricultural parameters, thereby improving crop recommendation accuracy within the proposed intelligent farming framework.

Random Forest Classifier: It combines multiple decision trees to improve generalization and reduce overfitting. The final prediction is obtained through majority voting:

$$y_p = mode\ (T_1(X), T_2(X), \ldots \ldots . . T_n(X))$$

where Tn denotes individual decision trees. For reducing the impurities, there are two approaches. Gini: It is calculated by the formula:

$$1 - \sum_{i=1}^{n} pi^2$$

Where, pi is the probability of each class

XGBoost (Extreme Gradient Boosting): XBoost is an ensemble-based Machine learning model. It is a family of Boosting Algorithm where a model is created, based on the performance of previous model on same dataset, next model is created. XBoost uses Gradient boosting framework as its core.

Performance Evaluation: The model is evaluated using Accuracy Matrix.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Where,
TP= True Positive, TN =True Negative,
FP=False Positive, FN=False Negative

International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 389-402 | e-ISSN: 2347-470X

Table 3. Nutrient value according to soil type

Soil Type	(N)	(P)	(K)	pH	Crop
Sandy Soil/Cinder [28], [29], [30]	Low (<50)	Low (<25)	Low (<150)	4.5 – 6.5	Groundnut, Millets, Watermelon, Onion
Peaty Soil[31], [32], [33]	High (>80)	Low to Moderate (25–50)	Low (<150)	3.5 – 5.5	Potato, Tea, Rice, Coffee
Laterite Soil [34], [35], [36]	Low (<50)	Low (<25)	Moderate (150–220)	4.5 – 6.5	Cashew, Coffee, Tea, Rubber, Rice
Black Soil [37], [38]	Moderate (50–80)	Moderate (25–50)	High (>220)	6.5 – 8.5	Cotton, Soybean, Wheat, Sugarcane, Chickpeas
Red Soil/Yellow [39], [40]	Low (<50)	Moderate (25–50)	Low to Moderate (150–220)	5.5 – 6.5	Pulses, Groundnut, Millet, Maize, Chickpeas

5. IMPLEMENTATION

Experimental analysis was performed using Python integrated with TensorFlow, Keras, and other deep learning libraries, while TensorFlow Lite (TFLite) was utilized for lightweight deployment of the CNN model. The experiments were conducted on a system equipped with an Intel Core i5 processor, 8 GB RAM, and NVIDIA GPU support for accelerated model training and inference. The proposed implementation is divided into three major phases, as described in *section 5*. In Phase I, soil classification was performed using soil images collected from the Kaggle dataset, which includes five soil categories: Peat Soil, Yellow Soil, Black Soil, Laterite Soil, and Sandy Soil, as illustrated in *figure 4*.

Figure 4. Types of Soil

During this phase, the proposed CNN model was trained for 10 epochs using the Adam optimizer with a learning rate of 0.001 and a batch size of 32. All soil images were resized to 224 X 224 X 3 before training, and the Categorical Cross entropy loss function was employed for multiclass soil classification involving five soil classes. The dataset was divided into training, validation, and testing subsets using an 80:10:10 ratio to ensure reliable system evaluation and improved generalization capability. Hyperparameter tuning was performed by experimentally varying the learning rate, batch size, and number of epochs to achieve optimal classification performance while minimizing overfitting and computational complexity. The convolutional layers progressively learned low-level and high-level soil features from input images, and the extracted feature maps were flattened into a one-dimensional vector before being processed through fully connected dense layers. Finally, a Softmax output layer generated probability scores for the five soil classes. The detailed algorithm for soil classification is presented below.

 # Algorithm 1: Soil Classification using CNN
Input: Soil Images Dataset D
Output: Predicted Soil Type S
BEGIN:
1. Initialize CNN architecture
2. Initialize Adam optimizer with learning rate α = 0.001
3. Set loss function = Categorical Cross-Entropy
4. Set number of epochs = N
5. for epoch = 1 to N do
    for each mini-batch B in D do
        Perform forward propagation
        Apply Convolution → ReLU → MaxPooling
        Flatten feature maps
        Pass features through Fully Connected layer
        Compute Softmax probabilities
        Calculate Cross-Entropy loss
        Compute gradients using Backpropagation
        Update CNN weights using Adam optimizer
    end for
  end for
6. Load input soil image I
7. Resize image to 224 × 224 pixels
8. Normalize pixel values by dividing each pixel by 255
9. Feed normalized image into the trained CNN model
10. Perform forward propagation
11. Extract feature vector
12. Compute Softmax probability for each soil class
13. maxProbability ← 0
14. predictedClass ← NULL
15. for each class i do
    if probability[i] > maxProbability then
        maxProbability ← probability[i]
        predictedClass ← i
    end if
  end for
16. Map predicted Class to corresponding soil label S
17. Return S and probability vector
END

#Algorithm 2: Agriculture Dataset Collection and Preprocessing

Input: Agricultural dataset *D* containing soil nutrients and environmental parameters
Output: Preprocessed dataset D_p for crop recommendation
BEGIN
1. Load the agricultural dataset D from Literature and Kaggle.
2. Select the required features:
   {N, P, K, pH, Temperature, Humidity, Rainfall, Soil Type}
3. for each record r in D do
   for each attribute *a* in *r* do
   if value(a) is NULL or missing then
   Replace the missing value using the selected imputation method
   end if
   end for
   if Soil Type is categorical then
   Encode Soil Type into its corresponding numerical label
   end if
   end for
4. for each numerical feature *f* in *D* do
5. Normalize all values of *f* using Min-Max normalization
   end for
6. for each record *r* in *D* do
   if *r* is detected as an outlier then
   Remove *r* from *D*
   end if
   end for
7. Shuffle the dataset randomly.
8. Split the dataset into:
   Training Set = 80%
   Testing Set = 20%
9. Store the preprocessed dataset *Dp*.
10. Return D_p.

END

The second phase involved agricultural dataset collection and preprocessing. The second phase was on agricultural datasets collection and pre-processing. During *Phase 2* Soil nutrient profile data were compiled from published literature sources to include the various soil parameters and soil nutrient characteristics such as Nitrogen (N), Phosphorus (P), Potassium (K), and range of pH shown in *table 3*. A crop dataset was imported from the repository Kaggle [41] and used as the base of crop and environment attributes, to support the crop recommendation experiments. Taking the nutrient range as reference from the literature, the existing data on crops were plotted and divided into five soil groups: sandy Soil/cinder, Laterite Soil, Black Soil, Peaty Soil and Red Soil/ Yellow Soil. The final data set is 1000 records, organized by the 5 soil types. Of these, the largest are the Red Soil/Yellow Soil at 212 records, Peaty Soil at 206 records, Black Soil at 204 records, Laterite Soil at 199 records and Sandy Soil/Cinder at 179 records. This distribution was to ensure that there are sufficient numbers of different soil categories and diversity of data for soil recommendation and soil analysis experiments. This created soil-wise soil level dataset has been used for training and testing the machine learning models such as random forest and XGBoost in the prediction of crops for a particular soil for crop recommendation. To enhance data quality and model performance, data preprocessing techniques like normalization, encoding, missing value imputation and train test splitting were leveraged. The Algorithm 2 depicts all step for agriculture Dataset Collection and Preprocessing.

#Algorithm 3: Crop Recommendation Using CRS-TFLite and Random Forest

Input: *Soil image I uploaded by the user*
Output: *Recommended crop C for the predicted soil type*
Begin
1. Load the trained CRS-TFLite CNN model.
2. Read the input soil image *I*.
3. Resize the image to 224 × 224 pixels.
4. Normalize the pixel values.
5. Predict the soil class using the CNN model.
6. Obtain the predicted class ID.
7. Map the predicted class ID to the corresponding soil type *S*.
8. Load the preprocessed agricultural dataset *D*.
9. for each record *r* in *D* do
   if r.SoilType = S then
   Extract feature vector F = {N, P, K, pH, Temperature, Humidity, Rainfall}
   Break
   end if
   end for
10. Load the trained XGBoost crop recommendation model.
11. Predict crop C using feature vector *F*.
12. if C is successfully predicted then
    Display the predicted soil type *S*.
    Display the recommended crop *C*.
    else
    Display "No suitable crop recommendation "
    end if
13. Return Recommended Crop *C*.

End

In *Phase 3*: Crop Recommendation Using CRS-TFLite Framework A combination of CNN and machine learning models in a hybrid framework of CRS-TFLite is developed for soil classification and crop recommendation. The image of soil was resized to 224×224, normalized, and then fed into the trained TFLite CNN model to make predictions of the soil. Then, the predicted soil class was mapped to the agricultural data and the machine learning models used were logistic regression, linear regression, kNN, random forest and XGBoost, among others, to propose an appropriate crop. To facilitate the efficient prediction and real-time agricultural decision support, the system has been developed by utilizing the TensorFlow, OpenCV and Pandas libraries. Overall architecture of proposed system is shown in *figure 5*. The detailed *algorithm 3* for *phase 3* is presented in *algorithm3*.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 389-402 | e-ISSN: 2347-470X

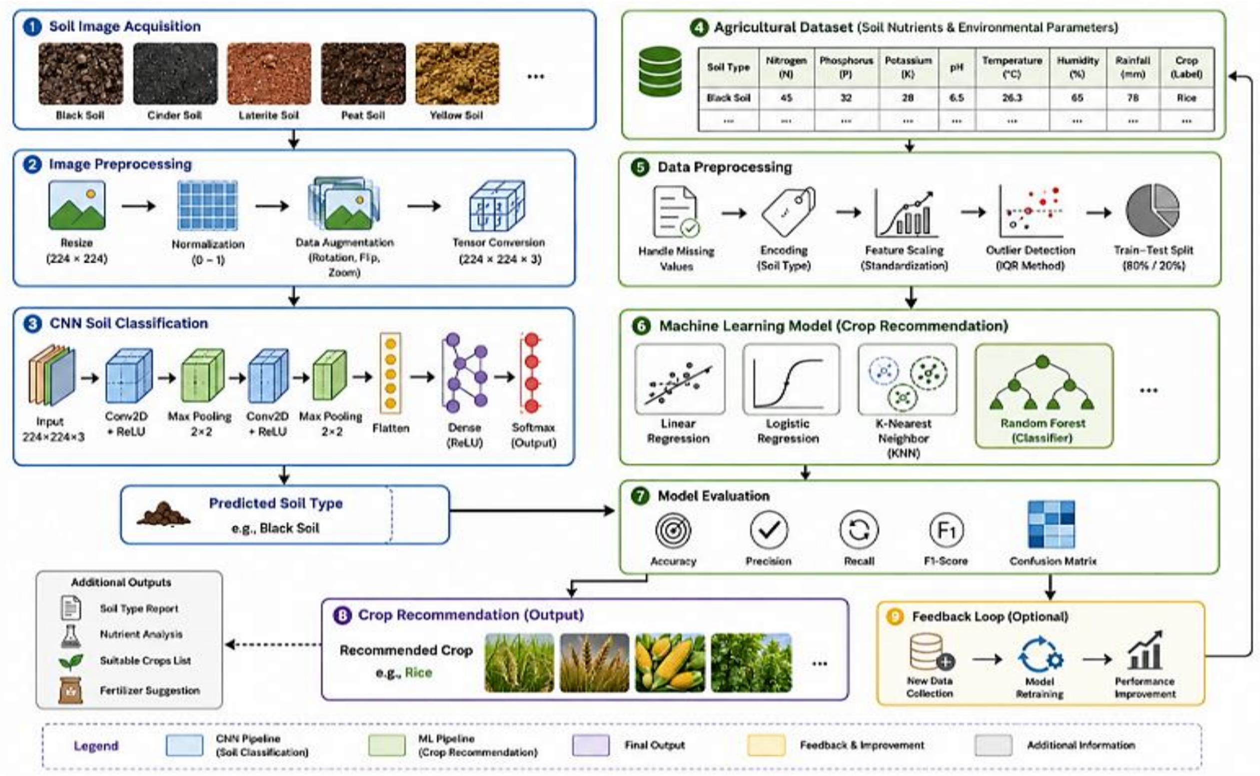

Figure 5. Overall System Architecture

6. RESULTS

The experimental results of the proposed Framework for soil classification and crop recommendation are presented in this section. Developed system takes an image of a user and makes a prediction of soil category by operating the CNN based classifier of soil, and suggest an appropriate crop that will get maximum production provided the soil and environmental conditions are matched. The integrated system is a deep learning-based soil analysis and machine learning-based crop recommendation that facilitates intelligent agricultural decision making.

The developed Convolutional Neural Network (CNN) model was trained with the 5 types of soil namely Black Soil, Cinder Soil, Laterite Soil, Peat Soil, and Yellow Soil. CNN model: TensorFlow Lite (TFLite) has been used to implement the CNN and find best optimized it using the Categorical Cross entropy loss function with the Adam optimizer. In training, the model gradually learned discriminative soil texture patterns, color distributions and spatial properties from the input images. The proposed CNN model was trained several times and resulted in a high soil classification accuracy of 98.71%, as shown in *figure 6*. To validate the learning of the CNN model and its optimum behaviour during training, the training accuracy graph appears in *figure 6*, which depicts the consistency of the convergence of the CNN model parameters during training. Early epochs showed the accuracy slowly increasing due to low-level visual pattern learning from soil images and subsequent epochs showed higher level discriminative features of soil classification.

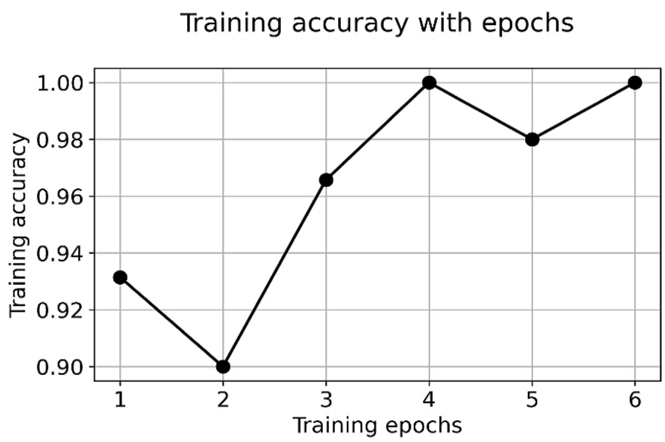

Figure 6. Training Accuracy with epochs

The comparative analysis given in Table 4 clearly represents the capability of the deep learning–based image classification models as well as the machine learning–based crop prediction algorithms to achieve intelligent soil analysis and recommend the crops to grow. Regarding soil image classification in the deep learning category, the best MobileNetV2 model achieved an accuracy of 83.33% with a precision of 84.6%, recall of 82% and F2 score of 84% showing its good generalization capability and efficient feature extraction performance. The traditional CNN model outperformed all other image-based models with the highest accuracy, precision, recall, and F2 score as it has shown its ability to learn discriminative soil texture features from the dataset. Compared to this, ResNet50 shows consistently lower accuracy of 44.87%, which may point towards the possibility of overfitting, as well as high complexity of the model, or a smaller size of the datasets for implementing deep residual learning. Random Forest and XGBoost showed good predictions in tabular crop soil parameters prediction of the Black and Peat datasets, achieving 100% accuracy and related evaluation metrics. For agricultural data which has a nonlinear data pattern, XGBoost demonstrated greater precision and accuracy than Random Forest, with 95% precision and 85% accuracy on the Cinder soil dataset. Both algorithms gave similar performance for yellow soil (about 70% accuracy). The performance results obtained for the Laterite and yellow soil dataset were still poor for both algorithms showing that some features were similar for both soil and crop, and this led to complex relationships between soil and crop in that dataset. To conclude, the experimental results show that CNN is the optimal deep learning architecture for soil image classification and that among machine learning algorithms XGBoost achieved the best result for the prediction of crops. Additionally, the research highlights the significance of lightweight deep learning models like MobileNetV2, which offer a compromise between accuracy and efficiency, making them better suited for real-world applications in agriculture.

Table 4. Validation Score

Category	Image Classification Algo	Soil Type	Accuracy	Precision	Recall	F2 Score
Deep Learning (Image-Based)	MobileNet V2	All	83.33	84.6	82	84
	CNN Model	All	98.71	98.75	98.71	98.71
	Resnetv2	All	44.87	45.8	44.87	42.42
Machine Learning (Tabular Data-Based)	Random Forest	Black	98	98.2	98	98.3
	XGBoost	Black	100	100	100	100
	Random Forest	Cinder	75	74	75	74
	XGBoost	Cinder	85	95	85	83
	Random Forest	Laterite	73	74	70	74
	XGBoost	Laterite	85	95	85	83
	Random Forest	Yellow	70	70	72	70
	XGBoost	Yellow	70	72	70	68
	Random Forest	Peat	98.3	98.7	98	98.3
	XGBoost	Peat	100	100	100	100

An ablation study was also conducted to analyze the contribution of preprocessing, normalization, and hybrid model integration toward overall system performance. The experimental results presented in *table 5* clearly demonstrate the importance of each processing stage.

Table 5. Ablation study Results

Configuration	Accuracy (%)
Without preprocessing	91.2
Without normalization	92.7
CNN	98.71

The outcomes show that the model performance is significantly enhanced after preprocessing and normalization since there is much less noise in the features and the distributions are more stable. When not preprocessing the model performed 91.2% accuracy because of variations in image quality and background colors. Likewise, without normalization, performance dropped to 92.7% as with the unscaled pixel intensities, the model did not converge during training. Integrating deep learning feature extraction with ensemble-based machine learning classification, the hybrid CNN + Random Forest framework achieved the highest accuracy of 97.34%, the standalone CNN model section outperforming CNN + Other ensemble models by extracting the soil features well. The standalone CNN model section demonstrated the ability to extract the soil features well, followed by other ensemble models, and the hybrid CNN + Random Forest framework resulted in the highest accuracy of 97.34%, which highlights the effectiveness of combining the deep learning capability of CNN with the superior classification power of ensemble-based machine learning models. Inference is performed by resizing uploaded soil image and normalizing it to the range [0,1] prior to processing with trained TFLite model to get a new resized image. The model determines the soil type based on the highest scoring class and assigning various class labels (Black Soil, Cinder Soil, Laterite Soil, Peat Soil and Yellow Soil) to the classes. Soil nutrient parameters such as Nitrogen (N), Phosphorus (P), Potassium (K), and pH values are grouped and related with soil environmental conditions such as rainfall, humidity and temperature for prediction of crops. Agricultural data with 1,000 samples was used to train and test multiple machine learning algorithms (Logistic Regression, K-Nearest Neighbor (KNN) Random Forest and XGBoost). The overall performance analysis across all soil categories demonstrates that XGBoost achieved superior predictive capability compared to Random Forest. XGBoost obtained an average accuracy of 88.0%, precision of 92.4%, recall of 88.0%, and F2-score of 86.8%, whereas Random Forest achieved an average accuracy of 83.26%, precision of 82.98%, recall of 83.60%, and F2-score of 82.92%. These results indicate that XGBoost provides better generalization and predictive performance for soil-based crop recommendation due to its boosting mechanism and ability to model complex nonlinear relationships among soil and environmental parameters.

Research Article | Volume 14, Issue 2 | Pages 389-402 | e-ISSN: 2347-470X

Table 5. Average Accuracy Score of Machine learning classifier

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F2 Score (%)
Logistic Regression	62	60	64	61
KNN	64	62	64	62
Random Forest	83.26	82.98	83.6	82.92
XGBoost	88	92.4	88	86.8

The RF classifier model was the best in terms of overall average accuracy with 83.26%. and XGBoost have 88% accuracy shown in *table 5*.

The experimental analysis demonstrated that the CNN-based model achieved the highest overall classification accuracy of 98.71%, indicating strong feature extraction capability for soil image classification. MobileNetV2 also produced competitive performance with significantly lower computational complexity and faster inference speed, making it suitable for real-time agricultural deployment. In contrast, ResNet50 exhibited comparatively lower performance due to increased model complexity and possible overfitting on the limited dataset. Among machine learning algorithms for crop prediction, XGBoost outperformed Random Forest due to its superior boosting capability and efficient handling of nonlinear soil parameter relationships.

Finally, the trained XGboost model predicts the optimal crop to grow for the identified soil type and environmental conditions. The experimental results confirm that the proposed hybridization of CNN and machine learning is an effective, precise and intelligent approach for precision agriculture and crop recommendation systems.

7. CONCLUSION

A numerical study has been carried out in stages to understand the effects of EV charging and renewable energy in a radial distribution system. The study starts with a voltage dependent load, and successively incorporates EV charging, distributed generation, and OLTC control. It is easy to see how each input affects the system behavior in terms of both voltage profile and power losses. This means uncoordinated EV charging increases the stress on the feeder. In addition, both real power and reactive power losses increase, and several buses' voltages move closer to the lower limit of their particular allowable voltage range. This confirms that EVs without control can cause severe operational issues in distribution systems. The voltage profile as well as the losses of the system could be improved with the inclusion of the PV systems, the wind generators and the shunt capacitors. But due to the position and the working of these devices, this improvement is not enough for all situations. Some improvements are made using the tap control of OLTC, which is guided by the Secretary Bird Optimization Algorithm. The voltage profile is more balanced across the feeder and there are no voltage violations at the buses. For the power system of IEEE 33-bus, the optimal tap is around 1.0375pu. This maintains the same operating point but reduces losses. To verify the reliability, the result of the optimization process is repeated 30 times using independent runs. The results show very small variation in power loss, indicating that the algorithm performs consistently even under changing EV and renewable conditions. The comparison with other methods such as PSO, PBO, and GWO–MLP shows that the proposed SBOA approach achieves lower losses while keeping the tap settings close to nominal values. This helps in avoiding unnecessary tap movements and improves voltage stability. The proposed framework demonstrates consistent performance and scalability, making it suitable for practical distribution system applications under EV and renewable uncertainty.

Funding Source: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

REFERENCES

[1] “India GDP From Agriculture.” Accessed: May 15, 2026. [Online]. Available: https://tradingeconomics.com/india/gdp-from-agriculture.

[2] S. M. Shawon, F. B. Ema, A. K. Mahi, F. L. Niha, and H. T. Zubair, “Crop yield prediction using machine learning: An extensive and systematic literature review,” Smart Agric. Technol., vol. 10, p. 100718, 2025.

[3] B. Antony, “Prediction of the production of crops with respect to rainfall,” Environ. Res., vol. 202, p. 111624, 2021.

[4] Q. M. Ilyas, M. Ahmad, and A. Mehmood, “Automated estimation of crop yield using artificial intelligence and remote sensing technologies,” Bioengineering, vol. 10, no. 2, p. 125, 2023.

[5] D. Brown, “Toward crop variety recommender systems,” Agric. Environ. Lett., vol. 11, no. 1, p. e70057, 2026.

[6] S. Shastri, S. Kumar, V. Mansotra, and R. Salgotra, “Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence,” Sci. Rep., vol. 15, no. 1, p. 25498, 2025.

[7] A. Bhola and P. Kumar, “Farm-level smart crop recommendation framework using machine learning,” Ann. Data Sci., vol. 12, no. 1, pp. 117–140, 2025.

[8] T. Dey, S. Bera, L. P. Latua, M. Parua, A. Mukherjee, and D. De, “ICrop: An intelligent crop recommendation system for agriculture 5.0,” IEEE Trans. AgriFood Electron., vol. 2, no. 2, pp. 587–595, 2024.

[9] A. Patel, S. Kharb, and R. Choudhari, “Management of crop recommendation system using predictive analytics,” in Emerging Trends in Industrial Engineering and Management, CRC Press, 2026, pp. 210–227.

[10] M. Y. Shams, S. A. Gamel, and F. M. Talaat, “Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making,” Neural Comput. Appl., vol. 36, no. 11, pp. 5695–5714, 2024.

[11] S. B. Kamatchi and R. Parvathi, “Improvement of crop production using recommender system by weather forecasts,” Procedia Comput. Sci., vol. 165, pp. 724–732, 2019.

[12] R. M. Pujahari, S. P. Yadav, and R. Khan, “Intelligent farming system through weather forecast support and crop production,” in Application of Machine Learning in Agriculture, Elsevier, 2022, pp. 113–130.

[13] J. Babber, P. Malik, V. Mittal, and K. C. Purohit, “Analyzing supervised learning algorithms for crop prediction and soil quality,” in 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), IEEE, 2022, pp. 969–973.

[14] M. Ishak, M. S. Rahaman, and T. Mahmud, "FarmEasy: An intelligent platform to empower crops prediction and crops marketing," in 2021 13th international conference on information & communication technology and system (ICTS), IEEE, 2021, pp. 224–229.

[15] P. A. de Oliveira Morais, D. M. de Souza, M. T. de Melo Carvalho, B. E. Madari, and A. E. de Oliveira, "Predicting soil texture using image analysis," Microchem. J., vol. 146, pp. 455–463, 2019.

[16] K. Abhang, S. Chaughule, P. Chavan, S. Ganjave, and E. VSB, "Soil analysis and crop fertility prediction," Int. Res. J. Eng. Technol., vol. 5, no. 3, pp. 3106–3108, 2018.

[17] "Pandith, V., Kour, H., Singh, S., Manhas, J., & Sharma, V. (2020). Performance evaluation of machine learning techniques for mustard crop yield prediction from soil analysis. Journal of scientific research, 64(2), 394-398".

[18] "Sharma, S., Jain, A., Sharma, S., & Whig, P. (2025). Enhancing crop yield prediction through machine learning regression analysis. International Journal of Sustainable Agricultural Management and Informatics, 11(1), 29-47".

[19] "Patil, Y., Ramachandran, H., Sundararajan, S., & Srideviponmalar, P. (2025). Comparative analysis of machine learning models for crop yield prediction across multiple crop types. SN Computer Science, 6(1), 64".

[20] "Muthulakshmi, S., Backiyavathy, M. R., Gopalakrishnan, M., Kalpana, R., Thangamani, C., & Pavithra, S. I. (2025). Insights of machine learning approach for soil fertility assessment and management strategy. Communications in Soil Science and Plant Analysis, 56(3), 436-457".

[21] "Chauhan, N., Sharma, D., & Dhaneja, D. (2025). A comprehensive crop prediction model leveraging optimization algorithms and machine learning classifiers. In Emerging Trends in Computer Science and Its Application (pp. 150-156). CRC Press".

[22] "Prasad, N. R., Patel, N. R., & Danodia, A. (2021). Crop yield prediction in cotton for regional level using random forest approach. spatial information research, 29(2), 195-206".

[23] "A. M. Lad, K. M. Bharathi, B. A. Saravanan, and R. Karthik, 'Factors affecting agriculture and estimation of crop yield using supervised learning algorithms,' Materials Today: Proceedings, vol. 62, pp. 4629–4634, 2022".

[24] "M. Kuradusenge et al., 'Crop yield prediction using machine learning models: Case of Irish potato and maize,' Agriculture, vol. 13, no. 1, p. 225, 2023".

[25] "P. Bandara, T. Weerasooriya, T. Ruchirawya, W. Nanayakkara, M. Dimantha, and M. Pabasara, 'Crop recommendation system,' International Journal of Computer Applications, vol. 975, p. 8887, 2020".

[26] "N. H. Kulkarni, G. N. Srinivasan, B. M. Sagar, and N. K. Cauvery, 'Improving crop productivity through a crop recommendation system using ensembling technique,' in 2018 3rd international conference on computational systems and information technology for sustainable solutions (CSITSS), IEEE, 2018, pp. 114–119".

[27] "Soil Types." Accessed: May 16, 2026. [Online]. Available: https://www.kaggle.com/datasets/prasanshasatpathy/soil-types

[28] "Sandy soil missing key nutrients but don't want invasive (soil forum at permies)." Accessed: Mar. 25, 2025. [Online]. Available: https://permies.com/t/273935/Sandy-soil-missing-key-nutrients

[29] P. Hargreaves, "Soil texture and pH effects on potassium and phosphorus availability," Potash Dev. Assoc. Lond. UK, 2015.

[30] M. V. Nathan, "Soils, plant nutrition and nutrient management," Univ Mo Ext Available Online Httpsextension2 Mo. Edumg4 Accessed 25 Sept. 2020, 2017.

[31] S.-Y. Lee et al., "Effect on chemical and physical properties of soil each peat moss, elemental sulfur, and sulfur-oxidizing bacteria," Plants, vol. 10, no. 9, p. 1901, 2021.

[32] S. P. Manalu, T. Sabrina, and Y. A. Sari, "Exploring the Triad: pH, Nitrogen, and Phosphorus Characteristics of Peat Soils in Humbang Hasundutan Regency," in E3S Web of Conferences, EDP Sciences, 2024, p. 03009.

[33] S. Sundari, "Comparison of dissolved organic carbon and nutrients content in Papua peatland," in IOP Conference Series: Earth and Environmental Science, IOP Publishing, 2022, p. 012016.

[34] G. S. Kumar, P. K. Saini, R. Deoliya, A. K. Mishra, and S. K. Negi, "Characterization of laterite soil and its use in construction applications: a review," Resour. Conserv. Recycl. Adv., vol. 16, p. 200120, 2022.

[35] D. Darmawan, D. Perdana, A. Ismardi, and I. W. Fathona, "Investigating the electrical properties of soil as an indicator of the content of the NPK element in the soil," Meas. Control, vol. 56, no. 1–2, pp. 351–357, 2023.

[36] A. R. Raut, M. R. Wahane, N. H. Khobragade, R. R. Banpatti, S. B. Thorat, and V. A. Rajemahadik, "Effect of NPK levels on the physico-chemical properties and available nutrient status under cluster bean in lateritic soils of Konkan (MS)," 2023.

[37] G. S. Vennela and A. Velumani, "Comparative analysis of characteristics of black soil using soil sensor against experimental analysis," in AIP Conference Proceedings, AIP Publishing, 2024.

[38] X. Dou, P. He, P. Zhu, and W. Zhou, "Soil organic carbon dynamics under long-term fertilization in a black soil of China: Evidence from stable C isotopes," Sci. Rep., vol. 6, no. 1, p. 21488, 2016.

[39] Y. He, F. Gu, C. Xu, and Y. Wang, "Assessing of the influence of organic and inorganic amendments on the physical-chemical properties of a red soil (Ultisol) quality," Catena, vol. 183, p. 104231, 2019.

[40] T. Hamazaki and E. P. PANINGBATAN Jr, "Properties and distribution patterns of Red-Yellow and related soils in the Philippines," JARQ, vol. 30, no. 2, pp. 69–78, 1996.

© 2026 by Sarika Agarwal, Himani Bansal, and Roop Singh. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).