2026 IEEE International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE 2026)
29–31 January 2026, Rajshahi-6204, Bangladesh

XAI-Crop: Explainable Counterfactual Crop Recommendation via Ensemble Learning

Abdullah Al Mahmud Joy*, Md. Faizer Islam†, Arifur Rahman‡, Md. Sunzid Islam§, Md. Mamun Hossain¶
Department of Computer Science and Engineering
Bangladesh University of Business and Technology, Dhaka, Bangladesh
abdullahalmahmudjoy39@gmail.com*, faizarislam@gmail.com†, arifurrahmanarif2003@gmail.com‡,
ebnajobaersunzid@gmail.com§, mamunhasan.cs@gmail.com¶

***Abstract*—Timely, data-driven crop selection is critical to crop productivity. While predictive accuracy is frequently emphasised in traditional crop recommendation systems, interpretability and actionable guidance features are lacking. This paper introduces XAI-Crop Ensemble, an interpretable and ensemble-based crop recommendation framework that uses a weighted soft-voting ensemble to achieve accuracy and robustness by integrating high-performance machine learning models, namely Random Forest, XGBoost, LightGBM, and CatBoost. A balanced dataset with seven agronomic features—nitrogen (N), phosphorus (P), potassium (K), temperature, humidity, pH, and rainfall—across 22 crop types was used to train and validate the framework. The ensemble outperformed or matched the best single model (CatBoost) with a classification accuracy of 99.77%.Temperature and pH had moderate effects, while humidity, potassium (K), and phosphorus (P) were the most influential features, according to SHapley Additive exPlanations (SHAP). DiCE and manual simulation counterfactual analyses produced useful farmer-level insights, such as modifying rainfall (+68 mm), N (+24 kg/ha), P (+25 kg/ha), and K (+37 kg/ha) to change the suitability of growing papaya instead of oranges. Consistent performance across all crop categories was confirmed by the fairness evaluation.All things considered, the XAI-Crop Ensemble provides an open and equitable framework that makes data-driven and understandable agricultural decision-making possible.**

Keywords: Crop recommendation*, *Explainable AI*, *SHAP*, *Ensemble learning* , *Counterfactual analysis*, *CatBoost*, *Sustainable agriculture

I. Introduction

Food security and economic growth are still largely dependent on agriculture, especially in developing nations like Africa. Precision farming is one of the smarter, data-driven methods that are required due to the world's expanding population and growing environmental uncertainty[1]. Data-informed crop selection is made possible by artificial intelligence (AI) driven by machine learning (ML), which examines rainfall patterns, climate, and soil nutrients[2]. But even with their high predictive accuracy, the majority of machine learning models are opaque and provide little insight into how predictions are made. This opacity restricts their use in agriculture, where the model's usability and credibility depend on interpretable and actionable feedback. This limitation is addressed by Explainable Artificial Intelligence (XAI) methods like SHAP, LIME, and DiCE, which measure how much each input feature contributes to the model's conclusion[3]. Though SHAP and LIME concentrate on the reasons behind decisions, counterfactual reasoning bridges the gap between predictive insight and practical decision support by extending explainability to how decisions can change through "what-if" scenarios (e.g., increasing nitrogen and rainfall could make papaya more suitable than orange).

In this work, we introduce XAI-Crop, a hybrid framework that generates transparent, equitable, and useful crop recommendations by fusing interpretability, predictive accuracy, and counterfactual analysis. In addition to identifying the best crops, XAI-Crop provides comprehensible and fact-based suggestions for sustainable farming methods by combining CatBoost with SHAP-based feature attribution and DiCE-driven counterfactual explanations[4].

To ensure fairness and transparency in agricultural AI, XAI-Crop combines counterfactual "what-if" reasoning, SHAP interpretability, and CatBoost accuracy in a unique way to offer explanations and practical crop improvement strategies[5].

This study's primary contributions include:

- A XAI framework that combines counterfactual reasoning and ensemble learning to provide highly accurate and interpretable decision support.
- Actionable agronomic interventions are produced by a two-stage counterfactual explanation mechanism that combines automated and manually simulated analyses based on DiCE.
- Thorough assessment of interpretability and fairness for 22 crop types that show reliable performance in actual agronomic settings.

The structure of the paper is as follows: Background study on crop recommendation Section II. Our dataset,

979-8-3315-6135-2/26/$31.00 ©2026 IEEE

preprocessing procedures, and model architectures are described in Section III. SHAPE visualisations, confusion matrices, and quantitative results are presented in Section IV. Lastly, implications, limitations, and future research directions are covered in Section V.

II. Related Works and Research Gaps

In order to maximise prediction accuracy based on soil nutrients and environmental features, crop recommendation systems have been extensively researched using a variety of machine learning (ML) techniques. For class balancing, Shraban *et al.*[6] used CatBoost with SMOTE, attaining a 99.51% accuracy rate. M. Karthik *et al.*[7] achieved 96.5% accuracy by combining Random Forest with LIME to enhance interpretability. Shams *et al.*[8] shows an Improved Random Forest, Gradient Boosting, Decision Tree model that attained 94.15% accuracy. Likewise, Kumari *et al.*[9] assessed a number of machine learning algorithms, including Decision Tree, Random Forest, Naïve Bayes, SVM, and XGBoost, with hybrid ensembles yielding up to 99% accuracy. They were all assessed using the same dataset, as shown in Table I.

TABLE I
Comparative Analysis of Existing Studies

Study - year	**Technique**	**Results**
Chana et al. [10] - 2023	RandomForest	Accuracy: 99%
Medagbe et al. [11] - 2025	RandomForest, KNN, SVM	Accuracy: 99.22%
Umamaheswari et al. [12] - 2025	Parallel Random Forest (PRF)	Accuracy: 89.7%, Precision: 88.6%, Recall: 87.5%
Musanase et al. [13] - 2023	ML models (DT, RF, NB, SVM, XGBoost) and NN	Accuracy: 97%
Bhati et al. [14] - 2023	RandomForest, DT	Accuracy: 95%
Dahiphale et al. [15] - 2025	RandomForest	Accuracy: 99.5%
M VAMSI et al. [16] - 2025	Random Forest	Accuracy: 99.3%
H Gunasekaran et al. [17] - 2024	LR, SVM, DT, NB, KNN, RF and Stacking Ensemble	Accuracy: 99.36%
H Gunasekaran et al. [18] - 2025	LR, SVM, DT, NB, KNN, and stacking ensemble model with RF, Extra Tree, and XGBoost	Accuracy: 99.5%
S Shastri et al. [19] - 2025	KNN, NN, LDA, DT, RF, NB, SVM, QDA, LR	Accuracy: 99.27%, Precision: 99.32%, Recall: 99.36%

A. Cartolano *et al.*[20] used SVM, XGB, and MLP and reported an accuracy of 99.32%. With an accuracy of 98.99%, Nirajan Acharya *et al.*[21] created an ensemble soft-voting model that combined Naïve Bayes, SVM, Decision Tree, and Random Forest.

Even though these studies show excellent predictive performance, they ignore the opaqueness of ML models in favour of concentrating only on accuracy. Few studies have looked at counterfactual reasoning or Explainable AI (XAI) to improve interpretability and offer useful insights. This gap is filled by the suggested **XAI-Crop** framework, which combines counterfactual reasoning, SHAP-based explanations, and CatBoost accuracy to produce transparent and equitable crop recommendations.

III. Methodology

A. Dataset Analysis

We used the Crop Recommendation Dataset by Atharva Ingle[22], which is 2,200 instances across 22 crop classes and was made available on Kaggle (2020), for our study. The seven continuous features that make up each entry are temperature, humidity, pH, rainfall, phosphorus (P), potassium (K), and nitrogen (N). The dataset provides 100 samples for each of the 22 crop types, thus making it possible to conduct a fair and balanced representation of the data. As shown in Fig. 1, the heatmap illustrates the correlation among the features in the dataset.

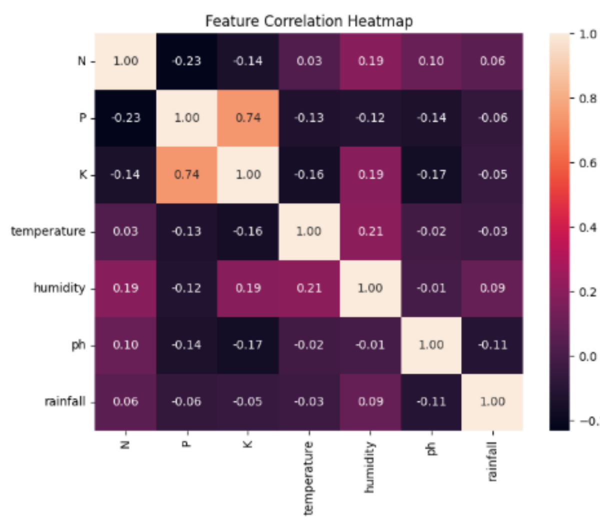

Fig. 1. Correlation heatmap of the Crop Recommendation Dataset's seven input features: temperature, humidity, pH, rainfall, N, P, and K.

B. Data Preprocessing

The dataset underwent scrutiny to check for the absence or inconsistency of data, and it was found to be clean. Crop labels were converted into numbers with the help of the LabelEncoder, while the continuous features (N, P, K, Temperature, Humidity, pH, and rainfall) were standardized by StandardScaler to be on the same scale as the other variables. The division of the stratified train-test set of 80:20 maintained the distribution of classes intact throughout the 22 crops. A 5-fold Stratified Cross-Validation was performed to ensure trustworthiness and combat over-fitting. Preprocessing tools (scalers and encoders) were stored to ensure the model reproducibility. For reproducibility, the complete codebase, trained model weights, and preprocessing scripts will be made publicly available through an open-access repository, upon acceptance. This ensures transparency and enables Other researchers should validate and extend our study.

C. System Overview

The general workflow of the suggested XAI-Crop framework is shown in Fig. 2, emphasising the integration of explainable AI analysis, ensemble modelling, and data preprocessing.

D. Model Selection

We compared four of the latest classifiers: Random Forest, XGBoost, LightGBM, and CatBoost. Among them, CatBoost was the best as it had the best capabilities in dealing with complex feature interactions and running regularization methods to prevent overfitting. The hyperparameter tuning process was performed using Optuna, optimizing parameters such as the learning rate, depth, and iterations.

1) CatBoost

CatBoost is a gradient boosting algorithm developed by Yandex that automatically handles categorical features and prevents overfitting using ordered boosting techniques. Even on datasets with a large number of categorical variables, it performs well and requires little data preprocessing.

2) Ensemble Model

The Ensemble Model improves prediction accuracy and generalisation by combining several classifiers through a weighted soft voting mechanism. Overfitting is successfully decreased, and the complementary strengths of individual learners are utilised by combining models like Random Forest, XGBoost, LightGBM, and CatBoost, which improves performance overall.

3) Performance Evaluation Setup

Every experiment was carried out on a workstation running Python 3.12 and outfitted with an Intel® Core™ i7 processor (2.6 GHz) and 16 GB of RAM. Scikit-learn 1.4.2, XGBoost 2.1, LightGBM 4.3, and CatBoost 1.2 libraries were used in the implementation.

Four common classification metrics—accuracy (ACC), precision (P), recall (R), and F1-Score (F1)—formed the basis of the performance evaluation. The F1-Score provided the harmonic balance between precision and recall, recall quantifies the sensitivity to true positives, accuracy measures overall correctness, and precision assesses the dependability of positive predictions.

Matplotlib 3.9 and Seaborn 0.13 were used to create confusion matrices and SHAP-based feature importance visualisations for interpretability and diagnostic evaluation, guaranteeing transparent and repeatable outcomes.

E. Explainability: SHAP

SHapley Additive exPlanations (SHAP) were calculated on the final Ensemble Model to measure both local and global feature contributions in order to improve interpretability. The most important characteristics influencing crop predictions are highlighted in the global SHAP summary plot (Fig. 3). At the local level, SHAP values offer clear insights into model behaviour across instances by elucidating how differences in individual features impact particular crop classifications.

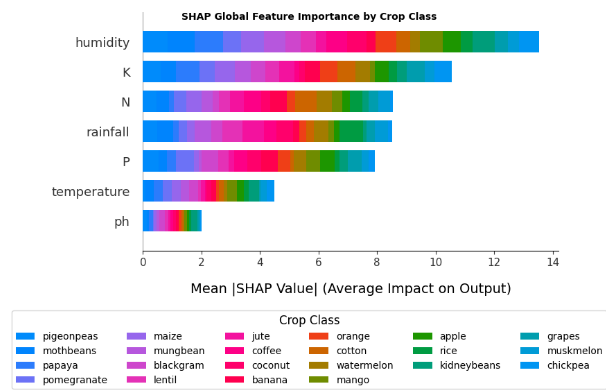

Fig. 3. Global SHAP summary plot showing input feature impact and relative importance on the ensemble crop prediction model.

F. Counterfactual Explainability

Two approaches were implemented:

Automated DiCE Counterfactuals: Different "what-if" scenarios that could change the anticipated crop classes were automatically generated using the DiCE framework. The model's sensitivity to important agronomic parameters was demonstrated, for example, when slight changes in rainfall and nitrogen (N) caused the prediction to change from orange to papaya.

Manual Counterfactual Simulation: To improve the domain realism and reproducibility, a manual simulation technique was implemented. The mean feature deltas were calculated and mapped to practical agronomic interventions by using binary filtering between particular crop pairs.

IV. Results & Discussion

Four baseline machine learning models for crop recommendation were compared in this study: CatBoost, Random Forest, XGBoost, and LightGBM. With an accuracy of 99.77%, CatBoost outperformed the other classifiers, closely followed by Random Forest (99.55%), while XGBoost and LightGBM achieved 98.86%. Every

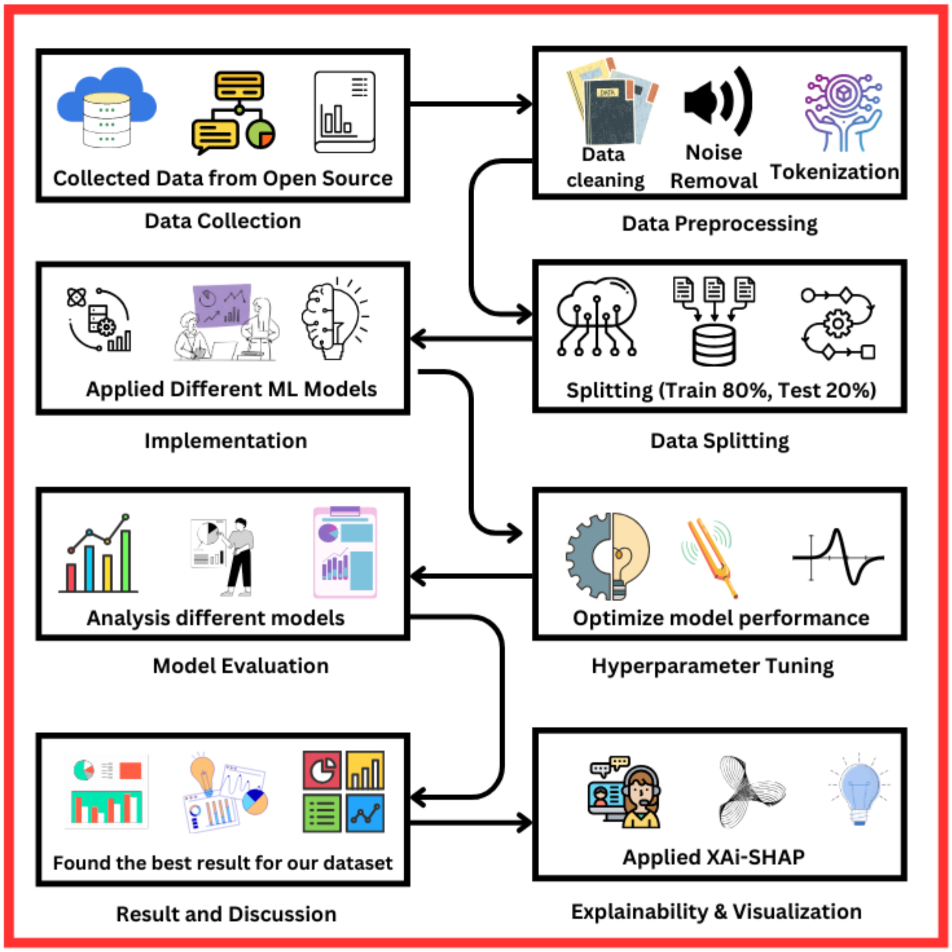

Fig. 2. Overall workflow of the suggested XAI-Crop framework, showing explainability, ensemble learning, and data preprocessing components.

model showed robustness in capturing the relationships between agronomic and climatic features, as evidenced by their high precision and recall across the 22 crop categories in Table II.

TABLE II
COMPARATIVE PERFORMANCE METRICS OF THE MODELS

Model	Accuracy	Precision	Recall	F1-Score
Ensemble Model	**0.9977**	**0.9978**	**0.9977**	**0.9977**
CatBoost	0.9977	0.9978	0.9977	0.9977
RandomForest	0.9955	0.9957	0.9955	0.9955
XGBoost	0.9886	0.9894	0.9886	0.9885
LightGBM	0.9886	0.9891	0.9886	0.9886

The four models were combined to create a weighted soft voting ensemble, which further improved the predictive stability. The ensemble's ability to combine complementary model strengths and enhance generalisation performance is demonstrated by its identical top accuracy of 99.77% (Fig. 4).

TABLE III
DETAILED RESULTS OF 5-FOLD VALIDATION (IN %)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Random Forest	100.00	99.72	98.86	98.86	99.43
XGBoost	99.15	98.86	99.72	99.15	98.86
LightGBM	99.15	99.43	98.58	99.72	98.30
CatBoost	99.15	99.72	98.01	98.86	99.15
Ensemble Model	100.00	100.00	99.15	99.43	99.15

With the exception of one misclassification in the lentil category, the soft voting ensemble accurately predicted every instance, achieving nearly perfect classification across all 22 crop categories. This result demonstrates the ensemble's strong generalisation ability, stability, and dependability in multi-class crop prediction (Table III).

The feature-level contribution of a representative instance predicted to be orange is depicted in the SHAP waterfall visualisation (Fig. 5). With the highest positive

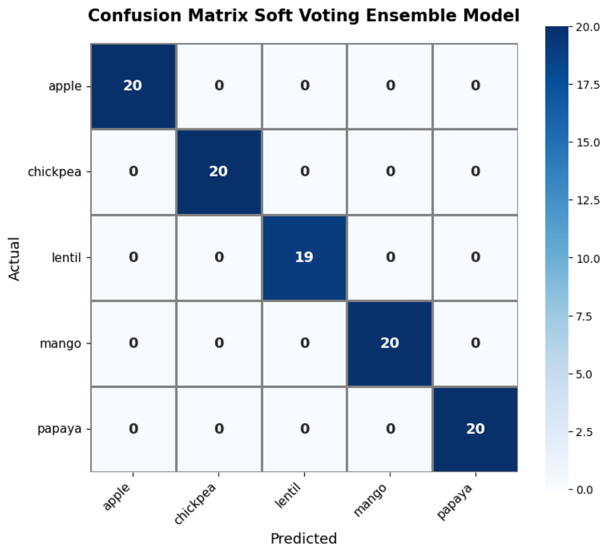

Fig. 4. Weighted soft-voting ensemble model for crop classification confusion matrix.

SHAP value, humidity had the biggest impact on the model's judgements. Additionally, potassium (K) and phosphorus (P) made positive contributions, supporting the prediction result. Temperature and nitrogen (N), on the other hand, had marginally detrimental effects that somewhat offset prediction confidence.Overall, the visualisation showed that the main factors influencing the ensemble model's crop classification decision were high humidity and balanced nutrient conditions.

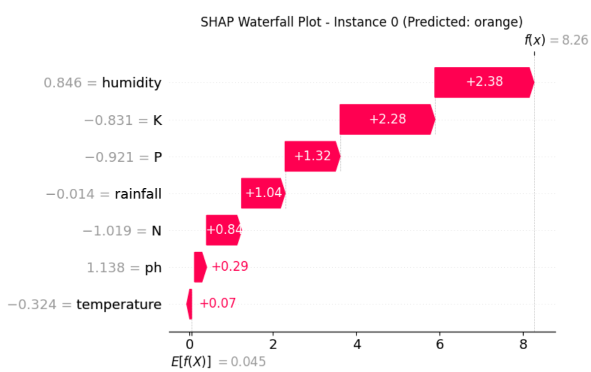

Fig. 5. SHAP waterfall plot showing feature-level contributions for an orange-classified representative instance.

Counterfactual Results: Counterfactual analysis revealed that **+24 kg/ha N**, **+25 kg/ha P**, **+37 kg/ha K**, and **+68 mm of rainfall** were necessary increases in the *orange* $\rightarrow$ *papaya* case. These modifications imply that growing papayas requires more moisture and higher nutrient levels. The results offer practical advice for maximising irrigation and soil fertility to accomplish desired crop transitions.

The interpretability and usability of the suggested model are further supported by a comparative visualisation of these feature variations (Fig. 6), which shows how counterfactual adjustments correspond to actual farming actions.

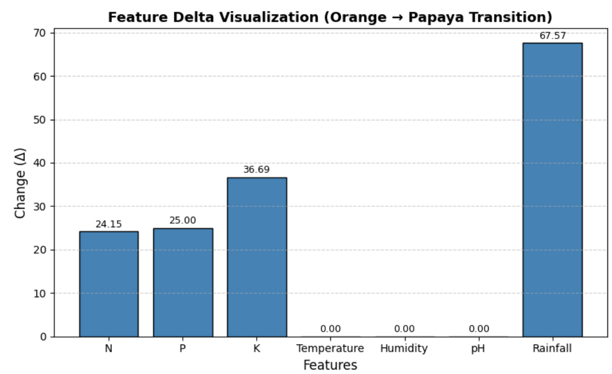

Fig. 6. Visualization of feature variations resulting from counterfactual adjustments and their correspondence to practical farming actions

Model Fairness: To confirm the consistency of the model behaviour, a fairness assessment was carried out for each of the 22 crop classes. With a precision variance of 0.0476 and a recall variance of 0.05, the ensemble model maintained negligible disparity, suggesting that no crop class was misclassified or disproportionately favoured. This demonstrates that the suggested framework is appropriate for real-world agricultural applications since it not only attains high accuracy but also guarantees equitable predictive performance.

V. Conclusion & Future Directions

This study presented XAI-Crop, an interpretable framework for crop recommendation that blends counterfactual reasoning, SHAP explainability, and CatBoost-based prediction. With a 99.7% accuracy rate, the model consistently showed fairness across 22 crop categories.

By pinpointing exact changes in nitrogen, phosphorus, K, and rainfall to maximise crop suitability, the counterfactual module provides actionable insights beyond conventional feature-importance techniques.

Future research will concentrate on implementing web advisory platforms, integrating IoT-based real-time sensing, and integrating regional soil-climate data. Overall, XAI-Crop bridges predictive modeling with human-understandable decision support, thereby advancing interpretable precision agriculture.

References

[1] R. J. Martin, R. Mittal, V. Malik, F. Jeribi, S. T. Siddiqui, M. A. Hossain, and S. Swapna, "Xai-

powered smart agriculture framework for enhancing food productivity and sustainability," *IEEE Access*, 2024.

[2] M. M. Hossain, M. A. Rahman, S. Chaki, H. Ahmed, and M. S. Rahman, "Smart-agri: A smart agricultural management with iot-ml-blockchain integrated framework," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 7, 2023.

[3] V. Vishwarupe, P. M. Joshi, N. Mathias, S. Maheshwari, S. Mhaisalkar, and V. Pawar, "Explainable ai and interpretable machine learning: A case study in perspective," *Procedia Computer Science*, vol. 204, pp. 869–876, 2022.

[4] A. M. Salih, Z. Raisi-Estabragh, I. B. Galazzo, P. Radeva, S. E. Petersen, K. Lekadir, and G. Menegaz, "A perspective on explainable artificial intelligence methods: Shap and lime," *Advanced Intelligent Systems*, vol. 7, no. 1, p. 2400304, 2025.

[5] S. Kawakura, M. Hirafuji, S. Ninomiya, and R. Shibasaki, "Analyses of diverse agricultural worker data with explainable artificial intelligence: Xai based on shap, lime, and lightgbm," *European Journal of Agriculture and Food Sciences*, vol. 4, no. 6, pp. 11–19, 2022.

[6] S. K. Apat, J. Mishra, K. S. Raju, and N. Padhy, "An artificial intelligence-based crop recommendation system using machine learning," *Journal of Scientific & Industrial Research (JSIR)*, vol. 82, no. 05, pp. 558–567, 2023.

[7] M. Karthik, A. Nandini, and P. V. Vedha, "A data-driven crop recommendation system with explainable ai for precision agriculture," in *2025 Third International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*. IEEE, 2025, pp. 774–779.

[8] M. Y. Shams, S. A. Gamel, and F. M. Talaat, "Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making," *Neural Computing and Applications*, vol. 36, no. 11, pp. 5695–5714, 2024.

[9] P. Kumari, G. Chandan, D. Agarwal, S. Joshi, and P. Vekariya, "Combining predictive models: Hybrid approaches in crop recommendation systems."

[10] A. M. Chana, B. Batchakui, and B. B. Nges, "Real-time crop prediction based on soil fertility and weather forecast using iot and a machine learning algorithm," *Agricultural Sciences*, vol. 14, no. 5, pp. 645–664, 2023.

[11] Y.-C. N. Medagbe, O. V. Ocama, and S. K. Kasereka, "Leveraging smart crop recommendation systems and climate-resilient practices for sustainable agriculture in developing regions: A brief review," *Procedia Computer Science*, vol. 265, pp. 209–216, 2025.

[12] R. Umamaheswari and E. Kannan, "Empowering precise crop recommendation," *International Research Journal of Multidisciplinary Technovation*, vol. 7, no. 2, pp. 50–59, 2025.

[13] C. Musanase, A. Vodacek, D. Hanyurwimfura, A. Uwitonze, and I. Kabandana, "Data-driven analysis and machine learning-based crop and fertilizer recommendation system for revolutionizing farming practices," *Agriculture*, vol. 13, no. 11, p. 2141, 2023.

[14] A. Bhati, A. Pandey, A. Agrawal, A. Khandelwal, and N. Gupta, "Crop recommendation system," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 5, no. 5, May 2023.

[15] D. Dahiphale, P. Shinde, K. Patil, and V. Dahiphale, "Smart farming: Crop recommendation using machine learning with challenges and future ideas," in *Artificial Intelligence and Applications*, 2022.

[16] M. VAMSI and G. P. BABU, "Comprehensive crop advisory system for sustainable and resilient agriculture."

[17] H. Gunasekaran, D. Kanmani, and R. Krishnamoorthi, "Optimized ensemble learning for enhanced crop recommendations: Leveraging ml for smarter agricultural decision-making," *Engineering Proceedings*, vol. 82, no. 1, p. 95, 2024.

[18] H. Gunasekaran, K. Ramalakshmi, S. Debnath, and D. K. Swaminathan, "Physics-aware ensemble learning for superior crop recommendation in smart agriculture," *Sensors*, vol. 25, no. 19, p. 6243, 2025.

[19] S. Shastri, S. Kumar, V. Mansotra, and R. Salgotra, "Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence," *Scientific Reports*, vol. 15, no. 1, p. 25498, 2025.

[20] A. Cartolano, A. Cuzzocrea, and G. Pilato, "Analyzing and assessing explainable ai models for smart agriculture environments," *Multimedia Tools and Applications*, vol. 83, no. 12, pp. 37 225–37 246, 2024.

[21] N. Acharya, P. Khatiwada, R. Pandey, S. Niroula, and P. Chapagain, "Crop recommendation system using machine learning: a comparative study," *International Journal on Engineering Technology*, vol. 1, no. 2, pp. 302–311, 2024.

[22] A. Ingle, "Crop recommendation dataset," 2020.