Journal of IoT in Social, Mobile, Analytics, and Cloud (ISSN: 2582-1369)
www.irojournals.com/iroismac/

AI-Based Smart Agriculture System

Suhas K P.[1], Shivashanker[2], Karunakar Reddy[3], Channabasava Gouda H.[4]

Department of Computer Science Engineering in Data Science, Dayananda Sagar Academy of Technology & Management, Bangalore, India

Email:[1]1dt23cd022@dsatm.edu.in, [2]1dt23cd047@dsatm.edu.in, [3]1dt23cd056@dsatm.edu.in, [4]1dt24cd400@dsatm.edu.in

Abstract

Crops selection process is one of the crucial factors involved in increasing agricultural production while reducing wastage. Conventional techniques that are based on farmer's experiences tend to produce inconsistent results since the environment varies from place to place. In this research, an ensemble approach based on the stacking technique is suggested. The hybrid approach incorporates Adaptive Boosting (AdaBoost) and Light Gradient Boosting Machine (LightGBM) as base models and Logistic Regression as the meta-model. In particular, the hybrid approach makes use of seven input variables which include nutrient contents in soils (nitrogen, phosphorus, and potassium), pH level, temperature, humidity, and rainfall. Various evaluation metrics such as accuracy, precision, recall, and F1-scores are used to evaluate the suggested model. Results indicate that the suggested hybrid model produces an accuracy rate of 99.12%, which beats individual model's performance. The superior performance of the model is due to adaptive boosting and gradient boosting algorithms. Moreover, the model is implemented using a graphical user interface, which is used for predicting crops in real-time

Keywords: Smart Agriculture, Crop Recommendation, AdaBoost, LightGBM, Machine Learning, Ensemble Learning, Precision Farming.

1. Introduction

Selecting right crop for a piece of land is one big challenge in agriculture, especially when population is growing fast and climate keep changing. Growing proper crops in each season and region by looking at soil condition and weather is very important to increase yield and save resources [1], [6]. In precision farming, data-driven systems can help farmers to choose better crops and reduce the risk of crop failure. One research also said that smart crop selection can increase yield and reduce wastage of resources [3]. Usually, farmers depend on

DOI: https://doi.org/10.36548/jismac.2026.2.003
Received:23.02.2026, received in revised form:25.03.2026, accepted:10.04.2026, published: 21.04.2026

their local knowledge for crop planning, but sometimes it gives wrong results. AI-based systems give another option which can help farmers to make more accurate and faster decisions. learning from historical crop and environmental data to suggest optimal crops and management strategies [7]. In the past few years, artificial intelligence (AI) and machine learning (ML) are being used a lot in modern farming. These ML models are now helping in many farming tasks such as weather prediction, disease finding, irrigation control, and crop yield prediction, and also for suggesting which crop to grow [1].

Modern crop recommendation systems inputs which include soil nutrients, pH Value, moisture, temperature, and rainfall will help to find which crop will grow best. They use many types of data such as sensors, satellite images, and old yield records, and apply learning algorithms to rank crops based on how suitable they are [8]. Models using sensors looked at soil and weather data and helped farmers to choose crops that give better profit and support sustainability [3]. These kinds of Methods utilize the power of ML in identifying unseen patterns between weather and crop growth, helping farmers in planning crops more wisely than old traditional methods. Among different types of ML models used for crop recommendation, Decision Trees (DTs) are the most commonly used models. A decision tree is like a small flow chart in which data is separated according to certain criteria to formulate "if-then" type rules. It can take both number and category type data without doing too much extra work, and it can also handle missing values with ease.

Due to these reasons, DTs fit well for agriculture data which is mostly mixed and incomplete. But one main issue is that a single decision tree can easily overfit the data and become unstable. Even a Small change of dataset can make totally different tree. When the tree is too deep, it starts learning random noise instead of an actual pattern, then it gives poor results. DT finds its application along with many crop recommendations works along with other models such as KNN, SVM, and AdaBoost [3]. To AdaBoost tries mainly to fix the problem of single trees used Ensemble learning is a method that incorporates many small weak models, mostly small trees, one by one in order to make a strong and more accurate model [4]. In AdaBoost, each tree is trained to correct the errors made by the one before it. It will achieve this by giving more weight to the data points that were predicted wrong so the next tree can focus more on them [4]. By doing this, many times, AdaBoost combines many weak models and makes one strong and accurate model [4]. In crop selection, AdaBoost helps in improving the accuracy by using multiple trees together rather than just one. Another helpful algorithm is

Light Gradient Boosting Machine or, in short, LightGBM. This is a newly invented, and also an emerging boosting approach, especially useful for processing large amounts of data. LightGBM utilizes a tree growth algorithm that grows trees in a leaf-wise manner instead of growing levels at once, resulting in faster learning with lower memory consumption [5]. The algorithm has many benefits that include increasing speed, using minimal memory resources, and providing accurate results when working with large data sets [5]. All this is extremely important in agriculture due to data consisting of n number of records.

2. Related Works

Modern advances in precision agriculture have shown that it is necessary to leverage ML and IoT techniques in crop recommendation and yield prediction. With smart sensing and data-based approach, it is possible to achieve more accurate analysis of soil conditions, which helps increase the efficiency of decision making in agriculture [1], [2], [8], [16]. In addition, the use of real-time and previous data allows increasing prediction precision. Various researchers have explored the idea of utilizing ML algorithms for crop recommendation purposes. Thus, Senapaty et al. [3] offered a decision support model based on the application of classification technique, whereas Mahesh and Soundrapandiyan [5] used gradient-based methods for predicting the yield. Graph and deep learning models were also used for crop yield predictions [12], [16]. Nevertheless, the disadvantages of a single algorithm model include such problems as overfitting, poor generalization, and vulnerability to changes in data.

In light of these shortcomings, several techniques involving ensemble learning have been extensively utilized. Gunasekaran et al. [4] introduced an enhanced ensemble system that is capable of enhancing the precision of predictions using different algorithms. Hasan et al. [9] illustrated the superiority of ensemble-based crop suggestion systems, which can account for intricate interrelationships among agricultural parameters. Likewise, Bakr et al. [10] proved that ensemble models lead to improved stability and reliability when used in smart agriculture. Another set of references [14], [15] further confirms that multiple learners can enhance the results of predictions. In recent literature, the use of explainable AI along with advanced machine learning frameworks has also gained popularity. Shastri et al. [7] explained the significance of adopting explainable AI in ML systems used in agriculture. On the other hand, Mancer et al. [13] discussed interpretable crop selection processes. Furthermore, Kumar and Joshi [14] suggested some innovative ML techniques for suggesting crops to farmers.

However, some challenges still exist within these advances. One of the common issues in the current literature is that most studies used small data sets and a few algorithms, limiting the ability to generalize their findings [4], [11]. Furthermore, some models have not been able to combine the factors of soil properties and environmental characteristics, resulting in less effective recommendations [1], [3]. To address these challenges, the present study proposes a hybrid stacking ensemble model that integrates AdaBoost, LightGBM, and Logistic Regression to achieve improved accuracy, stability, and scalability in crop recommendation systems.

3. Proposed Methodology

The proposed system employs a stacking ensemble framework that integrates AdaBoost and LightGBM as base learners with Logistic Regression as the meta-learner. In this approach, multiple models are trained in parallel, and a higher-level model combines their predictions to improve overall accuracy [4], [9]. Soil and environmental inputs (nitrogen, phosphorus, potassium, pH, temperature, humidity, rainfall) are preprocessed and provided to the base models, which generate predictions independently. These outputs are then used as inputs to the Logistic Regression model, which learns optimal weights to combine them. This two-layer (base–meta) design improves predictive performance, and recent studies confirm that stacked ensembles achieve higher accuracy and R^2 compared to individual models [4], [7], [9], [10].

A. AdaBoost

AdaBoost (Adaptive Boosting) is an ensemble learning technique that constructs a strong classifier by combining multiple weak learners, typically decision stumps or shallow decision trees. Initially, all training samples are assigned equal weights, ensuring uniform importance across the dataset. Above Figure 1 explaining how the weaker models are gradually combined to form the final stronger model.

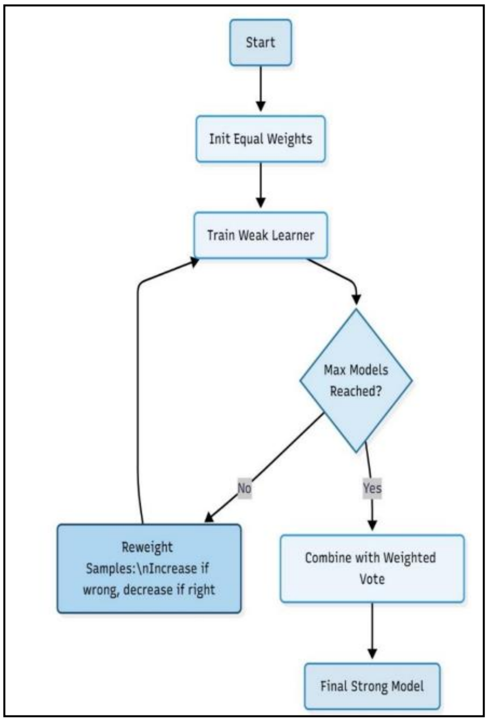

Figure 1. The Procedure of AdaBoost

At each iteration t, a weak classifier $h_t(x)$is trained to minimize the weighted classification error, denoted as ϵ_t, which reflects the proportion of misclassified samples considering their associated weights. Based on this error, the contribution of the weak learner is determined by its corresponding weight, calculated as:

$$\alpha_t = \frac{1}{2}\ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$$

where $\alpha_t$represents the weight of the t-th weak classifier, and $\epsilon_t$denotes the weighted classification error at iteration t. A lower value of $\epsilon_t$results in a higher α_t, thereby increasing the influence of more accurate classifiers in the ensemble. Subsequently, the weights of the training samples are updated such that misclassified instances receive higher weights, while correctly classified instances are assigned lower weights. This adaptive reweighting mechanism enables subsequent weak learners to focus more on difficult samples, thereby improving overall model performance.

After completing T iterations, the final strong classifier is obtained by aggregating the weighted predictions of all weak learners. The ensemble decision function is expressed as:

$$F(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t \ h_t(x)\right)$$

where $F(x)$denotes the final predicted class label, Trepresents the total number of weak learners, $\alpha_t$is the weight assigned to the t-th classifier, and $h_t(x)$is the prediction of the t-th weak classifier for input x. The function $\text{sign}(\cdot)$determines the final class label based on the weighted majority vote of all classifiers.

In this study, AdaBoost was implemented using 100 estimators, which provides a balance between model complexity and generalization capability while reducing the risk of overfitting and ensuring efficient learning.

B. LightGBM

LightGBM is a gradient boosting framework built specifically for efficient processing of large amounts of data. It builds decision trees by minimizing a smooth loss function via gradient descent. The key difference of LightGBM from standard boosting algorithms is the algorithm used for decision trees formation, which operates in the manner of expansion of the leaf with the largest loss decrease. As a result, better accuracy and faster convergence rates are achieved. At the same time, LightGBM retains excellent ability to generalize despite its power to detect nonlinear correlations between features and targets, as demonstrated in Figure 2 below. The objective function of LightGBM is defined as:

$$L = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

where Lrepresents the overall objective function, $l(y_i, \hat{y}_i)$denotes the loss function measuring the difference between the true value $y_i$and the predicted value $\hat{y}_i$, $n$is the total number of training samples, $K$is the number of trees, and $\Omega(f_k)$is the regularization term applied to the k-th tree to control model complexity. The regularization term is given by:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \parallel w \parallel^2$$

where γ is a parameter that penalizes the number of leaves T in the tree, thereby reducing overfitting, λ is the L2 regularization coefficient, and $\parallel w \parallel^2$ represents the squared magnitude of the leaf weights vector w. This regularization term ensures a balance between model complexity and generalization performance.

In this study, the implementation of LightGBM was done using 100 estimators and a learning rate of 0.1. The two parameters were chosen after thorough consideration in order to optimize both the accuracy and speed of computation. Because of its fast training speed and capability to deal with very large feature data sets, LightGBM has become popularly used in various areas, including forecasting yield of crops and smart agriculture [5], [7], [8].

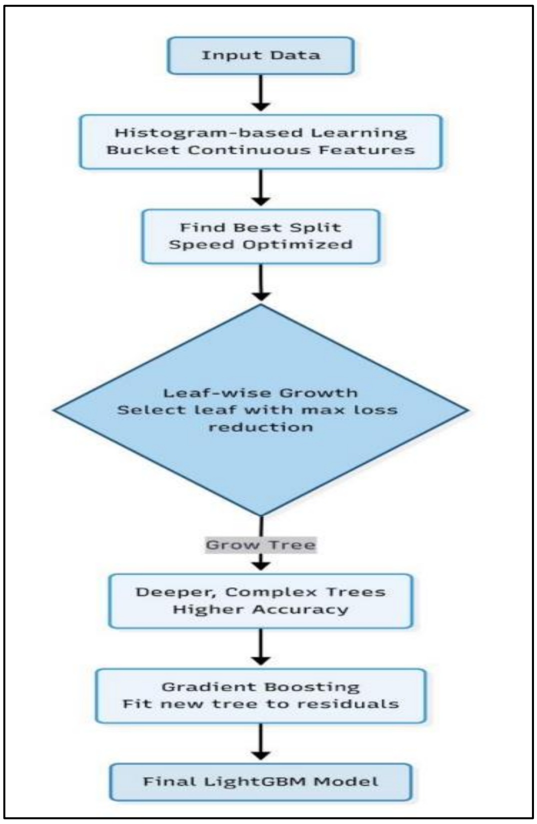

Figure 2. Workflow of the LightGBM Model

C. Meta-Learner (Logistic Regression)

The meta-classifier used for the stacking method is the Logistic Regression model, where the predictions made by individual models such as AdaBoost and LightGBM are combined to generate the final class probability. The Logistic Regression model makes use of the sigmoid function that maps the linear combination of the input data to values lying between 0 and 1, as shown below:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

where σ(z) is the sigmoid activation function and z is the linear combination of the input variables.

The linear combination of base learner outputs is expressed as:

$$z = w_0 + w_1 h_1(x) + w_2 h_2(x) + \cdots + w_n h_n(x)$$

Where z represents the total input to the Sigmoid function, w_0 represents the bias value, w_1, w_2, … wn represent the weights assigned to each base learner, and $h_1(x)$, $h_2(x)$, … hn(x) represent the output predictions from the respective base models. These values are calculated during training to appropriately weigh the predictions from the base learners.

Logistic Regression is chosen because of its ease of use, interpretability, and excellent capability in solving probabilistic classification problems. It helps to appropriately weigh the outputs from several base learners and reduce overfitting, as shown in Figure 3. Additionally, it gives accurate probability outputs, which help in ranking crop suitability. Its ability to give accurate results while using simple computation makes it extremely applicable in stacking ensembles. Previous research has shown that logistic regression improves the accuracy of ensembles significantly in predicting the best crops to recommend [4], [9].

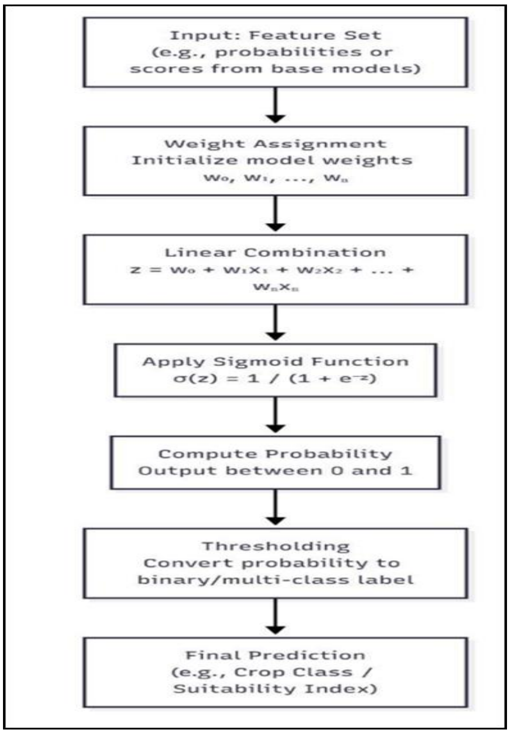

Figure 3. Working of the Meta Learner

D. Ensemble Modeling

Ensemble learning is one of the core concepts in machine learning, where an ensemble of models or classifiers known as base learners works together to form a better performing and more accurate prediction model. The main problem faced by many individual models can be high variance or bias or sensitivity to noise in the data. The combination of multiple base learners addresses these concerns and leads to improved accuracy and stability [4], [6]. The ensemble prediction can be mathematically represented as:

$$H_{ens}(x) = F(H_1(x), H_2(x), \dots, H_n(x))$$

where $H_{ens}(x)$ is the final prediction from the ensemble on the input x, while $H_1(x)$, $H_2(x)$,.,$H_n(x)$ are the individual predictions from the base learner models, and F(.) denotes the aggregation method used for aggregating these predictions. Different functions can be used as F(.), depending on the kind of problem, e.g., in classification tasks, F(.) could be the simple majority rule, while in regression problems, F(.) could be simply an average of the predictions. The application of such techniques in ensemble learning allows the learner to produce more accurate results by minimizing prediction errors.

E. Boosting

AdaBoost and LightGBM are some of the boosting algorithms that build their models in a sequential manner, wherein each successive model learns from the mistakes made by the previous models. For instance, the training method employed in the AdaBoost algorithm assigns weights to each data sample. The weights are then adjusted repeatedly depending on the accuracy of classification. Let h_t(x)be the weak hypothesis trained at the t-th iteration of the algorithm. Its significance is computed as follows:

$$\alpha_t = \frac{1}{2}\ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$$

where α_t stands for the weight allocated to the t-th weak hypothesis, while ϵ_t refers to the weighted error at the t-th iteration. As the weighted error decreases, the value of α_t increases.

The final strong classifier in AdaBoost is defined as:

ISSN: 2582-1369

$$F(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t \ h_t(x)\right)$$

where F(x) represents the predicted class label, T represents the total number of weak classifiers, α_t represents the weight of the t-th classifier, and h_t (x) represents the result produced by the t-th weak classifier. In such a way, the formula uses a method for minimizing an exponential loss while assigning more weight to incorrectly classified samples.

The second model called LightGBM utilizes gradient boosting but optimizes the following objective function:

$$L = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

where L represents the objective function, l$(y_i, \hat{y}_i)$ represents the loss function calculated between the real output y_i and the expected one $\hat{y}_i$, n represents the sample size, K represents the total number of trees, and Ω(f_k) represents the regularized part of the k-th tree.

The leaf-wise growing mechanism of trees used in LightGBM ensures the choice of split that minimizes the loss function, resulting in quicker training and more accurate performance. Another important feature of LightGBM is the presence of a regularization term Ω(f_k), which enables the control of complexity and thus prevents overfitting [5], [7].

F. Stacking

Stacking, or stacked generalization, involves combining the results from several base learners using a meta-model, which is sometimes called a meta-learner. In this case, all of the base models make predictions $h_i(x)$, which become input features for the meta-learner. It allows the meta-learner to effectively combine the results of different algorithms.The combined input to the meta-learner is expressed as:

$$z = w_0 + w_1 h_1(x) + w_2 h_2(x) + \cdots + w_n h_n(x)$$

where z represents the aggregated input to the meta-learner, w_0 is the bias term, $w_1, w_2, \ldots, w_n$ are the weights assigned to each base learner, and $h_1(x), h_2(x), \ldots, h_n(x)$ denote the predictions of the corresponding base models. These weights are learned during training to optimally combine the outputs.

The final prediction is obtained using the sigmoid activation function:

$$\hat{y} = \sigma(z) = \frac{1}{1 + e^{-z}}$$

where yˆ is the predicted probability of the target class and σ(z) is a logistic function mapping the aggregate input z into a range of 0 and 1. This function is widely used in Logistic Regression, which can be chosen as the meta-learner in classification problems because of its simplicity, easy interpretation, and good calibration property [4], [9].

G. Hybrid Model

Ensemble learning approaches have greatly enhanced the precision and reliability of crop prediction, yield estimation, and soil analysis in contemporary agriculture practices. The combination of algorithms like AdaBoost, LightGBM, and Logistic Regression has proven efficient and achieved remarkable accuracy greater than 99.12% [3], [4], [9]

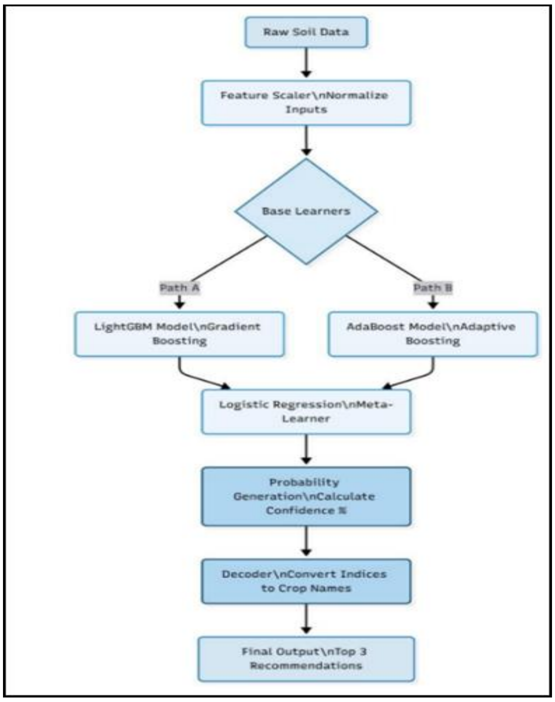

Figure 4. Work Flow of the Hybrid Model

The algorithms make use of various data features like soil nutrients, pH, temperature, humidity, and rainfall to predict the intricate relationship between the environment and appropriate crops [2], [8], [10]. With a decrease in bias and error in prediction, the ensemble learning approach has facilitated the smart farming system to provide reliable crop predictions, as depicted in Figure 4 [1], [6].

H. Dataset

The dataset employed in the current research is the Crop Recommendation Dataset available in Kaggle [17]. This is an elaborate and structured dataset, intended for classification purposes in precision agriculture. The dataset comprises 2,200 observations, with seven attributes as input and one class as the output variable. These input variables include N (Nitrogen), P (Phosphorus), K (Potassium), temperature, humidity, pH value, and rainfall the necessary parameters that affect crop growth.

There are 22 categories of crops within the dataset. They include but not limited to rice, maize, chickpea, kidney beans, cotton, coconut, and coffee. Each observation in the dataset represents different soil and environmental conditions, along with the recommended crop that fits them. The preprocessing process includes normalization of the values and division into the training and test datasets.

4. Results and Discussion

The proposed hybrid crop recommendation model has been implemented using the Streamlit framework. The platform ensures a user-friendly and interactive interface to predict crops in real time [2], [7]. As illustrated in Figure 5, the system enables users to enter seven critical agronomic features: Nitrogen (N), Phosphorus (P), Potassium (K), pH level, temperature, humidity, and rainfall. Such parameters constitute the vital soil and climatic factors that significantly impact the productivity of crops. Using the entered parameters, the model processes the data using the trained ensemble system and predicts appropriate crop recommendations.

The proposed hybrid crop recommendation model has been implemented using the Streamlit framework. The platform ensures a user-friendly and interactive interface to predict crops in real time [2], [7]. As illustrated in Figure 5, the system enables users to enter seven critical agronomic features: Nitrogen (N), Phosphorus (P), Potassium (K), pH level, temperature, humidity, and rainfall. Such parameters constitute the vital soil and climatic factors that significantly impact the productivity of crops. Using the entered parameters, the model processes the data using the trained ensemble system and predicts appropriate crop recommendations.

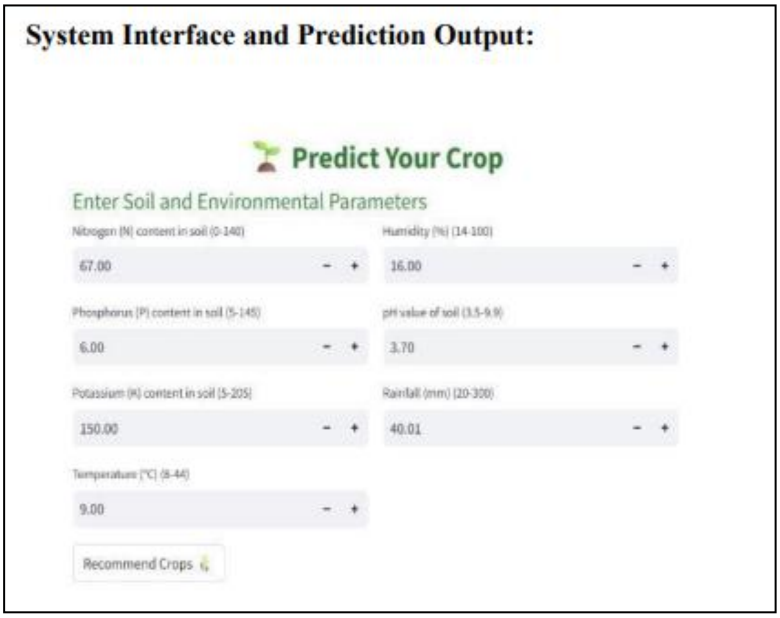

Figure 5. Crop Recommendation Interface

Once the "Recommend Crops" button is clicked, the application goes ahead and processes the input data using the stacked ensemble machine learning algorithm, providing the best three crop recommendations according to their suitability. Figure 6 below displays an example of the output provided by the model, where the crop recommendations and confidence levels are given.

Figure 6. Output Provided by the Model

From the sample provided, we have the following results: Chickpea with 88.79% suitability level, Rice with 1.29% suitability level, and Jute with 0.68% suitability level. Such type of output makes it easier for the users to analyse different options, as well as gain an understanding of the model's confidence in its recommendations [2], [4], [9].

Model Performance Evaluation

Performance of different machine learning algorithms is presented in Figure 7 and Table 1 below. Performance of the hybrid ensemble algorithm, which is the combination of AdaBoost and LightGBM, proves to be much better when compared to performance of individual algorithms used for prediction. It can be noted that the hybrid model produces more accurate results in terms of different measures, such as accuracy, precision, recall, and F1 score (Table.1).

Model	Accuracy (%)	Precision	Recall	F1-Score
Decision Trees	94.73	0.93	0.94	0.94
Adaboost	96.87	0.97	0.96	0.96
LightGBM	98.54	0.98	0.99	0.98
Hybrid (AdaBoost + LightBGM)	99.12	0.99	0.99	0.99

Table 1. Performance Comparison of Implemented Algorithms

Table 1 provides the performance of four machine learning models based on evaluation criteria like accuracy, precision, recall, and F1-score [3], [4], [5], [7], [10]. Of all individual models, the best performance is achieved by LightGBM with 98.54% accuracy, in addition to high levels of precision (0.98), recall (0.99), and F1-score (0.98) [4], [5]. The next best performer in terms of performance is AdaBoost with 96.87% accuracy, having balance in other measures [3]. On the contrary, the performance of Decision Tree Model is found poor with 94.73% accuracy [7]. It can be clearly seen that our proposed hybrid model (AdaBoost+LightGBM) outperforms others and provides 99.12% accuracy along with excellent values of precision, recall, and F1-score (0.99) [10].

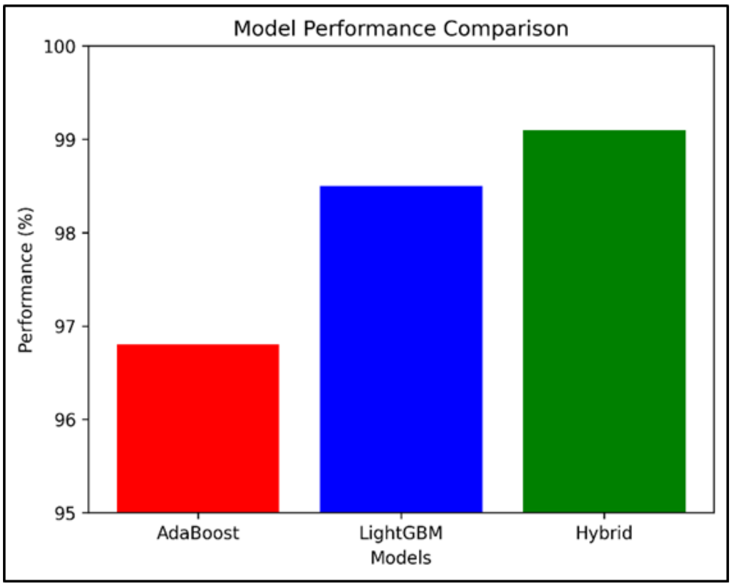

Figure 7. Model Performance Comparison

Figure 7 shows the performance comparison between various machine learning algorithms used for the crop recommendation system. The algorithm comparison includes Decision Tree, AdaBoost, LightGBM, and Hybrid (AdaBoost + LightGBM) machine learning algorithms. The performance is compared using evaluation parameters like accuracy, in which the performance of each algorithm is plotted separately using a bar. It can be seen from the chart above that the hybrid machine learning algorithm has obtained the maximum value of accuracy as compared to other models. While the Decision Tree is the least efficient model, AdaBoost and LightGBM have performed better than others individually, yet they are outperformed by the hybrid model. The high performance of the hybrid model proves that ensembles are quite effective in improving the performance of the model.

Limitations

Despite the fact that the suggested hybrid model yields accurate results, several shortcomings can be identified in this regard. First of all, the model has been designed based on a particular set of training data, and thus its application in other geographical locations with various types of soil and weather cannot be guaranteed. Second, the model considers only certain environmental variables but ignores such important aspects as current market trends and economic considerations. In addition, using historical data may become problematic due to the changing climate.

5. Conclusion and Future Work

This research introduces a novel hybrid crop recommendation system using the stacking ensemble approach by employing AdaBoost and LightGBM along with logistic regression. By incorporating the capabilities of both adaptive boosting and gradient boosting methods, the suggested model leverages the soil nutrients and environmental parameters to provide precise crop recommendations. According to experimental results, the proposed hybrid model outperforms all individual models in terms of accuracy, precision, recall, and F1-score with an accuracy score of 99.12%. The superior performance of the suggested model is due to its capability to identify the difficult examples and capture complex patterns in data using AdaBoost and LightGBM, respectively. Further, the implementation of the model takes place through an interactive web interface that uses Streamlit, providing ease and real-time interactions that are necessary for the actual usage of the model for agricultural purposes. For future works, the inclusion of real-time environment data by use of IoT sensors, crop datasets from different regions, and price analysis would make the model more effective.

References

[1] Soussi, Abdellatif, Enrico Zero, Roberto Sacile, Daniele Trinchero, and Marco Fossa. "Smart sensors and smart data for precision agriculture: a review." Sensors 24, no. 8 (2024): p. 2647.

[2] Bouni, Mohamed, Badr Hssina, Khadija Douzi, and Samira Douzi. "Integrated IoT approaches for crop recommendation and yield-prediction using machine-learning." IoT 5, no. 4 (2024): 634-649.

[3] Senapaty, Murali Krishna, Abhishek Ray, and Neelamadhab Padhy. "A decision support system for crop recommendation using machine learning classification algorithms." Agriculture 14, no. 8 (2024): 1256.

[4] Gunasekaran, Hemalatha, Deepa Kanmani, and Ramalakshmi Krishnamoorthi. "Optimized ensemble learning for enhanced crop recommendations: Leveraging ML for smarter agricultural decision-making." Engineering Proceedings 82, no. 1 (2024): p. 95.

[5] Mahesh, Pavithra, and Rajkumar Soundrapandiyan. "Yield prediction for crops by gradient-based algorithms." Plos one 19, no. 8 (2024): e0291928.

[6] Singh, Gurpreet, and Sandeep Sharma. "Enhancing precision agriculture through cloud based transformative crop recommendation model." Scientific Reports 15, no. 1 (2025): p. 9138.

[7] Shastri, Sourabh, Sachin Kumar, Vibhakar Mansotra, and Rohit Salgotra. "Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence." Scientific Reports 15, no. 1 (2025): 25498.

[8] Sharafat, MD Shaifullah, Nilavro Das Kabya, Rahimul Islam Emu, Mehrab Uddin Ahmed, Jakaria Chowdhury Onik, Mohammad Aminul Islam, and Riasat Khan. "An IoT-enabled AI system for real-time crop prediction using soil and weather data in precision agriculture." Smart Agricultural Technology (2025): 101263.

[9] Hasan, Mahmudul, Md Abu Marjan, Md Palash Uddin, Masud Ibn Afjal, Seifedine Kardy, Shaoqi Ma, and Yunyoung Nam. "Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation." Frontiers in Plant Science 14 (2023): 1234555.

[10] Bakr, Muhammad Abu, Ahmad Jaffar Khan, Sultan Daud Khan, Mohammad Haseeb Zafar, Mohib Ullah, and Habib Ullah. "Evaluation of Learning-Based Models for Crop Recommendation in Smart Agriculture." Information 16, no. 8 (2025): 632.

[11] Prity, Farida Siddiqi, Md Mehadi Hasan, Shakhawat Hossain Saif, Md Maruf Hossain, Sazzad Hossain Bhuiyan, Md Ariful Islam, and Md Tousif Hasan Lavlu. "Enhancing agricultural productivity: a machine learning approach to crop recommendations." Human-Centric Intelligent Systems 4, no. 4 (2024): 497-510.

[12] Ayesha Barvin, P., and T. Sampradeepraj. "Crop recommendation systems based on soil and environmental factors using graph convolution neural network: A systematic literature review." Engineering Proceedings 58, no. 1 (2023): 97.

[13] Mancer, M'hamed, S. Labib, and Soheyb Ayad. "Interpretable crop selection for optimized farming decisions." International Journal of Computing and Digital Systems (2025): 17, no.1

[14] Gautam Bisodiya, K., Ahmed, L., Bathla, M., Sharma, N., Vashisth, S. "A Novel Machine Learning Approach to Design Crop Recommendation System." In International Journal of Novel Research and Development (IJNRD) 2025, Vol. 10, Issue 4-480–481.

[15] Maram, S. R., Geetha, M., Sai, G. P., Chava, J. “Machine Learning Based Crop Recommendation System and Rainfall Prediction”. TIJER - INTERNATIONAL RESEARCH JOURNAL, 10(6), June-2023, Vol.10, Issue 6-639-644.

[16] Rao, PAPPALA MOHAN, and KUNJAM NAGESWAR Rao. "Optimizing crop recommendation systems using advanced deep learning techniques." Journal of Theoretical and Applied Information Technology (2024), 102, no. 23.

[17] Crop Recommendation Dataset - https://www.kaggle.com/code/niteshhalai/crop-recommendation-dataset/input.