

ISSN 1840-4855
e-ISSN 2233-0046

Original scientific article
<http://dx.doi.org/10.70102/afts.2026.1836.897>

A HYBRID ENSEMBLE–DEEP LEARNING SYSTEM FOR MULTI-CROP RECOMMENDATION USING LIGHTGBM, CATBOOST, AND MLP

B. Ammanni^{1*}, Dr. Sreerama Murty Maturi²

^{1*}Department of Computer Science and Engineering, GITAM (Deemed to be University), Hyderabad, India, e-mail: bammanni@gitam.in, orcid: <https://orcid.org/0009-0007-4904-0555>

²Assistant Professor, Department of Computer Science and Engineering, GITAM (Deemed to be University), Hyderabad, India. e-mail: sreeramssit@gmail.com, orcid: <https://orcid.org/0000-0003-0029-1672>

Received: June 11, 2026; Revised: July 22, 2026; Accepted: August 28, 2026; Published: September 29, 2026

SUMMARY

It is essential in the context of precision agriculture to be able to give accurate advice about the types of crops to grow, as the suitability of a crop depends on the complicated, nonlinear interactions between soil fertility, nutrient levels, and climatic conditions. Although the existing body of literature currently evaluates each recommender system on its own, it fails to offer a single framework for comparing classical methods with more advanced learning methods, particularly when it comes to the individual components and the way in which they are combined within a consistent evaluation environment. In this study, Bagging, Random Forest, RBF-SVM, LightGBM, CatBoost, and a Keras multilayer perceptron (MLP) are systematically compared using a standard preprocessing procedure and an 80:20 method of evaluation, with soil pH, nitrogen, phosphorus, potassium, temperature, humidity, and rainfall acting as the features. Missing values are dealt with, outliers are removed, encoding is performed, and z-score normalisation is applied in the same manner in all cases. The accuracies of Bagging, Random Forest and RBF-SVM are 47.97%, 50.41% and 60.16% respectively, while their Top-3 accuracies are 78.86%, 78.05% and 85.37% respectively. LightGBM achieves 92.61% accuracy and 97.13% Top-3 accuracy; CatBoost reaches 90.14% and 97.33%; and the Keras MLP attains 96.84% and 98.50%. Since there is a class imbalance in the test subset, both macro and weighted metrics are included. A probabilistic fusion architecture is proposed for the three more advanced components. However, the final fusion weights and the results from the strict leave-one-component-out procedure are not given in the experimental records provided, so no performance figures are reported for a fused model that is not supported. The results indicate that boosting and neural network models are good choices for ranked crop decision support, on the condition that the full dataset provenance is available and independent geographic validation has been performed.

Key words: crop recommendation, precision agriculture, ensemble learning, LightGBM, CatBoost, multilayer perceptron, decision support system.

INTRODUCTION

Precision agriculture makes use of data-driven decision support in order to take into account the variations in climate, soil conditions, and the suitability of different crops [1][2][3]. The nutrients in the soil, the pH level, temperature, humidity, and rainfall generally interact in a nonlinear manner, and more than one crop may be agronomically suitable for a given field. Multi-crop recommendation also involves a multiclass problem, and offering ranked recommendations gives decision makers a wider range of choices than simply one recommendation. Soil and environmental data have recently been used in studies on decision support, explainable, and ensemble crop recommendation, with positive outcomes [4][5][6]. Research into broader crop prediction, as well as into crop and fertiliser recommendations and cloud-enabled systems, has applied the same objective to different datasets and deployment situations [7][8][9]. Explainable and multicriteria recommender frameworks have been found to increase interpretability and the variety of possible decisions [10][11][12]. Moreover, various studies have used AI, regional predictive methods, and deep learning to improve crop outcomes by means of nonlinear models [13][14][15]. The results have been further improved through the use of LSTM forecasting, IoT-assisted recommendations, and comparative evaluations based on different models, including those in temporal, connected, and model-selection contexts [16][17][18]. These methods include ensemble learning and the application of analytics in precision agriculture to design and test solutions based on soil sensing [19][20][21]. Other research has looked at feature selection, stacking, and cloud-based recommendations, demonstrating that performance depends not only on which data features are selected or which model framework is employed, but also on how the predictions are made and how the solutions are deployed [22][23][24]. It is difficult to compare the performance results from different studies because they use different datasets, have different class balances, cover different geographical areas, and employ different methods for evaluating and presenting the results. Relying exclusively on accuracy can cause the performance for a small number of crops to be overlooked. The benchmark is set by subjecting all the experimental comparisons of the classical (Bagging, Random Forest, and RBF-SVM) and state-of-the-art recommendation components (LightGBM, CatBoost, and Keras MLP) to the same processing and evaluation methodology. The gradient boosting models (CatBoost and LightGBM) will be selected to provide the most up-to-date technology for learning from structured tabular data, and MLP will be chosen for nonlinear representation learning and as the third component of the recommendation system. For each crop type, as well as for all crop types, we will assess accuracy, accuracy in the list of recommended candidates, precision, recall, and the F1-score. The contributions of this study are the following: (1) the creation of a strict benchmark that compares classical and state-of-the-art frameworks in the field of multi-crop recommendation; (2) the development and evaluation of a probabilistic fusion of the recommendations, together with experimental evidence of its effectiveness for CatBoost, LightGBM, and Keras MLP; (3) the assessment of framework performance using Accuracy, Top-3 Accuracy, Precision, Recall, and F1-score for all the classes, with particular attention to those classes that have low support; and (4) the conduct of a component-level analysis by comparing the results of the individual recommender systems with those obtained from the integrated frameworks.

Section 2 examines the previous research in this field, and Section 3 describes the methodology and fusion approach. Section 4 gives the results together with the relevant limitations [4], before the Conclusions go on to discuss and summarise the future objectives, the challenges, and the proposed validation strategies in Section 5.

RELATED WORK

Recent original studies have verified the advantages of combining information on soil and weather when giving crop recommendations. Methods using Random Forest, boosting, and ensemble techniques demonstrate good classification performance [2][4][6]. Explainable recommendation systems increase transparency and help with decision-making [5][10][11]. The use of cloud and IoT technologies allow for large-scale implementation in the field [9][17][20]. Investigations that involve deep learning, LSTM, and comparative models provide additional evidence on nonlinear representation learning [15][16][18]. Ensemble and stacking approaches show complementary improvements amongst the various types of learners [19][23][24]. Research into crop and fertiliser recommendations as well as feature selection extends the range of possible decisions beyond that of a single classifier [8][12][22]. Yet the published

results are still difficult to compare because the datasets, locations, class distributions, and validation methods differ. Table 1 gives a summary of six studies that are directly relevant and matches each entry with its corresponding citation.

Table 1. Summarises a comparison of recent related studies

References	Model	Input Feature and Dataset	Reported Performance	Key Strengths	Limitations
[2]	LSTM-RNN + Random Forest	Weather and Soil Parameters; Telangana, India	97.235% Crop-Selection Accuracy	Strong Joint Use of Weather and Soil Variables	Regional Training Context; External Validation Required
[4]	ML Classification DSS; XGBoost best	Soil and Environmental Crop-Recommendation Attributes	99.31% Best Reported Accuracy	Broad Classifier Comparison and Decision-Support Framing	Benchmark-Data Dependence Limits Geographic Inference
[5]	Gradient Boosting + XAI	Nutrient and Environmental Attributes; 22 crop classes	99.27% Accuracy	High accuracy with explainability	Performance Established on a Defined Benchmark Setting
[6]	RFXG Ensemble (RF + XGBoost)	N, P, K, pH, Rainfall, Humidity, Temperature	98% Accuracy/Precision/Recall/F1	Ensemble Design and Cross-Validation	Balanced-Data Setting; Transferability Requires Further Study
[12]	AgriRec Multicriteria ML	Soil, Water, Farm Size, Market Value, Fertilizer Data	95.85% crop Recommendation; 92.11% Fertilizer	Multicriteria Crop and Fertilizer Decision Support	Region-Specific Gujarat Data and Task Formulation
[23]	Feature Fusion + Stacking/Voting Ensembles	Two Crop-Recommendation Datasets	Up to 99.56% on the Smaller Dataset	Feature Fusion and Multi-Ensemble Validation	Performance Varies Substantially with Dataset Scale

The literature identifies three recurring gaps: inconsistent evaluation procedures, no analysis of the minor crop types, and incomplete evidence concerning the contribution of each component in the hybrid designs. In order to deal with these problems, the current evaluation makes use of a common preprocessing pipeline, class-sensitive metrics, Top-3 accuracy, and comparisons at the component level. Yet it is not possible to carry out a strict leave-one-component-out fusion ablation since the necessary fused outputs were not included in the experimental records.

PROPOSED METHODOLOGY

The procedure includes carrying out data acquisition, preprocessing, developing the model and carrying out evaluation, as well as conducting a comparative analysis. All of the models used receive the same preprocessing and make use of the same train-test partitioning, which means that any differences in performance observed can be attributed to the models' behaviour and the experimental conditions rather than to inconsistencies in the way the data were processed.

Data Collection and Description

The dataset contains soil and environmental variables that were used in the multi-crop recommendation process, including soil pH, nitrogen, phosphorus, potassium, temperature, humidity and rainfall or precipitation. The crop types referred to are Bajra, Paddy, Banana, Bengal Gram, Black Gram, Cotton, Groundnut, Onion, Red Chilli, Sesamum, Soybean, Sunflower, Sweet Orange, Turmeric, Vegetables and

various other categories. The train:test split reported is 80:20. As shown in Tables 3, 5 and 7, the test set comprises 126 cases and displays a significant class imbalance, there being 29 test cases for Paddy, 19 for both Cotton and Black Gram, and several classes having only one or three cases. Because both macro and weighted metrics are considered, the manuscript and the experimental files do not give the exact name of the repository, the total number of the original samples, the period during which the data was collected, or the location where it was collected, and these particulars cannot be determined. It is necessary to obtain confirmation of these provenance details from the original experiment records prior to the submission being made.

Data Preprocessing

When carrying out the preprocessing, missing values are dealt with, the numerical values are standardised, the categorical variables are encoded and the outliers are addressed. The numerical variables are standardised by means of z-score normalisation so that features measured on different scales can make a comparable contribution. The outliers are detected using interquartile-range filtering and the categorical attributes are one-hot encoded where this is necessary. The same preprocessing procedure is applied to all of the models being compared in order to avoid any possible bias in the evaluation [8][9][17].

Model Development

It was possible to carry out a thorough performance analysis since both the conventional machine learning models and the proposed hybrid recommendation system were developed using the same preprocessing and evaluation conditions.

(a) Ensemble Model – Bagging

The individual base learner $h_i(x)$ is fitted on bootstrap sample D_i drawn from D . The final Bagging prediction is defined in equation (1).

$$\hat{y} = \text{majority_vote}\{h_1(x), h_2(x), \dots, h_M(x)\} \quad (1)$$

where M denotes the number of base estimators.

(b) Random Forest

Random Forest introduces feature randomness during tree construction, and its ensemble prediction is defined in equation (2).

$$\hat{y} = \frac{1}{M} \sum_{i=1}^M h_i(x; \Theta_i) \quad (2)$$

where Θ_i represents the random feature and sample subsets associated with the i th tree.

(c) Support Vector Machine (RBF Kernel)

SVM identifies a separating hyperplane with maximum margin by solving the optimization in equation (3).

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \quad (3)$$

The margin constraints are specified in equation (4).

$$y_i(\mathbf{w}^\top \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad (4)$$

The RBF kernel used by the SVM is defined in equation (5).

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2\right) \quad (5)$$

(d) Proposed Hybrid Recommendation System

For input x and crop class y , the final class probability is defined by equation (6). The validation weights α , β , and γ are non-negative and constrained by equation (7).

$$P_{\text{final}}(y | x) = \alpha P_{\text{LGBM}}(y | x) + \beta P_{\text{CatBoost}}(y | x) + \gamma P_{\text{MLP}}(y | x) \quad (6)$$

The fusion-weight constraint is specified in equation (7).

$$\alpha + \beta + \gamma = 1, \alpha, \beta, \gamma \geq 0 \quad (7)$$

Fusion weights are selected exclusively on validation data. Candidate combinations are evaluated on the validation partition, and the best validation configuration is retained before test evaluation. The supplied experimental record does not provide the final numerical values of α , β , and γ . Consequently, numerical fusion weights are neither inserted nor inferred in this revision.

The design proposed combines three sophisticated learners within a single probability-based recommendation system. LightGBM and CatBoost offer complementary boosting capabilities when it comes to structured data, whereas the MLP is responsible for nonlinear representation learning. The hybrid approach is evaluated using Top-1, Top-3, macro, and weighted metrics; however, a strict fused-output ablation is still needed for future validation.

Baseline Models

As classic baseline algorithms, support vector machine (SVM), bagging, and random forest are considered on the basis of the kernel-based approach and ensemble learning.

The Bagging Classifier (using a decision tree estimator) involves training trees on bootstrapped samples and then combining the results in order to reduce variance. Random Forest is similar to bagging in principle, with the difference that it decorrelates the trees by using random subsets of both samples and features and combines the predictions to enhance generalization. The last of the Classical Learning Baselines to be used is the Support Vector Machine. For the SVM an RBF kernel was selected and its hyper-parameters were determined by means of a grid search. The experiment also provides the SVM parameters $C = 6.8479$ and $\gamma = 0.000361$ that are to be used for the base SVM.

Proposed Hybrid Recommendation System

The recommended architecture at the end makes use of LightGBM, CatBoost, and a Keras MLP in order to capture complementary patterns in structured data and to achieve nonlinear learning [18][19][23].

1. LightGBM is a gradient-boosting model based on histograms and is known for its computational efficiency when dealing with structured tabular data and for dealing with high-capacity nonlinear decision boundaries.
2. CatBoost is a boosting algorithm capable of dealing with categorical features and of addressing prediction shift; in particular it makes use of an ordered boosting technique and employs symmetric tree construction.
3. The Keras MLP is a fully connected Multilayer Perceptron which uses ReLU activations and Dropout regularization in order to effectively learn the nonlinear relationships among the variables for each feature relating to soil and weather conditions.

Criteria for Assessment

The criteria used for performance evaluation are Accuracy, Top-3 Accuracy, Precision, Recall, and the F1-score. Since there is a class imbalance in the test dataset, both macro averages (in which the different

crops are given equal weight) and weighted averages (which take into account the frequency of the classes) are provided. Top-3 Accuracy involves checking whether the correct crop is included among the top three predictions, a criterion that is especially important when more than one crop that is agro-economically viable can grow under the same agro-environmental conditions.

Accuracy and the F1-score are given in equations (8) and (9) respectively.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

Top-3 Accuracy is defined in equations (10).

$$\text{Top3} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(y_i \in \text{Top-3 Predictions}) \quad (10)$$

Where $\mathbb{I}(\cdot)$ is the indicator function.

Comparative Evaluation

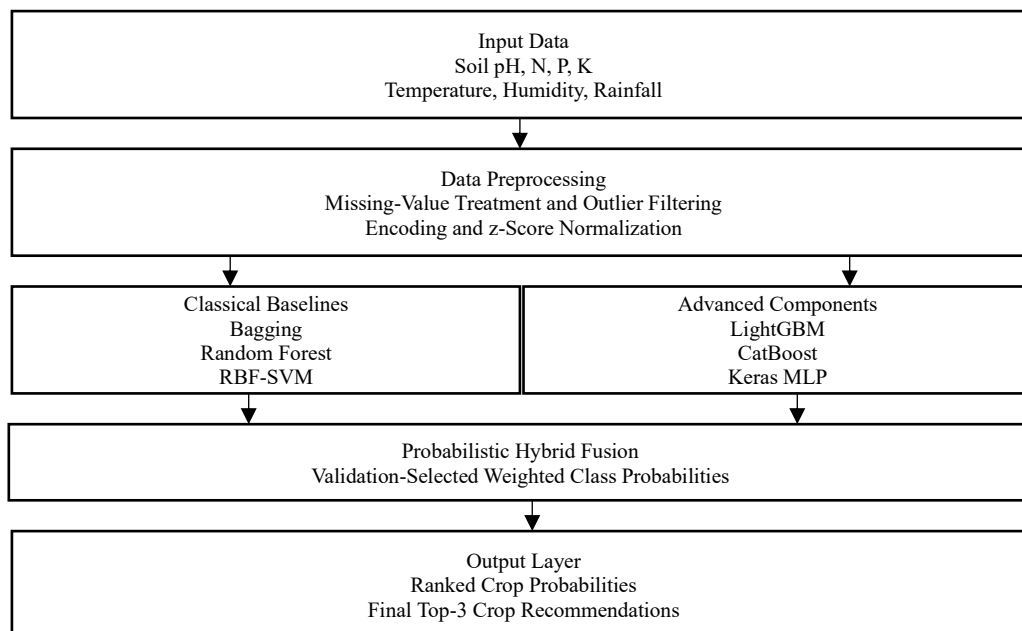


Figure 1. Proposed hybrid crop-recommendation architecture

Using the standard evaluation procedure, Bagging attains an accuracy of 47.97%, Random Forest 50.41% and RBF-SVM 60.16%. The respective Top-3 accuracies are 78.86%, 78.05% and 85.37%. The more advanced components show a marked improvement: LightGBM achieves 92.61% accuracy and 97.13% Top-3 accuracy, CatBoost reaches 90.14% and 97.33%, and the Keras MLP attains 96.84% and 98.50%. These results show that the boosting and neural models are able to capture the nonlinear relationships between soil and environment more effectively on the given test subset.

Because of the highly uneven class distribution in the test set, the overall accuracy is examined together with the macro, weighted, and Top-3 measures. Classes consisting of just one instance might produce extremely high recall figures and should therefore not be regarded as reliable signs of generalization.

Figure 1 shows that the proposed system incorporates the data, preprocessing, baseline, advanced-model, probabilistic-fusion, and ranked-output stages into one crop-recommendation workflow.

The input variables needed for crop recommendation are included in the Data Layer: soil pH, nitrogen, phosphorus, potassium, temperature, humidity, and rainfall.

Such inputs can be obtained from agricultural databases, weather services, and IoT-enabled field sensing platforms, as has been shown in recent precision-agriculture systems [9][17][20].

The preprocessing pipeline includes data validation, the handling of missing values, outlier filtering, encoding, and numerical normalization; z-score normalization is given in equation (11).

$$z = (x - \mu) / \sigma \quad (11)$$

In equation (11), x is the feature value, μ is the feature mean, and σ is the feature standard deviation.

One-hot encoding is used for categorical attributes when it is necessary, and irrelevant or redundant features can be identified by means of mutual information and correlation criteria in order to keep the model pipeline computationally manageable.

The model layer includes the learning algorithms which have been trained on the processed features and is made up of classical baselines and more advanced components.

(a) Classical ML Models

Bagging, Random Forest, and RBF-SVM give ensemble-tree and kernel-based starting points for the usual way of testing.

To reduce variance, aggregates the predictions from estimators that have been trained using the bootstrap method; Random Forest introduces additional randomness with respect to features in order to decorrelate the trees, while RBF-SVM produces nonlinear boundaries by using a radial-basis kernel.

(b) Advanced Components

While LightGBM and CatBoost provide complementary forms of gradient boosting for structured data, the use [18][19][23].

The fully connected hidden layers of the Keras MLP include ReLU activations and dropout regularization, enabling it to learn nonlinear representations $P(y|x)$.

In the probabilistic approach mentioned above, the class probability contributed by each sophisticated component—that is, the Keras MLP, the Keras CNN, and the Keras RNN—is combined using equation (6).

The Hybrid Fusion Layer carries out the weighted aggregation as given in equation (6).

The validation constraint must strictly be satisfied, as illustrated by equation (7); although the exact figures are not given here, the matter is left on purpose without specification.

The way in which the decisions are combined—each one being derived from an approach that models complementary decision behaviour—is intended to improve upon existing standalone solutions; however, all of them still yield a probability output score within a combined framework. For that reason, the results presented in this manuscript compare the individual runs with the fused solution.

The Top-3 crop recommendations are returned by the Output Layer, ranked according to posterior probability. Using a ranked output of this kind is useful since multiple crops may be agronomically suitable. However, other local factors such as the availability of irrigation facilities or rotation conditions, as well as the various risks that might need to be taken into account because of market and seasonal factors, must be considered.

RESULTS AND DISCUSSION

The results obtained using Bagging, Random Forest, and RBF-SVM are compared with those of LightGBM, CatBoost, and the Keras MLP, all of them undergoing the same preprocessing steps and test conditions. As shown in table 9, RBF-SVM is the best-performing method among the classical baselines, while the MLP achieves the highest accuracy and the highest Top-3 accuracy of all the more advanced models.

The aggregate metrics for Bagging are given in table 2 above, showing an accuracy of 47.97 per cent and a Top-3 accuracy of 78.86 per cent, together with a macro-F1 of 0.36 and a weighted F1 of 0.48. figure 2 displays the aggregate scores, while figure 3 in the table 3 shows the per-class results and the distribution of the F1 score.

The fact that the Top-3 accuracy is high shows that the correct crop is often among the top predictions, even if it is not listed first. When it comes to per-class metrics, they should be understood in light of the support counts.

Table 2. Bagging classifier performance metrics

Metric	Value
Accuracy	47.97%
Top-3 Accuracy	78.86%
Macro Avg (Precision, Recall, F1)	0.37 / 0.51 / 0.36
Weighted Avg (Precision, Recall, F1)	0.58 / 0.48 / 0.48

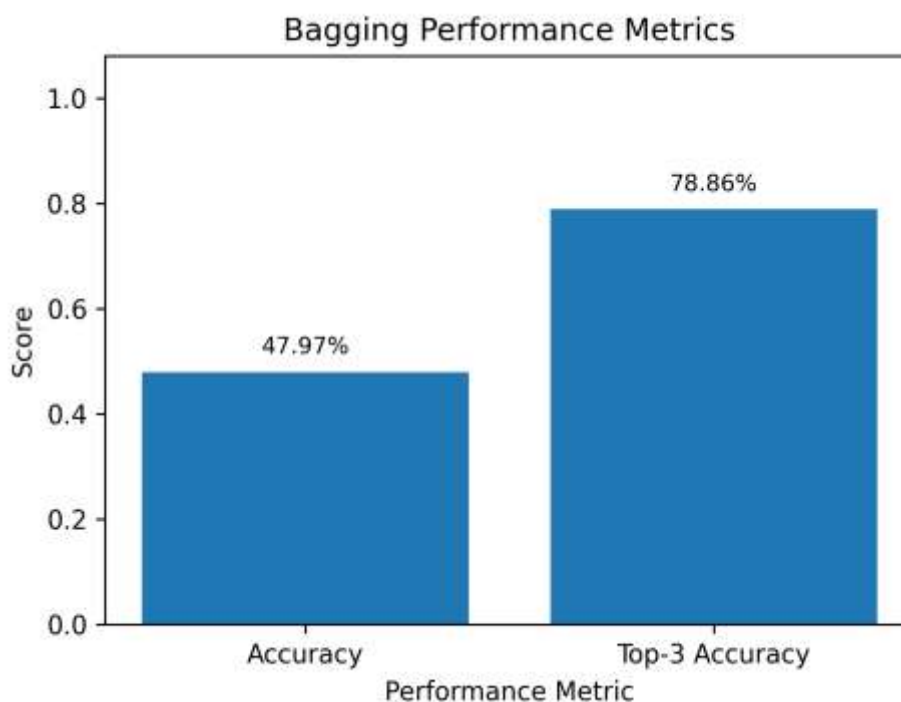


Figure 2. Bagging performance metrics

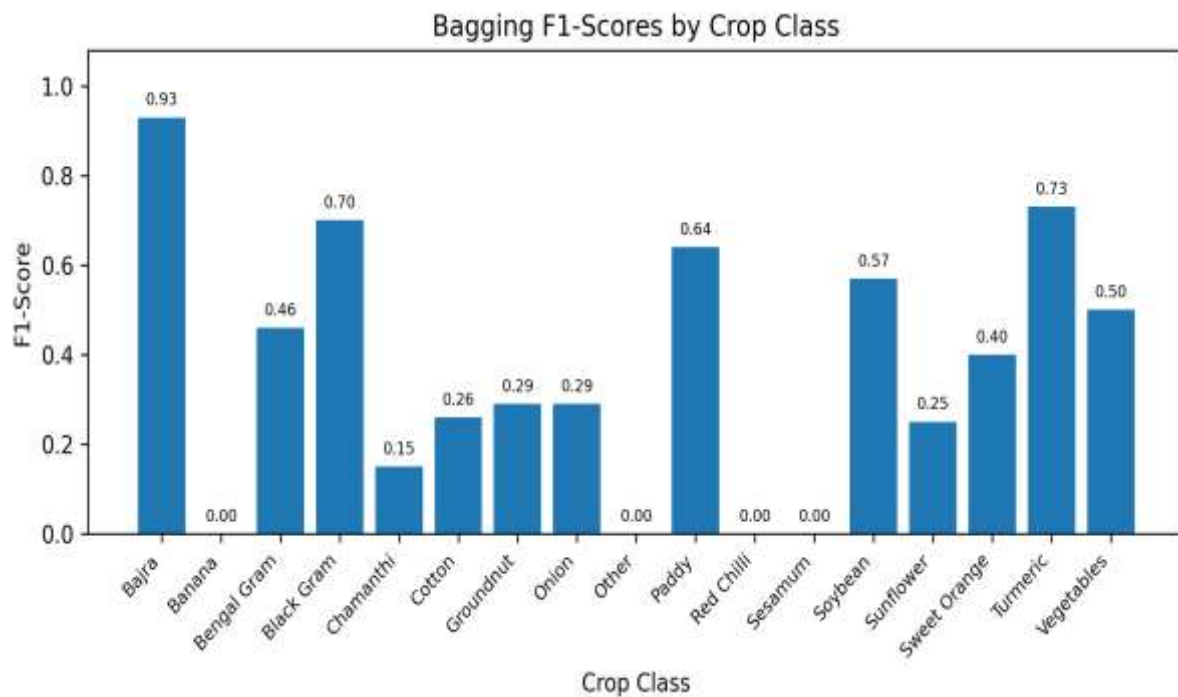


Figure 3. Bagging F1-score by crop class

Table 3. Bagging classifier per-class results

Class	Precision	Recall	F1-Score	Support
Bajra	0.88	1.00	0.93	7
Banana	0.00	0.00	0.00	3
Bengal Gram	0.50	0.43	0.46	7
Black Gram	0.72	0.68	0.70	19
Chamanthi	0.08	1.00	0.15	1
Cotton	0.75	0.16	0.26	19
Groundnut	0.33	0.25	0.29	12
Onion	0.17	1.00	0.29	1
Other	0.00	0.00	0.00	7
Paddy	0.67	0.62	0.64	29
Red Chilli	0.00	0.00	0.00	1
Sesamum	0.00	0.00	0.00	1
Soybean	0.40	1.00	0.57	2
Sunflower	0.20	0.33	0.25	3
Sweet Orange	0.29	0.67	0.40	3
Turmeric	1.00	0.57	0.73	7
Vegetables	0.33	1.00	0.50	1

The results obtained with the random Forest classifier are given in table 4: accuracy equals 50.41 per cent, top-3 accuracy is 78.05 per cent; the macro F1 score is 0.37 and the weighted F1 score is 0.50. Figure 4 displays the aggregated performance. The per-class calculation results are listed in table 5 and the F1-score distribution is shown in figure 5. It performed only slightly better than bagging due to the introduction of feature randomness and the decorrelation effect between the trees. Nevertheless, class imbalance still had a significant impact on the accuracy results.

The random forest method shows better performance when applied to crops that are more widely represented, for example black gram, paddy, and turmeric, while those crops which have limited data continue to be prone to misclassification. The fact that this is the case highlights the need for reporting class support and for using class-sensitive evaluation metrics.

Table 4. Random Forest performance metrics

Metric	Value
Accuracy	50.41%
Top-3 Accuracy	78.05%
Macro Avg (Precision, Recall, F1)	0.37 / 0.52 / 0.37
Weighted Avg (Precision, Recall, F1)	0.57 / 0.50 / 0.50

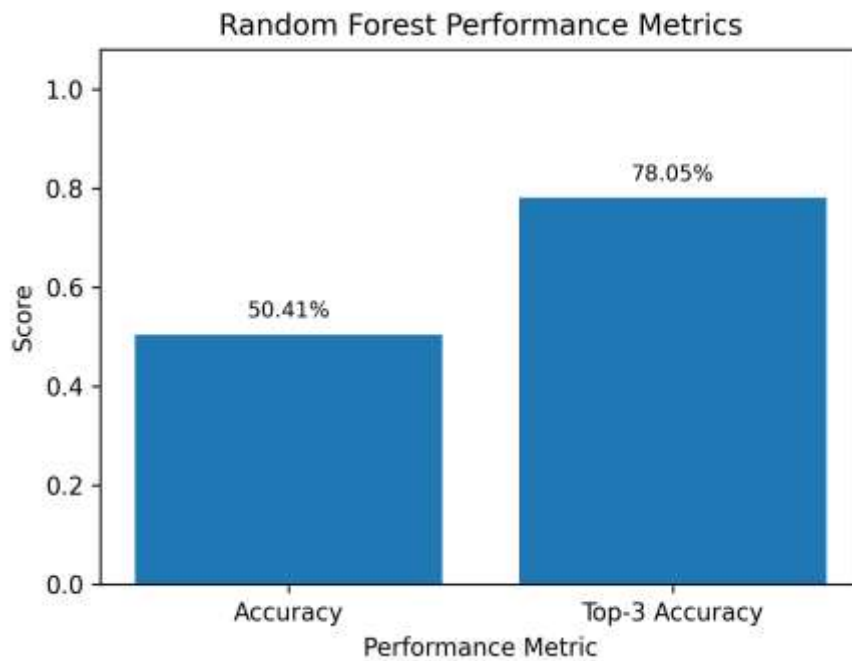


Figure 4. Random Forest performance metrics

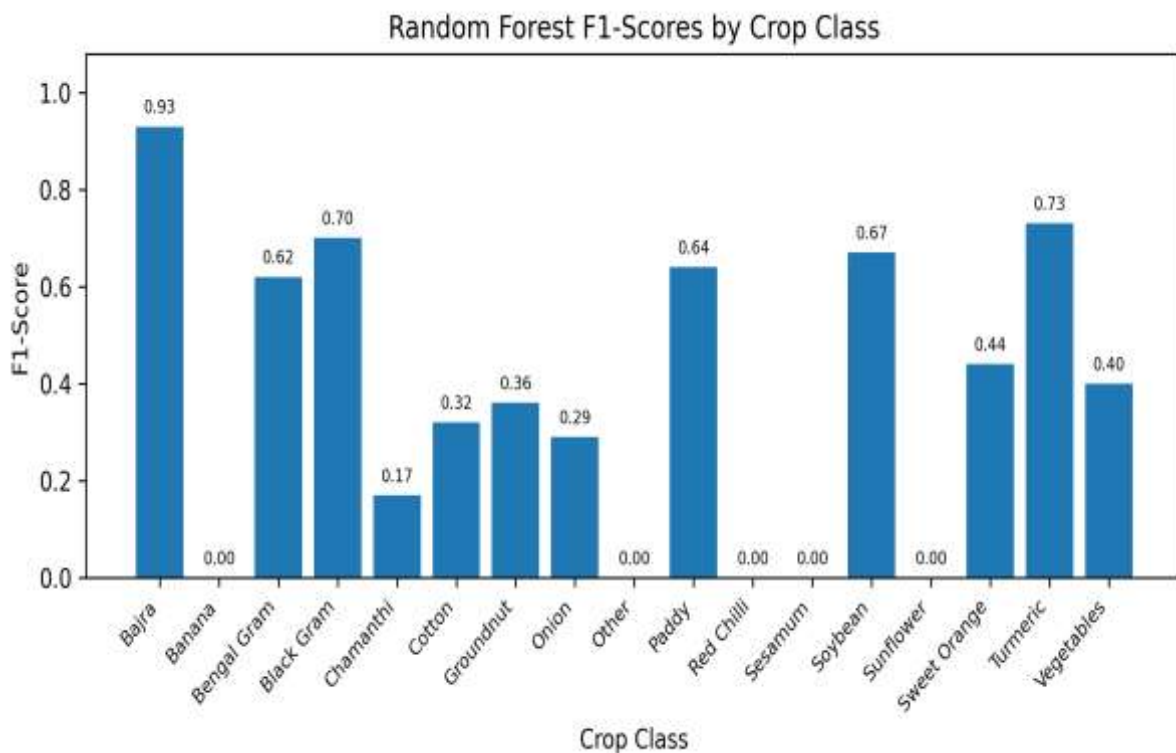


Figure 5. Random Forest F1-score by crop class

Table 5. Random Forest per-class metrics

Class	Precision	Recall	F1-Score	Support
Bajra	0.88	1.00	0.93	7
Banana	0.00	0.00	0.00	3
Bengal Gram	0.56	0.71	0.62	7
Black Gram	0.72	0.68	0.70	19
Chamanthi	0.09	1.00	0.17	1
Cotton	0.67	0.21	0.32	19
Groundnut	0.40	0.33	0.36	12
Onion	0.17	1.00	0.29	1
Other	0.00	0.00	0.00	7
Paddy	0.67	0.62	0.64	29
Red Chilli	0.00	0.00	0.00	1
Sesamum	0.00	0.00	0.00	1
Soybean	0.50	1.00	0.67	2
Sunflower	0.00	0.00	0.00	3
Sweet Orange	0.33	0.67	0.44	3
Turmeric	1.00	0.57	0.73	7
Vegetables	0.25	1.00	0.40	1

Table 6 gives the RBF-SVM performance, with an accuracy of 60.16% and a Top-3 accuracy of 85.37%. The parameters that were chosen are listed below. The score visualization is shown in figure 6. Table 7 displays the per-class metrics and figure 7 shows the F1-score distribution. The fact that the results exceed those of the tree baselines indicates that a kernel (nonlinear) boundary would be effective in this feature space. Bajra, Sesamum, Turmeric, and Vegetables show good performance at the class level, even though some classes with low support are still problematic. Consequently, the macro averages should be interpreted with greater caution than the weighted averages.

Table 6. Support vector machine (RBF kernel) performance metrics

Metric	Value
Accuracy	60.16%
Top-3 Accuracy	85.37%
Macro Avg (Precision, Recall, F1)	0.48 / 0.54 / 0.49
Weighted Avg (Precision, Recall, F1)	0.61 / 0.60 / 0.60

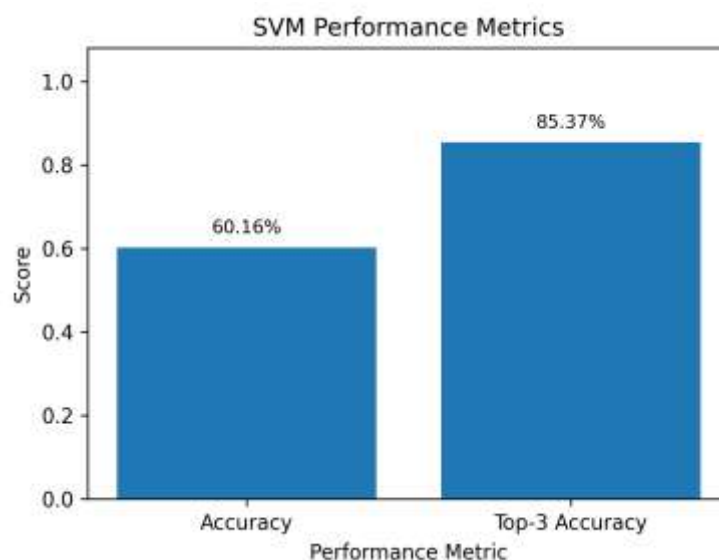


Figure 6. SVM performance metrics

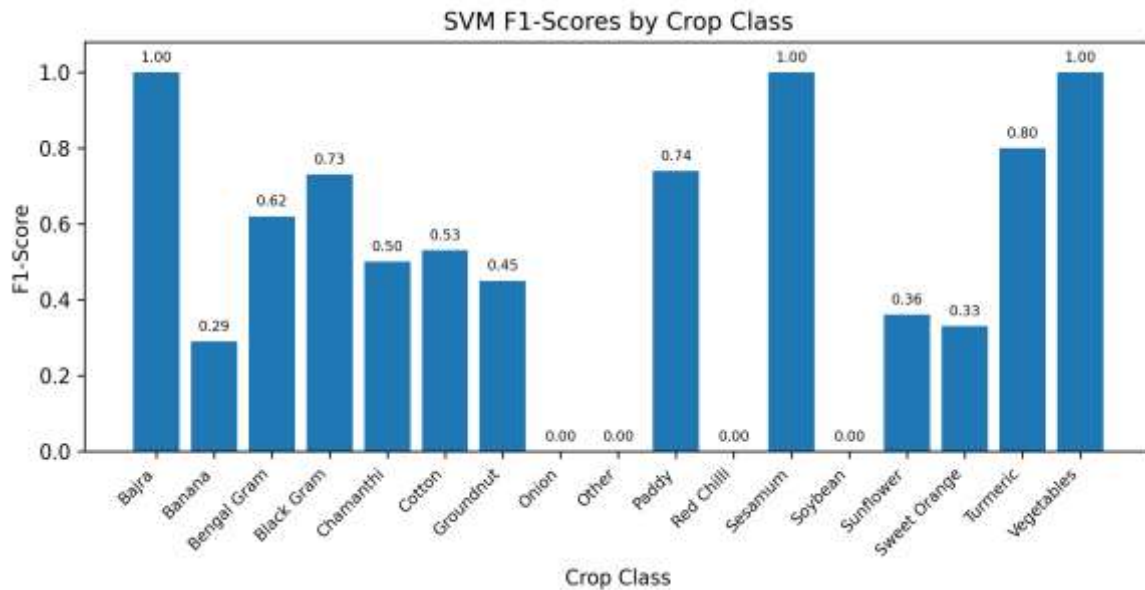


Figure 7. SVM F1-score by crop class

Table 7. SVM per-class metrics

Class	Precision	Recall	F1-Score	Support
Bajra	1.00	1.00	1.00	7
Banana	0.25	0.33	0.29	3
Bengal Gram	0.56	0.71	0.62	7
black gram	0.68	0.79	0.73	19
Chamanthi	0.33	1.00	0.50	1
Cotton	0.73	0.42	0.53	19
Groundnut	0.50	0.42	0.45	12
Onion	0.00	0.00	0.00	1
Other	0.00	0.00	0.00	7
Paddy	0.75	0.72	0.74	29
Red Chilli	0.00	0.00	0.00	1
Sesamum	1.00	1.00	1.00	1
Soybean	0.00	0.00	0.00	2
Sunflower	0.25	0.67	0.36	3
Sweet Orange	0.33	0.33	0.33	3
Turmeric	0.75	0.86	0.80	7
Vegetables	1.00	1.00	1.00	1

The advanced components outperform the classical baseline methods. LightGBM reaches an accuracy of 92.61%, CatBoost 90.14%, and the Keras MLP attains 96.84%. For Top-3 accuracy the figures are 97.13%, 97.33%, and 98.50% respectively. All the results are brought together in Table 9. Although the MLP has the highest individual performance, the gradient boosting models are still competitive and this has led to the decision to use a complementary probabilistic fusion.

Table 8 gives the accuracy and Top-3 accuracy figures in relation to the highest accuracy obtained by any of the best classical models. LightGBM achieves an improvement in accuracy of 32.45 pp, CatBoost an improvement of 29.98 pp, and the MLP an improvement of 36.68 pp. For Top-3 accuracy, LightGBM is better by 11.76 pp, CatBoost by 11.96 pp, and the MLP by 13.13 pp when compared with the best-performing classical RBF-SVM (which had 60.16% accuracy and 73.1% Top-3 accuracy). These figures are the standalone accuracies for each component and are thus not the exact leave-one-component-out results.

The fact that the advanced components have achieved such a strong position in the top three is especially important for providing practical decision support, since several crops could be appropriate under similar

soil and climatic conditions. The ranked recommendations can be combined with other factors such as irrigation capacity, crop rotation, market conditions, and seasonal risk. Comparisons with recently developed high-performing crop-recommendation systems show that it is necessary to take into account the design of the evaluation together with the overall accuracy [5][6][12]. Moreover, recent studies based on ensembles and cloud technology show that the context of the dataset and the deployment settings can have a significant effect on the reported performance [23][24].

The main limitations of this study are a biased test partition, inconsistent or missing dataset provenance metadata, the absence of independent geographic validation, and the lack of strictly performed leave-one-component-out fusion outputs. It is essential to verify from the original experiment records the source of the data repository, the total number of samples, the period during which the data were collected, and the area where the data were collected before the results can be submitted. In order to achieve more robust validation, future research should include stratified repeated cross-validation, confidence intervals, geographically independent datasets, probability calibration, and clear fusion ablation analyses, provided that time allows.

Table 8. Component-level contribution analysis based on reported standalone results

Configuration	Accuracy (%)	Top-3 Accuracy (%)	Gain vs. SVM Accuracy (pp)	Interpretation
SVM Baseline	60.16	85.37	0.00	Strongest classical baseline
LightGBM Only	92.61	97.13	+32.45	Strong structured-data boosting contribution
CatBoost Only	90.14	97.33	+29.98	Complementary ordered-boosting contribution
MLP Only	96.84	98.50	+36.68	Largest standalone nonlinear contribution

Table 9. Comparative performance of advanced and classical models

Model	Accuracy	Top-3 Accuracy	Notes
ML (LightGBM)	92.61%	97.13%	Baseline Boosting Model; Fast Inference
DL (Keras MLP)	96.84%	98.50%	Deep Neural Architecture with Softmax Output
ML-Pro (CatBoost)	90.14%	97.33%	Handles Categorical Data Efficiently
SVM (RBF)	60.16%	85.37%	Best Among Classical Models
Random Forest	50.41%	78.05%	Ensemble Tree Method; Good Interpretability
Bagging	47.97%	78.86%	Variance Reduction Through Bootstrap Aggregation

CONCLUSION

When consistent preprocessing and testing protocols are applied, the results show that the advanced learning methods do better than the traditional ones. The accuracy scores obtained were 47.97% for bagging, 50.41% for the random forest, and 60.16% for the RBF-SVM, whereas those for LightGBM, CatBoost, and the Keras MLP were 92.61%, 90.14%, and 96.84% respectively. This difference is even greater when looking at Top-3 accuracy: the classical models reached 78.86%, 78.05%, and 85.37%, compared with 97.13%, 97.33%, and 98.50% for the advanced models. Of the individual models, the MLP shows the best performance. Compared with the RBF-SVM, the accuracy gains for LightGBM, CatBoost, and MLP are 32.45, 29.98, and 36.68 percentage points respectively. The corresponding increases in Top-3 accuracy are 11.76, 11.96, and 13.13 percentage points.

The importance of class sensitivity is highlighted by the effect of class imbalance. In the test set (N=126: 29 Paddy, 19 for each of Cotton and Black Gram, and only 1 to 3 samples for several of the other classes), Bagging obtains macro and weighted F1-scores of 0.36 and 0.48 respectively, Random Forest gets 0.37 and 0.50, and RBF-SVM reaches 0.49 and 0.60, showing that overall accuracy is not a good measure of the performance on the minority classes. Although the proposed fusion architecture seems promising as a graded crop decision-support system, it is necessary to carry out further validation of the fusion weights and to use strict leave-one-component-out evaluations. Future research should make use of geographically independent datasets with full provenance, apply stratified cross-validation, include 95% confidence intervals, produce calibrated probabilities, carry out explicit fusion ablation, and perform

explainable prediction analysis. Moreover, it should be evaluated whether integrating data from IoT sensors, remote sensing, and weather forecasts leads to robust and context-aware field recommendations.

Additional Information

Funding: No external funding was received for this work.

REFERENCES

- [1] Lobell DB, Schlenker W, Costa-Roberts J. Climate trends and global crop production since 1980. *Science*. 2011 Jul 29;333(6042):616-20. <https://doi.org/10.1126/science.1204531>
- [2] Rani S, Mishra AK, Kataria A, Mallik S, Qin H. Machine learning-based optimal crop selection system in smart agriculture. *Scientific Reports*. 2023 Sep 25;13(1):15997. <https://doi.org/10.1038/s41598-023-42356-y>
- [3] Dey B, Ferdous J, Ahmed R. Machine learning-based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables. *Heliyon*. 2024 Feb 15;10(3). <https://doi.org/10.1016/j.heliyon.2024.e25112>
- [4] Senapaty MK, Ray A, Padhy N. A decision support system for crop recommendation using machine learning classification algorithms. *Agriculture*. 2024 Jul 30;14(8):1256. <https://doi.org/10.3390/agriculture14081256>
- [5] Shastri S, Kumar S, Mansotra V, Salgotra R. Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence. *Scientific Reports*. 2025 Jul 15;15(1):25498. <https://doi.org/10.1038/s41598-025-07003-8>
- [6] Afzal H, Amjad M, Raza A, Munir K, Villar SG, Lopez LA, Ashraf I. Incorporating soil information with machine learning for crop recommendation to improve agricultural output. *Scientific reports*. 2025 Mar 12;15(1):8560. <https://doi.org/10.1038/s41598-025-88676-z>
- [7] Elbasi E, Zaki C, Topcu AE, Abdelbaki W, Zreikat AI, Cina E, Shdefat A, Saker L. Crop prediction model using machine learning algorithms. *Applied Sciences*. 2023 Aug 16;13(16):9288. <https://doi.org/10.3390/app13169288>
- [8] Musanase, C., Vodacek, A., Hanyurwimfura, D., Uwitonze, A., & Kabandana, I. (2023). Data-driven analysis and machine learning-based crop and fertilizer recommendation system for revolutionizing farming practices. *Agriculture*, 13(11), 2141. <https://doi.org/10.3390/agriculture13112141>
- [9] Thilakarathne NN, Bakar MS, Abas PE, Yassin H. A cloud enabled crop recommendation platform for machine learning-driven precision farming. *Sensors*. 2022 Aug 22;22(16):6299. <https://doi.org/10.3390/s22166299>
- [10] Shams MY, Gamel SA, Talaat FM. Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making. *Neural Computing and Applications*. 2024 Apr;36(11):5695-714. <https://doi.org/10.1007/s00521-023-09391-2>
- [11] Naga Srinivasu P, Ijaz MF, Woźniak M. XAI-driven model for crop recommender system for use in precision agriculture. *Computational Intelligence*. 2024 Feb;40(1):e12629. <https://doi.org/10.1111/coin.12629>
- [12] Patel K, Patel HB. Multi-criteria Agriculture Recommendation System using Machine Learning for Crop and Fertilizers Prediction. *Current Agriculture Research Journal*. 2023 Apr 1;11(1):137. <https://doi.org/10.12944/CARJ.11.1.12>
- [13] Apat SK, Mishra J, Raju KS, Padhy N. An artificial intelligence-based crop recommendation system using machine learning. *Journal of Scientific & Industrial Research (JSIR)*. 2023 May 12;82(05):558-67. <https://doi.org/10.56042/jsir.v82i05.1092>
- [14] Ed-Daoudi R, Alaoui A, Ettaki B, Zerouaoui J. A predictive approach to improving agricultural productivity in morocco through crop recommendations. *International Journal of Advanced Computer Science and Applications*. 2023;14(3): 199–205. <https://doi.org/10.14569/IJACSA.2023.0140322>
- [15] Mythili K, Rangaraj R. Deep learning with particle swarm based hyper parameter tuning based crop recommendation for better crop yield for precision agriculture. *Indian Journal of Science and Technology*. 2021 May;14(17):1325-37. <https://doi.org/10.17485/IJST/v14i17.450>
- [16] Mahale Y, Khan N, Kulkarni K, Wagle SA, Pareek P, Kotecha K, Choudhury T, Sharma A. Crop recommendation and forecasting system for Maharashtra using machine learning with LSTM: a novel expectation-maximization technique. *Discover Sustainability*. 2024 Jun 28;5(1):134. <https://doi.org/10.1007/s43621-024-00292-5>
- [17] Bouni M, Hssina B, Douzi K, Douzi S. Integrated IoT approaches for crop recommendation and yield-prediction using machine-learning. *IoT*. 2024 Sep 30;5(4):634-49. <https://doi.org/10.3390/iot5040028>
- [18] Bakr MA, Khan AJ, Khan SD, Zafar MH, Ullah M, Ullah H. Evaluation of learning-based models for crop recommendation in smart agriculture. *Information*. 2025 Jul 24;16(8):632. <https://doi.org/10.3390/info16080632>

- [19] Gunasekaran H, Ramalakshmi K, Debnath S, Swaminathan DK. Physics-aware ensemble learning for superior crop recommendation in smart agriculture. *Sensors*. 2025 Oct 9;25(19):6243. <https://doi.org/10.3390/s25196243>
- [20] Akhter, R., & Sofi, S. A. (2022). Precision agriculture using IoT data analytics and machine learning. *Journal of King Saud University-Computer and Information Sciences*, 34(8), 5602-5618. <https://doi.org/10.1016/j.jksuci.2021.05.013>
- [21] Varshitha DN, Choudhary S. An artificial intelligence solution for crop recommendation. *Indonesian Journal of Electrical Engineering and Computer Science*. 2022 Mar;25(3):1688-95. <https://doi.org/10.11591/ijeecs.v25.i3.pp1688-1695>
- [22] Suruliandi A, Mariammal G, Raja SP. Crop prediction based on soil and environmental characteristics using feature selection techniques. *Mathematical and Computer Modelling of Dynamical Systems*. 2021 Jan 2;27(1):117-40. <https://doi.org/10.1080/13873954.2021.1882505>
- [23] Ahmed RA, El-Shafai W, Ahmed ZA, El-Rabaie ES, El-Samie FE. High-precision crop recommendation system with stacking ensemble classifiers for optimizing agricultural productivity. *Scientific reports*. 2025 Dec 9;15(1):43377. <https://doi.org/10.1038/s41598-025-09640-5>
- [24] Singh G, Sharma S. Enhancing precision agriculture through cloud based transformative crop recommendation model. *Scientific Reports*. 2025 Mar 17;15(1):9138. <https://doi.org/10.1038/s41598-025-93417-3>