IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 3, June 2026, pp. 2314~2324
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i3.pp2314-2324

A hybrid machine learning model for optimized mixed-crop recommendation

Ahmed Mohammed Gimba, Pradeep Kumar Mishra
Department of Computer Science and Applications, Sharda School of Computing Science and Engineering, Sharda University, Greater Noida, India

Article Info

Article history:

Received May 22, 2025
Revised Feb 20, 2026
Accepted Apr 20, 2026

Keywords:

Agriculture
Hybrid model
Machine learning
Mixed-crop
Nitrogen, phosphorus, potassium
Recommendation system

ABSTRACT

Farmers today encounter more challenges when selecting appropriate variety of crops depending on their farm soil nutrients and climate. This research will assist farmers in choosing suitable mixed-crops depending on the individual farms soil and climate conditions in Andhra Pradesh, India. Using the dataset sourced from Indian Institute of Soil Science (IISS), Bhopal with 2,552 entries. Previous studies focused on only single-crop recommendations. This work proposes a novel hybrid mixed-crop recommendation system (CRS) that incorporates several machine learning (ML) techniques comprise of random forest-ExtraTrees (RF-ExtraTrees), decision tree-C4.5 (DT-C4.5), extreme gradient boosting-gradient boosting (XGBoost-GBoost), quadratic discriminant analysis-linear discriminant analysis (QDA-LDA), and support vector machine-stochastic gradient descent (SVM-SGD) were utilized to recommend mixed-crops. To enhance the reliability of the training process, 20% of the dataset was held in reserve for validation to analyze model performance. The result of the proposed work shows that all the hybrid ML models applied were viable, and RF-ExtraTrees has achieved 95.91% best accuracy, 95.08% precision, and 95.91% recall, when contrasted to the other ML models.

Corresponding Author:

Pradeep Kumar Mishra
Department of Computer Science and Applications, Sharda School of Computing Science and Engineering
Sharda University
Greater Noida, India B-405, Rail Vihar Alpha 1, Greater Noida, Uttar Pradesh 210310, India
Email: pradeepkumar.mishra@sharda.ac.in

1. INTRODUCTION

Agriculture is one of the most imperative areas globally, and is a major contributor to the Indian economy, supporting livings and assuring food security by contributing 10% gross domestic product (GDP) worldwide and 17-20% GDP in India. With the rising challenges such as soil abasement, climate inconsistency, and growing demand for sustainable food production, intelligent decision-support systems have now become vital in present agriculture. In present-days machine learning (ML) techniques have stood commonly implemented for crop recommendation systems (CRS); however, most current systems are limited to recommend single-crop. Such approaches fail to replicate real-world farming practices, where mixed-cropping is regularly employed to enhance land-use efficacy, optimize resource utilization, decrease risks of production, and sustain long-term soil fertility. To address these hindrances, there is a demanding need for intelligent ML framework that can recommend suitable mixed-cropping using comprehensive soil and climatic data. Such a model can help farmers achieve increased productivity and resilience in changing and friendly, encourage sustainable farming methods, and increase precision agriculture decision-making [1]. An innovative CRS is proposed in this study to support farmers select the best crops for their individual farms. Shawon *et al.* [2] proposed soil nutrient analysis and a model for recommending crops using numerous types of ML techniques. This method addresses issues in traditional farming [3]. Precision agriculture is still in its primary periods and while it has resulted in more input that is efficient, yield, and better agricultural decision making, there are still many difficulties to address. Several models offer inputs for different agricultural fields of study, and crop recommendations is an important aspect of precision agriculture [4]. This model is tested and trained on datasets collected from various regions and assessed using common performance, comprising precision, accuracy, recall, F1-score, and confusion matrix [5]. This study employs a random forest (RF) ML model trained on a custom soil nutrient dataset containing diverse attributes, such as nitrogen, phosphorus, rainfall, temperature, sulphur, iron, pH, and zinc.

Despite technological advances, around 25% of worldwide crop yield is lost due to unproductive single-cropping arrangement practices (World Bank, 2023). A lot of studies from India expose those farmers using traditional recommendation systems face lower returns and greater pest vulnerability. These challenges emphasize the need for intelligent mixed-CRS that can adapt to local conditions and promote sustainable agriculture.

As a solution, we propose to develop a novel mixed-CRS using hybrid ML model trained on dataset acquired from Indian Institute of Soil Science (IISS) containing multiple variables such as nitrogen, phosphorus, potassium, temperature, humidity, pH value, and rainfall. The hybrid ML models will support farmers in choosing the best suitable variety of crops to be cultivated based on their farm soil nutrients and climate condition. To achieve this, we used hybrid ML algorithms. Major contribution to this work includes:

i) A novel hybrid mixed-CRS that will accurately suggest the best suited variety of crops based on the soil and climate condition of Andhra Pradesh District.
ii) Enhancing the number of crops and the dimension of the dataset.
iii) Accurate prediction of crops according the soil and climate using hybrid ML technique.

2. LITERATURE REVIEW

This part offers an in-depth review of existing research on crops recommendation systems. Several research articles provide significant contributions [6]. This article focuses on collecting soil-related information from farmers to provide accurate crop recommendations. It compares ML methods RF, extreme gradient boosting (XGBoost), and decision tree (DT) with RF achieving the highest accuracy among them [7]. This study collects soil data using various kinds of sensors, which is then sent to a mobile application that makes crop recommendations based on the parameters of the soil. Using a custom soil nutrients dataset, RF model was utilized [8]. This work presents a novel CRS that collects significant soil and environmental data from farmland using sensors. The model guarantees effective, real-time data transfer by utilizing the internet of things (IoT) and the message queuing telemetry transport (MQTT) protocol [9]. This work introduces a smart CRS that gathers and transmit data in real-time using sensors IoT and ML models provide precise and dynamic crop recommendations [10]. The authors created a neural network (NN) model that analyzes the data and provide agriculturalists with the accurate information to make informed crop selections [11]. This article demonstrates how to use IoT and AI technologies to enhance agricultural operations by recommending suitable crops based on real-time sensor data using image processing to predict possible diseases [12]. The researchers applied a regional dataset and ML techniques like support vector machine (SVM), RF, artificial neural network (ANN), and k-nearest neighbors (KNN) to predict appropriate crops [13]. In this research, the RF method was employed to develop a CRS. Model was able to train on a huge amount of data [14].

In this article a convolution neural network (CNN) model was created utilizing images to diagnose deficiencies in plant leaves. In order to process plant leaf images and precisely suggest the right nutrients to be given in order to overcome the deficiency, the study employs a ResNet-50 deep convolution neural network (DCC) [15]. The paper utilized decision-making process was based on ML techniques, and the full prototype was created with an STM32 ARM processor and Nucleo board to collect soil data. The result was shown on the LCD display and in the Blynk app [16]. The researchers developed ML that can recommend suitable crop and fertilizer utilizing diverse ML application algorithms [17]. The paper examines the changing landscape of crop suggestion systems and demonstrates the use of ML algorithms and data analytics to improve crop selection based on environmental and historical data [18]. The authors developed a multi-class classification model using IoT data. The data consists mostly of environmental variables such as nitrogen, phosphorus, potassium (NPK) levels, temperature, humidity. The proposed model, a hybrid of long short-term memory (LSTM) and time series-(transmission system operator security cooperation network services (TSCNET)), the model recommends crops based on location and time-space date [19]. The researchers implemented a system to help farmers combine their current farming practices with cutting-edge technologies such as deep learning (DL), ML, and remote sensing, using an open technology stack consisting of Angular for the frontend, Flask for the backend, and MySQL and Google Earth database [20]. Presents a crop yield recommendation system based on a precise comparison of various ML regression models, demonstrating an overall performance enhances with of 3.6% [21].

The authors used DL techniques and analyzed the agricultural metrics tracked by the IoT and fed them into the DL algorithm [22]. The article employed ML techniques like DT, RF, and light gradient boosting machine (LGBM) to choose crops, and LGBM outperformed the other two techniques [23]. In this work, a soil nutrient prediction system is built that predicts NPK levels and recommends which crop to be cultivate. The Boruta regression model has the accuracy of 91% [24]. The proposes fertilizer recommendation system uses ML with real-time soil data comprise of nitrogen, phosphorus, potassium, pH, temperature, humidity, moisture, and rainfall captured an image-based soil fertility mapping system. ML algorithms were applied, with a sequential CNN model achieving 91% accuracy. Similarly, in [25]. Compared DT, KNN, and RF algorithms in a crop recommendation model. Table 1 demonstrate the summary of recent studies on CRSs, highlighting datasets used, applied techniques, performance attained and their key limitations. The research gaps of this study are:
- Existing studies for crop recommendation have inadequate scope. These studies have repeatedly focused completely on recommending single crop.
- There is need to develop mixed-CRS.
- The existing works focus on recommending a few varieties of crops.
- There is need to develop a hybrid ML model for improved accuracy.

Table 1. Comparison among the current studies on CRS

Reference	Dataset	Techniques used	Limitations	Performance
[1]	Kaggle dataset	ML and DL	22 crops are considered	TCN achieved 93%
[26]	Indian agriculture dataset	ML and IoT	18 different variety of crops	Voting base ensemble with the accuracy of 93.91%
[2]	Kaggle dataset	ML	22 crops are considered	RF and DT attained the highest accuracy of 90.7%
[16]	Indian Chamber for Food and Agriculture	ML and NN	10 crops various crops are utilized	Adaptive neuro-fuzzy inference system (ANFIS) achieved 93.10% accuracy
[3]	Kaggle dataset	ML	22 crops are considered	RF outperformed with 92.09%
[6]	Dataset of Uttara Kannada	ML	20 different variety of crops	RF achieved highest accuracy of 90%
[7]	IoT sensor dataset	ML	12 unique crops have been utilized	Hybrid model RF+CNN achieved 89.23% accuracy
[9]	Dataset from Northern Bengal	ML and IoT	9 different variety crops	KNN achieved 91.97% accuracy
[15]	IoT sensor dataset	ML and IoT	6 crops are considered	DT outperformed with 87.98% accuracy
[19]	Kaggle dataset	DL and ML	22 crops are considered	DL model attains 90% accuracy and ML model achieved 93% accuracy

3. METHOD

The important aim of our research is to develop a novel mixed-CRS for Andhra Pradesh soil and climate conditions capable of producing extremely accurate results using hybrid ML models and carefully obtained dataset from IISS, Bhopal. The recommended approach uses hybrid ML algorithms for mixed-crop recommendation. These models incorporate robust features and implement a recommendation system employing cutting-edge ML techniques such as RF-ExtraTrees, DT-C4.5, XGBoost-gradient boosting (GBoost), quadratic discriminant analysis-linear discriminant analysis (QDA-LDA), and SVM-stochastic gradient descent (SGD) for suitable mixed-crop recommendation. Algorithm 1 presents mixed-CRS, a hybrid ensemble framework that processes regional soil datasets through preprocessing, trains five classifier combinations (RF+ExtraTrees, DT+C4.5, XGBoost+GBoost, QDA+LDA, SVM+SGD), evaluates performance metrics, and recommends optimal mixed-crops. Figure 1 show the complete structure of the proposed model framework, illustrating the sequential flow and interconnection of all the major modules of the study.

Algorithm 1. Mixed-CRS
Step 1: Collect the dataset
(Regional dataset)
Step 2: Perform data-preprocessing
(Replacing missing values, data cleaning, feature selection, and engineering)
Step 3: Model training and testing
(Divide the data into 80% for training and 20% for testing)

Step 4: Building hybrid models
(RF+ExtraTrees, DT+C4.5, XGBoost+GBoost, QDA+LDA, SVM+SGD)
Step 5: Model performance evaluation
(Accuracy, precision, recall, confusion matrix)
Step 6: Recommended mixed-crops

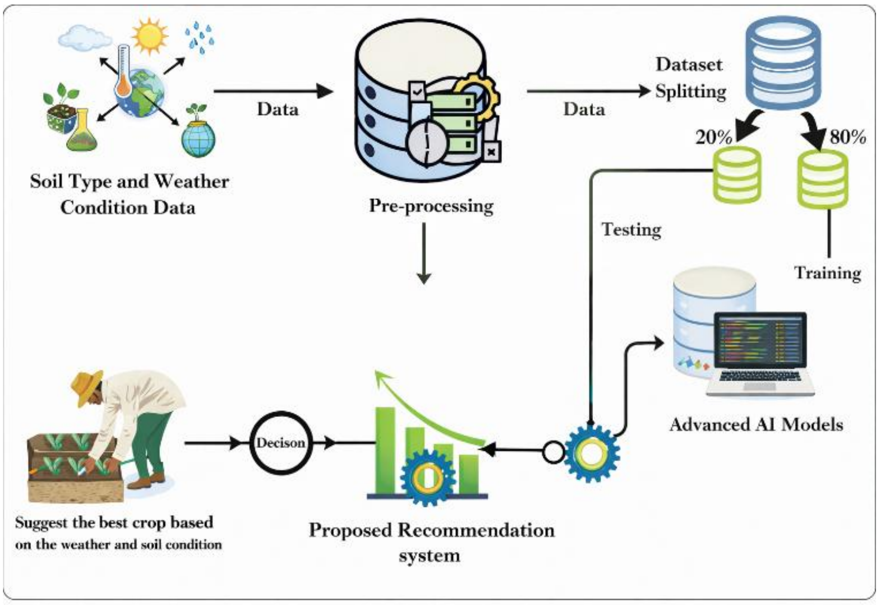

Figure 1. Proposed model architecture

3.1. Machine learning

It is a branch of computer science that leverages statistical methods and algorithms to allow computers to detect patterns, make informed decisions, or generate predictions from input data.

i) RF classifier: builds several DTs while the training phase integrates their outputs using the most vote for classification tasks or averaging for regression activities to make more accurate and robust predictions. Let $h1(x), h2(x), \dots, hT(x)$, be the predictions of T individual DTs in the forest for input x. Then the RF predictions as in (1).

$$\bar{y} = mode\{h1(x), h2(x), \dots, hT(x)\} \quad (1)$$

ii) ExtraTrees classifier: builds many unpruned DTs using the full training dataset. Unlike traditional DTs or RF, it introduces additional randomness by selecting split thresholds randomly instead of choosing the optimal split, enhancing diversity among trees and reducing overfitting. The mathematical equation is given in (2).

$$\bar{y}(x) = \frac{1}{M} \sum_{m=1}^{M} hm(x) \quad (2)$$

iii) DT classifier: it can be used for classification and regression problems. The predictions are made by traversing from the root node through decision rules to a final leaf node, which symbolizes decisions and their results in a classified, tree-structure. The mathematical equation is given in (3).

$$\begin{aligned} Entropy &= \textstyle\sum_{i=1}^{c} - pi * log2(pi) \\ Gini &= 1 - \textstyle\sum_{i=1}^{c} (pi)2 \end{aligned} \quad (3)$$

iv) C4.5 classifier: builds a DT by selecting features that yield the maximum normalized information obtained, which is expressed as the entropy variance at each decision point. It can handle both continuous and categorical variables, efficiently handles missing data, and once constructed, has a pruning mechanism to make the tree simpler and less prone to overfitting. The mathematical equation is given in (4).

$$Entropy(S) = -\sum_{i=1}^{c} pi \; log2(pi) \tag{4}$$

v) XGBoost classifier: is an ensemble leaning method using GBoost that repeatedly reduces the errors made by earlier models while combining the output of various weak learners, usually DTs to make a strong predictive model. The mathematical equation is given in (5).

$$\text{Obj} = \sum_{i=1}^{n} l(y_i, \bar{y}_i) + \sum_{k=1}^{k} \Omega(fk) \tag{5}$$

vi) GBoost classifier: is a method that repeatedly integrates several weak learners, typically decision, to create a robust predictive model. Gradient descent is used to lessen a particular loss in each iteration as a new model is trained to predict the residual errors of the previously model. The mathematical equation is given in (6).

$$\text{function.}\; Fm(x) = \sum_{m=1}^{M} Ym^{h}m(x) \tag{6}$$

vii) QDA: is a probabilistic classification method which generates a nonlinear, quadratic decision boundary by representing for each class feature distribution using a multivariate Gaussian distribution with a different covariance matrix. The mathematical equation is given in (7).

$$\delta k(x) = -\frac{1}{2}\log|\textstyle\sum k| - \frac{1}{2}\,(x - \mu k)T \textstyle\sum_k^{-1}(X - \mu k) + log\pi k \tag{7}$$

viii) LDA: is a probabilistic classification technique that makes an assumption that every class has a common covariance matrix and a multivariate Gaussian distribution. As a result of this, classes are successfully divided based on their mean variances using linear decision boundaries. The mathematical equation is given in (8).

$$\delta k(x) = x^{T} \textstyle\sum^{-1} \mu k - \frac{1}{2}\,\mu T/k \textstyle\sum^{-1} \mu k + log\pi k \tag{8}$$

ix) SVM: is an algorithm that classifies the optimal hyperplane to distinct data points of diverse classes by maximizing the margin between the closest points, known as support vectors, ensuring the most distinct boundary for classification. The mathematical equation is given in (9).

$$\text{Subject to: } min/(w, b) \quad 1/2 \;\; |(|W|)| \;⟦\text{^}2⟧ \tag{9}$$

$$yi(\mathrm{w}^{n}xi + b) \geq 1\, for\ all\ i$$

x) SGD: is an iterative optimization approach that updates model features using the gradient of the loss function calculated from one training example (or a mini-batch) at a time, allowing for faster and more scalable learning on large datasets. The mathematical equation is given in (10).

$$\theta := \theta - \eta\,.\,\nabla\theta L(\theta; xi, yi) \tag{10}$$

The datasets have been collected from IISS, Bhopal. The dataset includes soil and climate parameters consist of nitrogen, phosphorus, potassium, K, temperature, humidity, pH, and rainfall. Then, we consult and work together with soil scientist expert and analyze the soil and climate characteristics. Based on this analysis we perform feature engineering, and determine which crops can be cultivated together under the same farmland conditions. Figure 2 illustrate the spatial distribution of main towns in the Andhra Pradesh region from which soil and climatic samples were obtained and used in this work.

The dataset samples were composed from all the major towns in Andhra Pradesh, India with 36 diverse crops and 2,552 data entries. The study also considered the aggregation of all agricultural seasons. Figure 3 show the sample of dataset that were employed in this study for the development of models and analysis, indicating key features such as soil nutrient and environmental.

Figure 2. Study region from which soil and climate were obtained, i.e., Andhra Pradesh

```
df.tail()
```

	N	P	K	temperature	humidity	pH	rainfall	label
2548	140	14	350	22.107190	78.583201	6.364730	74.941366	cotton, soybean, pigeon pea
2549	160	18	400	23.038140	76.110215	6.913679	91.496975	cotton, soybean, pigeon pea
2550	180	22	450	24.547953	75.397527	7.766260	63.880799	cotton, soybean, pigeon pea
2551	200	26	500	23.738680	75.775038	7.556064	76.636692	cotton, soybean, pigeon pea
2552	220	30	550	22.318719	83.861300	7.288377	65.357470	cotton, soybean, pigeon pea

Figure 3. Sample of soil and climate dataset

3.2. Data pre-processing

This is one of the most significant phases in ML since the accuracy and performance of the ML model are dependent on the quality of the data that we keep [13]. In data pre-processing, we convert the collected data so that we may use it for analysis. Data pre-processing consists of selecting some important features, fixing missing data values, removing duplicates, overfitting and underfitting, dealing with outliers, and feature engineering. Feature selection the process of selecting or choosing the significant input parameters from the dataset that contribute the most to the prediction output of a ML model.

– Handling missing values: this stage deals with data entries where some information is missing, like empty cells in the dataset.
– Normalization: is a technique that used to transform the values of numeric features so they all have a similar range of distribution. This method helps ML models train faster and more accurate.
– Feature engineering: the process of producing, transforming, or selecting features to make a ML model more accurate and effective.

3.3. Model training and testing

The dataset is fragmented into training and testing batches, which enhances memory efficiency and accuracy. A validation set continuously analyzes the model's performance to shun overfitting and selects the optimal model based on the validation parameters. After training is completed, we tested on a different test set to see how well it generalizes to new data.

3.4. Model performance evaluation

The model assessment refers to the method in ML during which performance of a model is analytically assess and evaluated. The purpose is to find out how good the prediction quality of the model is and how well. Each mixed-crop recommendation model's accuracy, precision, and recall score were thoroughly assessed as define in (11) to (13).

– Accuracy: a simple technique to assess accuracy is to look at how frequently the classifier predicts accurately. The ratio of the total of precisely predicted outcomes to all of the model's predictions can be utilized to calculate accuracy.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{True\ Positives + True\ Negatives + False\ Positives + False\ Negatives} \quad (11)$$

– Precision: describes as the ratio of true positives to the sum of true positives and false positives, which shows the positive predictions accuracy.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (12)$$

– Recall: is the ratio of true positives to the sum of true positives and false negatives, which specifies how successfully the model classifies all appropriate events.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (13)$$

– Confusion matrix: is a performance assessment tool that illustrate the accuracy of a classification model. In order to assess model performance, recognize misclassifications, and improve overall prediction accuracy, it indicates the sum of true positives, true negatives, false positive, and false negatives. The mathematical equation is given in (5).

$$CM = \begin{bmatrix} True\ Positives & False\ Positives \\ False\ Negatives & True\ Negatives \end{bmatrix} \quad (14)$$

4. RESULTS AND DISCUSSION

The results of the proposed hybrid mixed-crop recommendations attained using RF-ExtraTrees, DT-C4.5, XGBoost-GBoost, QDA-LDA, and SVM-SGD models show that the proposed RF-ExtraTrees outperforms the others with 95.91% accuracy. The hybrid ML models were compared using several performance metrics. The proposed hybrid model offers tailored mixed-crop recommendations by analyzing the unique soil and climatic conditions of each farmer's farmland, enabling the cultivation of various and suitable mixed-crop. This has the possibility to significantly increase farm production and increase the economic returns for farmers. Table 2 demonstrate the performance comparison of the ML models using accuracy, precision, and recall. It can be experimentally shown that the hybrid RF-ExtraTrees model attain the peak accuracy, indicating its predictive capability and robustness for mixed-crop recommendation task. In the same sequence, Figure 4 demonstrates the RF-ExtraTrees model achieving the peak accuracy, precision, and recall among all ML models, demonstrating its superior performance for mixed-crop recommendation and Figure 5 shows the confusion matrix of the same RF-ExtraTrees hybrid model, proving the distribution of accurate and inaccurate classified mixed-crop instances.

Table 2. Hybrid mixed-crop recommendation performance metrics

Hybrid models	Accuracy (%)	Precision (%)	Recall (%)
RF-ExtraTrees	0.95.91	0.95.08	0.95.91
DT-C4.5	0.91.59	0.91.72	0.91.59
XGBoost-GBoost	0.94.13	0.94.25	0.94.13
QDA-LDA	0.83.17	0.83.73	0.83.17
SVM-SGD	0.88.45	0.88.97	0.88.45

Table 3 present the performance comparison of various existing models, highlighting their respective accuracy levels, is provided. Figure 6 shows the detailed assessment of the hybrid RF-ExtraTrees model, which outperformed all other relevant studies by achieving the highest performance. The proposed methods results are shown against several significant existing studies, as illustrated in Table 3 and Figure 6. This research work introduced an innovative hybrid ML technique, including RF-ExtraTrees, DT-C4.5, XGBoost-GBoost, QDA-LDA, and SVM-SGD, to the collected dataset. The model's performance has been evaluated using metrics consisting of accuracy, precision, and recall. Based on the results, the proposed hybrid approach performs more effectively than the existing models, by attaining an astounding accuracy of 95.91.%.

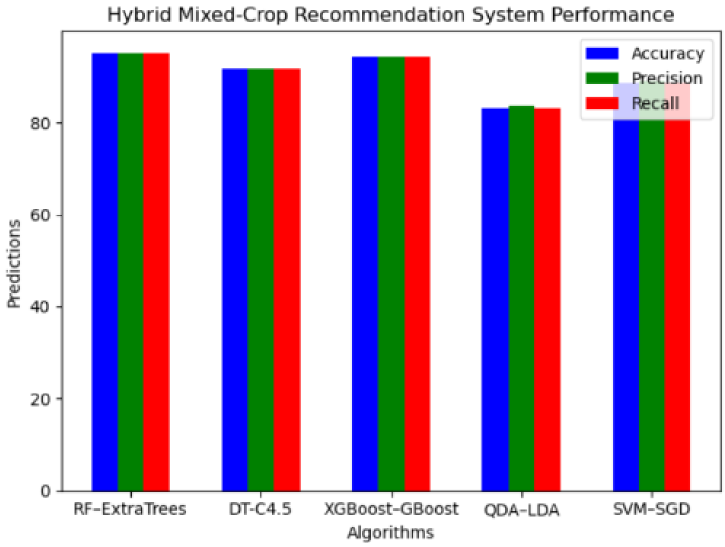

Figure 4. Hybrid mixed-crop recommendation bar chart

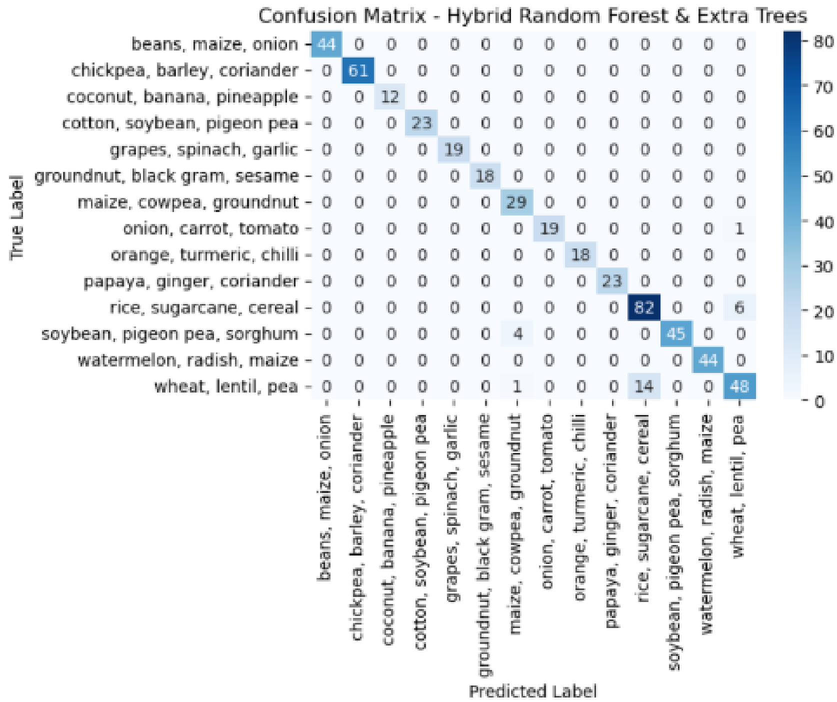

Figure 5. Hybrid mixed-crop recommendation confusion matrix of RF-ExtraTrees

Table 3. Accuracy for all present models and the proposed model

Models	Performance (%)
TCN	93
Voting base ensemble	93.91
RF+DT	90.70
ANFIS	93.10
RF	92.09
RF	90
RF+CNN	89.23
DT	87.98
Proposed RF+ ExtraTrees	95.91

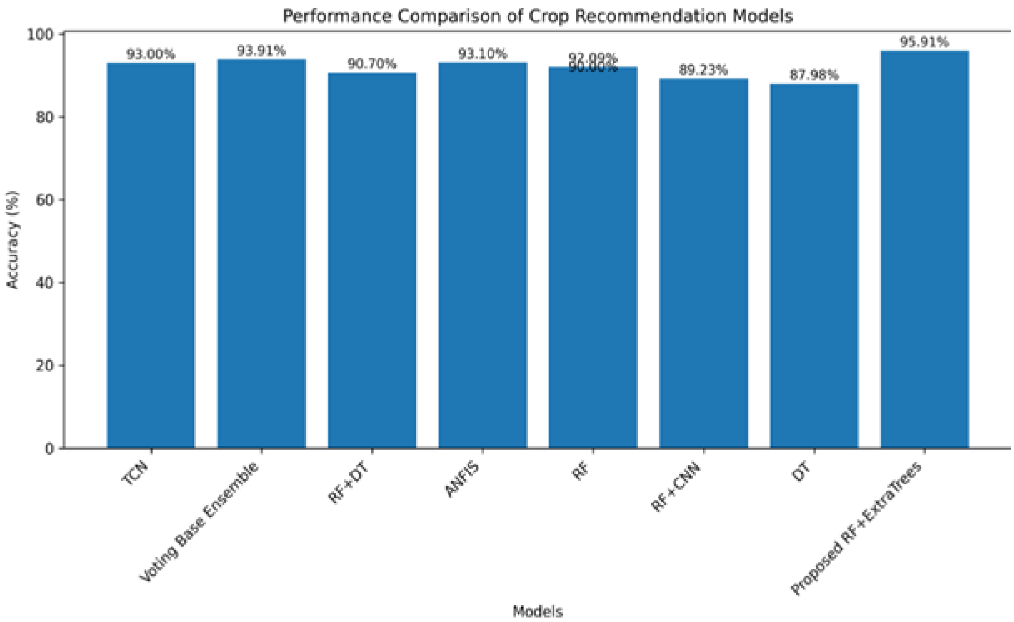

Figure 6. Performance comparison between the existing models and proposed model

Considering their demonstrated efficacy in agricultural datasets, RF and ExtraTrees and combined into hybrid model. Complex patterns in soil and climate features can be captured by RF because of its capability to handle high-dimensional data efficiently, provide feature importance, and be robust to noise. ExtraTrees, on the hand, works well and is ideal for fine-tuning decision boundaries in mixed-crop classification. Their combination ensures improved generalization and prediction accuracy, as supported by [13], [22], who demonstrated similar benefits in agri-intelligence applications.

This work discusses the importance of hybrid ML models for mixed-crop recommendation in modern agriculture, which is intended to enhance output in farming and minimize resource waste through intelligent agricultural decision-making. The time required to develop and evaluate each classification algorithm is also presented, along with findings from experiments illustrating the accuracy, precision, and recall of data analysis techniques. This experiment used various kinds of hybrid ML algorithms for evaluating mixed crops based on common features. The results were useful because they provided recommendations for situations where certain crop varieties are unknown or difficult to identify. The findings of our study proved that when assessing agricultural datasets, ML algorithms' accuracy is significantly improved by effective feature selection. RF-ExtraTrees generated the highest accuracy of 95.91% for mixed-crop recommendation using the soil and climate features dataset. This study provides a lot of insight on potential benefits of these technologies in modern agriculture, and further research and advancement in this field may improve mixed-crop recommendation, reduce resource waste, and improve global food security.

5. CONCLUSION

The proposed mixed-crop recommendation using hybrid ML algorithms has proved promising results in precision agriculture decision-making. By employing various soil and climate features evaluated. The application of hybrid ML algorithms, RF-ExtraTrees with 95.91% accuracy proved effective in recommending suitable mixed-crop. The future work can include various technique to it by incorporating IoT for taking real-time soil and climate data and employing explainable artificial intelligence (XAI) for farmers detail and proper explanation about the recommended crops.

ACKNOWLEDGMENTS

The authors express their propound appreciation to Tertiary Education Trust Fund (TetFund), Nigeria and Sharda University, Greater Noida, India.

FUNDING INFORMATION

This research was conducted without any financial support from any organizations or funding agencies.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Ahmed Mohammed Gimba	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Pradeep Kumar Mishra		✓		✓	✓	✓		✓	✓	✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

No any conflict of interest associated to this study.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [PKM], upon reasonable request.

REFERENCES

[1] A. Ghosh, S. K. Mohapatra, P. Pattanaik, P. K. Dash, and S. Chakravarty, "A comprehensive crop recommendation system integrating machine learning and deep learning models," in *2024 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU)*, 2024, doi: 10.1109/IC-CGU58078.2024.10530724.

[2] S. M. Shawon, N. I. Neha, A. N. Jui, N. Dey, and M. M. S. Raihan, "AgriAI: machine learning frameworks for tailored crop recommendations using soil nutrient parameters," in *2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems: Innovation for Sustainability (iCACCESS)*, 2024, doi: 10.1109/iCACCESS61735.2024.10499482.

[3] S. Gawade, G. Rout, P. Kochar, V. Ahire, and T. Namboodiri, "AGROFERDURE: intelligent crop recommendation system for agriculture crop productivity using machine learning algorithm," in *2023 International Conference on Computer, Electronics and Electrical Engineering and their Applications (IC2E3)*, 2023, doi: 10.1109/IC2E357697.2023.10262476.

[4] S. M. Badhusha, H. A. Athmashree, E. R. Bhoomika, Anusha, and M. Bhuvan, "An effective and comprehensive crop recommendation system using machine learning," in *International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS)*, 2024, doi: 10.1109/ICETCS61022.2024.10543400.

[5] M. Mehta, A. Saha, S. Fulia, T. Jain, and N. Shekokar, "An innovative model for soil testing with intelligent crop recommendation system," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2023, doi: 10.1109/ICCCNT56998.2023.10307726.

[6] B. Mahalakshmi, V. Sakthivel, B. S. Devi, and S. Swetha, "Agricultural crop and fertilizer recommendations based on various parameters," in *International Conference on Sustainable Computing and Smart Systems, ICSCSS* 2023, pp. 735–739, doi: 10.1109/ICSCSS57650.2023.10169320.

[7] P. Singh, M. Kaur, and R. Bajaj, "An IoT-enabled crop recommendation system utilizing MQTT for efficient data transmission to AI/ML model," in *2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT)*, 2024, pp. 315–320, doi: 10.1109/ICCPCT61902.2024.10672704.

[8] A. Kayum and M. O. Rahman, "Appropriate crop recommended system for cultivation using IoT and ML," in *2023 5th International Conference on Sustainable Technologies for Industry 5.0 (STI)*, 2023, doi: 10.1109/STI59863.2023.10464865.

[9] S. N. Pasha, D. Ramesh, S. Mohmmad, and G. Deepak, "Automatically suggesting suitable crop from soil samples using optimized neural networks," in *2024 3rd International Conference for Innovation in Technology (INOCON)*, 2024, doi: 10.1109/INOCON60754.2024.10511669.

[10] K. S. Priya, J. A. Jenifer, S. P. Janani, M. S. Aarthi, and T. Kavitha, "Crop recommendation and disease prediction using IOT and AI," in *2024 10th International Conference on Communication and Signal Processing (ICCSP)*, 2024, pp. 807–812, doi: 10.1109/ICCSP60870.2024.10543366.

[11] B. N. Mohapatra and V. Kale, "Crop recommendation system using machine learning," *ITEGAM-Journal of Engineering and Technology for Industrial Applications*, vol. 10, no. 48, pp. 1728–1733, 2024, doi: 10.5935/jetia.v10i48.1186.

[12] S. R. Sani, S. V. S. Ummadi, S. Thota, N. Muthineni, V. S. S. Swargam, and T. S. Ravella, "Crop recommendation system using random forest algorithm in machine learning," in *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, May 2023, pp. 501–505, doi: 10.1109/ICAAIC56838.2023.10141384.

[13] K. Prathik, J. Madhuri, and J. K. Bhanushree, "Deep learning-based fertilizer recommendation system using nutrient deficiency prediction," in *2023 7th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*, Nov. 2023, pp. 1–6, doi: 10.1109/CSITSS60515.2023.10334122.

[14] J. S. Jeyanathan, B. Veerasamy, B. Medha, G. T. V. Sai, R. B. Kumar, and V. Sahu, “Design of crop recommender system using machine learning and IoT,” in *7th International Conference on Trends in Electronics and Informatics*, 2023, pp. 1127–1132, doi: 10.1109/ICOEI56765.2023.10125963.

[15] K. D. Priya, A. S. Samyogitha, A. V. K. Reddy, and B. D. Sri, “Ensembled cropify-crop and fertilizer recommender system with leaf disease prediction,” in *International Conference on Innovative Data Communication Technologies and Application*, 2023, pp. 600–604, doi: 10.1109/ICIDCA56705.2023.10100117.

[16] C. Shewale, A. Sardeshmukh, P. Shinde, O. Sapkal, and S. Shinde, “Intelligent system for crop recommendation and disease identification,” in *2024 IEEE 9th International Conference for Convergence in Technology, I2CT 2024*, 2024, doi: 10.1109/I2CT61223.2024.10544110.

[17] K. S. Sandeep, K. L. Sahithi, S. S. Azmi, J. R. Devi, and B. N. Jyothi, “IoT-driven precision crop recommendation system using LSTM & TSC-NET for sustainable agriculture,” in *2024 1st International Conference on Innovative Sustainable Technologies for Energy, Mechatronics and Smart Systems (ISTEMS)*, 2024, doi: 10.1109/ISTEMS60181.2024.10560216.

[18] S. Mahalakshmi, A. J. Anand, and T. Sampath, “Recommendation of crops and fertilizer, detection of crop weed, pest and diseases using machine learning,” in *2023 International Conference on Data Science, Agents and Artificial Intelligence (ICDSAAI)*, 2023, doi: 10.1109/ICDSAAI59313.2023.10452616.

[19] 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*, vol. 12, no. 1, pp. 1–20, 2021, doi: 10.4018/IJAEIS.20210101.oa1.

[20] M. Altalak, M. A. Uddin, A. Alajmi, and A. Rizg, “Smart agriculture applications using deep learning technologies: a survey,” *Applied Sciences*, vol. 12, no. 12, 2022, doi: 10.3390/app12125919.

[21] A. Sunilkumar and A. A. Jovith, “Smart irrigation system by recommending crop and fertilizers,” in *2nd IEEE International Conference on Networking and Communications 2024 (ICNWC)*, 2024, doi: 10.1109/ICNWC60771.2024.10537564.

[22] R. Madhumathi, K. A. Kumar, N. C. Kumar, A. Dhyanesh, and T. Arumuganathan, “Soil nutrient prediction and crop recommendation system,” in *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)*, 2023, pp. 976–981, doi: 10.1109/ICCPCT58313.2023.10245685.

[23] A. Shahzad, M. Ibrahim, R. M. Saleem, S. Zia, S. Habib, and M. Mahmood, “Unveiling soil fertility patterns via image analysis and machine learning for accurate crop recommendations,” *Journal of Computing and Biomedical Informatics*, 2024, pp. 1-15.

[24] U. Ahmad and L. Sharma, “A review of best management practices for potato crop using precision agricultural technologies,” *Smart Agricultural Technology*, vol. 4, 2023, doi: 10.1016/j.atech.2023.100220.

[25] J. Gao *et al.*, “A fertilization decision model for maize, rice, and soybean based on machine learning and swarm intelligent search algorithms,” *Agronomy*, vol. 13, no. 5, 2023, doi: 10.3390/agronomy13051400.

[26] R. Ed-daoudi, A. Alaoui, B. Ettaki, and J. Zerouaoui, “A predictive approach to improving agricultural productivity in Morocco through crop recommendations,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 3, pp. 199–205, 2023, doi: 10.14569/IJACSA.2023.0140322.

BIOGRAPHIES OF AUTHORS

Ahmed Mohammed Gimba holds a Bachelor of Science in Computer Science from Islamic University in Uganda in 2015. He also received Masters of Science in Information Technology in Stafford University, Uganda in 2017. He is also a Ph.D. scholar in the Department of Computer Science and Applications in Sharda University, India. He is presently lecturing with the Department of Information Technology at Nigerian Army University, Biu, Borno State, Nigeria. He is a member of IEEE. His research areas include ML, neural networks, and cyber security. He can be contacted at email: ahmedmohammedgimba@gmail.com.

Pradeep Kumar Mishra working as an assistant professor in the Department of Computer Science and Application at Sharda University, Greater Noida. He has more than 22 years of broad knowledge in academics. His research area is ML and artificial intelligence. He is guiding 1 Ph.D. research scholar, published 30+ in SCI, Scopus indexed journal. Guided M.Tech. students, 100+ B. Tech students, 30 MCA and M.Sc. students, and one Ph.D. student. He can be contacted at email: pkmishra28@gmail.com.