International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(6s) 2026 pp. 124-143
https://doi.org/10.70917/ijcisim-2026-2874
Article

An Explainable Hybrid Ensemble Framework for Region-Adaptive Crop Recommendation Using Multi-Source Agricultural Data

Bhavana Gowda D. M.[1], K. S. Arvind[2], Nachappa M. N.[3], Madhumala R. B.[4], Sunil Kumar Y. C.[5]

[1] Department of Computer Science and Engineering, School of Engineering and Technology, JAIN (Deemed-to-be University), Bengaluru, Karnataka, India.
Email: dmbgowda@gmail.com
[2] Department of Computer Science and Engineering, School of Engineering and Technology, JAIN (Deemed-to-be University), Bengaluru, Karnataka, India.
Email: ks.arvind@jainuniversity.ac.in
[3] School of Computer Science and Information Technology, JAIN (Deemed-to-be University), Bengaluru, Karnataka, India.
Email: mn.nachappa@jainuniversity.ac.in
[4] Department of Computer Science and Engineering (ICB), Dayananda Sagar Academy of Technology and Management, Bengaluru, Karnataka, India.
Email: dr.madhumala-cse@dsatm.edu.in
[5] Department of Chemistry, Dayananda Sagar Academy of Technology and Management, Bengaluru, Karnataka, India.
Email: sunilkumaryc@gmail.com

Abstract: Choosing an appropriate crop for a set of soil and environmental conditions exists one of most important decisions affecting agricultural productivity and long-term sustainability. Despite much of advances in machine learning-based agricultural analytics, larger existing crop recommendation models function as black-box systems and often face struggle to adapt across numerous geographic locations. Larger predictive performance may not always translate into practical considerations which farmers and agricultural planners can readily make use. In order to solve those challenges, the review gives an explainable hybrid ensemble system for region-adaptive crop recommendation which merges Random Forest, XGBoost, and Deep Neural Network models within a Bayesian optimization scheme. Proposed system integrates information from different agricultural sources, consisting soil nutrient properties, historical cultivation records, weather observations and satellite-derived vegetation indicators, permitting diverse environmental parameters to be simultaneously considered. The data quality was improved via missing-value imputation, dimensionality reduction, feature normalization and feature relevance analysis before the model training. The proposed framework was evaluated making utilizing agricultural datasets gathered from India and Africa under a 10-fold cross-validation strategy. The ensemble model performed an accuracy of 93.1%, an RMSE of 5.46, and an R^2 value of 0.95, outperforming the individual learning models. Evaluation on geographically different datasets further consistently demonstrated predictive capability, accuracy of 91.3% during cross-regional validation. Explainability analysis based on SHAP revealed that rainfall and vegetation-related indicators exerted the strongest influence on crop suitability predictions. The outcome demonstrates that integrating ensemble learning with explainable artificial intelligence can reliably provide transparent crop recommendations, offering support practically for resource-efficient farming, precision agriculture and climate-resilient agricultural planning.

Keywords: Crop Recommendation, Explainable Artificial Intelligence, Precision Agriculture, Bayesian Optimization, Ensemble Learning, SHAP, Remote Sensing, Sustainable Agriculture.

1. INTRODUCTION

Agriculture is the foundation for production of food and rural livelihoods, while also contributing importantly to national economies and well-being of the societies. As the global population extends to grow and is approaching

10 billion within few decades, the demand for agricultural outcome is increasing at higher rate. Achieving this demand is becoming difficult as the farming systems are confronting a range of environmental and resource-related constraints. Irregular patterns of rainfall, temperatures rise, declining of soil fertility, shrinking of freshwater resources, and rapid land-use changes are disturbing agricultural productivity in larger regions of world. In such constraints, conventional farming practices alone are often insufficient for sustainable production support. There is a need for intelligent decision-support systems which can assist farmers and agricultural planners in getting informed choices regarding the selection of crops and resource management [8], [19].

Digital transformation of agriculture has led new options for addressing such challenges. Technologies like Internet of Things (IoT) sensors, platforms for remote sensing, geographic information systems, and automated monitoring tools are causing larger volumes of agricultural data on daily basis. Collected data from weather stations, satellite imagery, soil surveys and farm management records leads to valuable insights into conditions which influence growth and productivity of crop. Increasing accessibility of such data sources has encouraged the adoption of information-driven methods capable of supporting higher precise and efficient agricultural management [33], [38]. Instead of relying solely on experience-dependent decisions, farmers can now benefitted from analytical frameworks which leverage historical and real-time data to guide agricultural functions.

Out of the different analytical methods employed in modern agriculture, Machine Learning (ML) has gained some attention because of its capability to uncover complicated relationships within high and heterogeneous datasets. Agricultural frameworks are influenced by various interacting factors, maximum of which exhibit nonlinear behavior that is hard to represent utilizing conventional statistical mechanisms. ML algorithms like Random Forest(RF), Support Vector Machines (SVM), XGBoost, and Deep Neural Networks (DNN) have been successfully applied to jobs consisting crop suitability assessment, irrigation scheduling, yield forecasting and disease prediction [23], [40]. Ensemble learning approaches recently have attracted interest as they combine multiple predictive models and often deliver higher reliable performance than specific algorithms while dealing with numerous agricultural conditions [1], [28], [51].

Improvements in Earth observation methodologies have majorly extended scope of agricultural analytics. Satellite-derived vegetation indices, climatic records, and soil measurements provide complementary information for understanding crop–environment interactions [20], [36], [39]. Many of the studies have reported that models trained utilizing integrated datasets outperform consistently those grown from isolated data streams [8], [49]. As a outcome, fusion of multi-source agricultural data has emerged as a key in research of the precision agriculture, enabling the growth of much robust and context-aware support systems for the decision making.

Considerable part of the literature has focused on determining crop yield instead of assisting farmers in making decisions that crop must be cultivated under specific environmental conditions. Although yield forecasting gives valuable data about expected production, this does not directly support decision-making method that happens before planting. Farmers regularly require mentoring on crop choices based on prevailing soil and climatic constraints, resources available, and regional constraints. Consequently, there exists a requirement for recommendation-oriented systems which can convert predictive insights into practical cultivation strategies [15], [23].

One more concern relates to the limited geographic scope of the majority of existing studies. Agricultural constraints change substantially across places due to varations in soil composition, patterns of rainfall, temperature fluctuations, topography, and practices of farming. Frameworks built making use of information from a specific region generally perform better within the environment but face challenges as applied other places. Therefore, a model trained in one region may not perform with the same reliability when applied to another agro-climatic zone. This makes the development of region-adaptive recommendation frameworks an important research challenge [7], [35], [46].

A further concern is the limited transparency of advanced machine learning and deep learning models. Although these models often achieve high predictive performance, their internal decision-making process is not always easy to interpret. This becomes a serious issue in agriculture, where crop recommendations can influence farm investments, input use, and production outcomes. The growth in Explainable Artificial Intelligence (XAI) have sought to represent this issue by revealing how single variables contribute to predictions of the model and by providing higher insight into the underlying decision approach [13], [42]. Therefore, integrating explainability into crop recommendation systems is necessary to improve stakeholder confidence and support wider adoption.

A third significant challenge is the lack of integration among distinct agricultural information sources. Soil attributes, climatic variables, crop management data, and remote sensing predictors are typically studied separately. However, they are inherently interconnected under field conditions. Existing solutions that lack fusion across information sources may miss important relationships between crops and environment [2], [20], [49]. There are also limitations due to many solutions relying on single-model architectures that are vulnerable to noise, class imbalance, and regional shifts [21], [28].

Recent developments in explainable ensemble learning provide a useful direction for addressing these limitations. Hybrid ensemble frameworks can combine the strengths of different machine learning models and reduce the limitations of relying on a single algorithm. Also, XAI, such as SHAP, improve the transparency of these models by highlighting the contribution of specific features to the final consideration [13], [42]. This helps users understand how the model arrives at its decisions and builds greater confidence in the recommended outcomes. Furthermore, combining explainability with heterogeneous agricultural datasets and adaptive optimization methods supports the development of crop recommendation systems that are accurate, scalable, transparent and capable of solving the complexities of real-time farming environments.

Based on the above discussion, the following research gaps are identified:

G1: Existing crop recommendation systems do not fully integrate heterogeneous agricultural data, such as soil attributes, climatic variables, historical crop records, and remote sensing indicators, within a single analytical framework.

G2: The adaptability of crop recommendation models across different agro-climatic regions remains limited, as many models are developed and validated only within specific geographical contexts.

G3: Many existing approaches provide recommendations without adequate explainability, making it difficult to interpret the influence of individual features on the final crop suitability decision.

G4: Most current studies focus mainly on improving prediction accuracy, while comparatively less attention is given to generating actionable recommendations that can support practical agricultural decision-making.

The above gaps presents an Explainable Hybrid Ensemble Framework for Region-Adaptive Crop Recommendation Using Multi-Source Agricultural Data in response to these research gaps. To increase forecast reliability under numerous agricultural settings, system integrates RF, XGBoost, and DNN models into a Bayesian-optimized ensemble architecture. Crop suitability takes into consideration the relationship between several factors. Our framework incorporates information about soil features, weather conditions, historical cropping patterns, and satellite-based vegetation indices. SHAP-based explainability has been introduced to facilitate transparency for better decision-making by understanding the impact of each parameter on each recommendation and gaining insights into model prediction. Moreover, datasets from regions with varying agro-climatic conditions in India and Africa are employed to validate the generality and adaptability of the system. The contributions of this study can be highlighted:

i. To present an explainable hybrid ensemble system with RF, XGBoost, and DNN models fused with Bayesian optimization-based weights assignment.

ii. To propose a holistic approach to fuse heterogeneous sources of agriculture-related information such as soil properties, weather metrics, historical cropping patterns, and satellite imagery data for crop recommendation.

iii. To propose a region adaptive framework that offers dependable recommendations for different agro-climatic regions by enhancing generalization ability and prediction stability.

iv. SHAP-based explainability is incorporated to improve the transparency and interpretability of crop recommendations.

v. Cross-regional datasets acquired from different locations are used to evaluate the predictive ability, robustness, transferability, and practicality of the proposed framework.

2. LITERATURE REVIEW

Agricultural production is becoming increasingly vulnerable to changing climate, soil degradation and increasing demands on vital resources like water and nutrients. These challenges demand higher requirements for supporting decision systems in order to select crops according to ground truths about actual field and environmental conditions. Machine learning algorithms, ensemble models, multi-source agricultural data fusion and Explainable AI solutions have been broadly explored for crop prediction/recommendation tasks over the past few years.

Decision Trees, Support Vector Machine (SVM), Artificial Neural Networks (ANN), and Random Forest (RF) classifiers are supervised learning based ML models that have also been adopted in several studies to predict the set of feasible crops given certain environmental conditions. The reviewed studies indicate that ML models can learn meaningful patterns from agricultural data and often outperform conventional rule-based crop recommendation approaches. Moreover, researchers have shifted their focus from standalone prediction models to smart crop recommendation systems that intelligently combine multiple environmental and agronomic parameters.

Authors in [23] designed an ensemble-based recommendation system which resulted in better predictions with smaller confidence intervals across different climates. Authors from [15] employed more than one property of soil such as soil nutrients, soil pH and weather variables to create their recommendation system for crops. Their results suggest improved recommendation reliability when evaluating cumulative impact of multiple environmental variables rather than a standalone indicator. Soil attributes are important determinants of crop choice because they impact nutrient uptake, water availability, and root development in plants. The research conducted in [2] shows that results can be augmented by providing rich soil information to be considered during learning. Similarly, work from [35] created AI-based soil–crop suitability models specific to each region in India. They also mentioned that having region specific environmental features is beneficial for crop recommendation systems. Due to differences in climate, soil, farming techniques, and regional growth conditions, learned models may not generalize well to different regions that were not studied [7], [35], [46]. Much of the current literature also focuses on creating models that achieve higher predictive accuracy with very little emphasis on model interpretability.

Agricultural Decision Support using Ensemble Learning

A number of existing works have attempted to employ ensemble learning for agricultural tasks. For instance, the authors in [1] demonstrated that ensemble-driven methods surpassed the baseline performances for crop recommendation and disease detection. Concurrently, the study conducted in [51] proposed a hybrid RF–XGBoost ensemble architecture for crop recommendation and yield prediction. Their results illustrated an enhanced accuracy and stability due to effective learning of nonlinear relationships among cultivable parameters. Consequently, more recent studies have strived to create hybrid ensemble designs that incorporate diverse learning capabilities. For instance, combining RF and DNN models with an uncertainty-aware learning paradigm in [28] led to enhanced prediction dependability despite the inclusion of noisy/incomplete data records. Moreover, hybridizing RF–XGBoost ensembles and training them to predict cotton yield showed good generalization ability over multiple farming sites [21].

Stacking- and weighted-based ensemble models have also been used to enhance prediction performance. For example, [26] designed a stacking-framework that leverages information from both soil profiles and weather histories to assess crop suitability. Concurrently, [5] integrated multiple environmental attributes within an ensemble-based model to better generalize over multiple environments.

Current ensemble learning approaches tend to determine ensemble weights in an ad-hoc or trial-and-error fashion rather than through explicit optimization. As a result, these models tend to lack robustness and do not generalize well to multiple agricultural environments [5], [21], [28]. This necessitates the development of AI systems that can be explained, employing ensemble weights that adjust on the fly to ensure both precise predictions and straightforward guidance for decision-making.

Fusion of Agricultural Data from Multiple Sources

Sensorial data collected from soil sensors, weather stations, satellite imagery, crop management, and IoT-enabled agricultural equipment, produce information about different parameters that affect crop growth and production. Remote sensing has emerged as a crucial aspect of agricultural analytics due to its ability to monitor crop conditions continuously across vast spatial expanses. Vegetation indexes like NDVI and SAVI have been effectively employed in ML models for crop yield prediction [36]. [20]. Authors in [36] reported that employing vegetation indices with ML models increased the prediction accuracy. Authors in [20] mentioned that fusion of NDVI and climatic parameters using DL models decreased uncertainty along with improving accuracy.

Authors in [39] combined Sentinel-2 imagery with soil information to predict corn yield. Similarly, a framework was proposed in [49] that integrated remote sensing information, Crop Growth Models (CGMs), and ML for precision nitrogen management in maize. Authors in [13] presented that fusion of vegetation indexes, soil properties, and weather parameters helped them to attain higher prediction accuracies while lowering uncertainty. However, several previous studies fuse limited sources of data or make use of independent sources of data for their analysis. Independent or limited sources are unable to capture the diversity found in practical agricultural systems [8],[13], [49]. This highlights the need for data fusion methods that can seamlessly merge various agricultural data streams, paving the way for crop recommendation systems that are both precise, dependable, and capable of handling growth.

Explainable Artificial Intelligence (XAI) for Agriculture Decision Support Systems

Deep learning (DL) and machine learning (ML) models tend to perform exceptionally well in most agricultural prediction tasks. However, as these models become more complex in architecture, it becomes challenging to understand their decision-making process. Increasing the transparency and interpretability of these models has been a growing concern. Explainable Artificial Intelligence (XAI) has emerged as a potential solution to bridge this gap.

More studies have emphasized the relevance of explainability in agricultural applications. For example, the study in [42] highlighted that explainability enables agricultural stakeholders to better understand the reasoning behind complex predictive models, thereby nondecreasing confidence in automated decision-support systems. IN the same way, the work presented in [13] demonstrated that explainable frameworks improve the adoption of ML techniques by providing higher transparency into the process via which recommendations are generated.

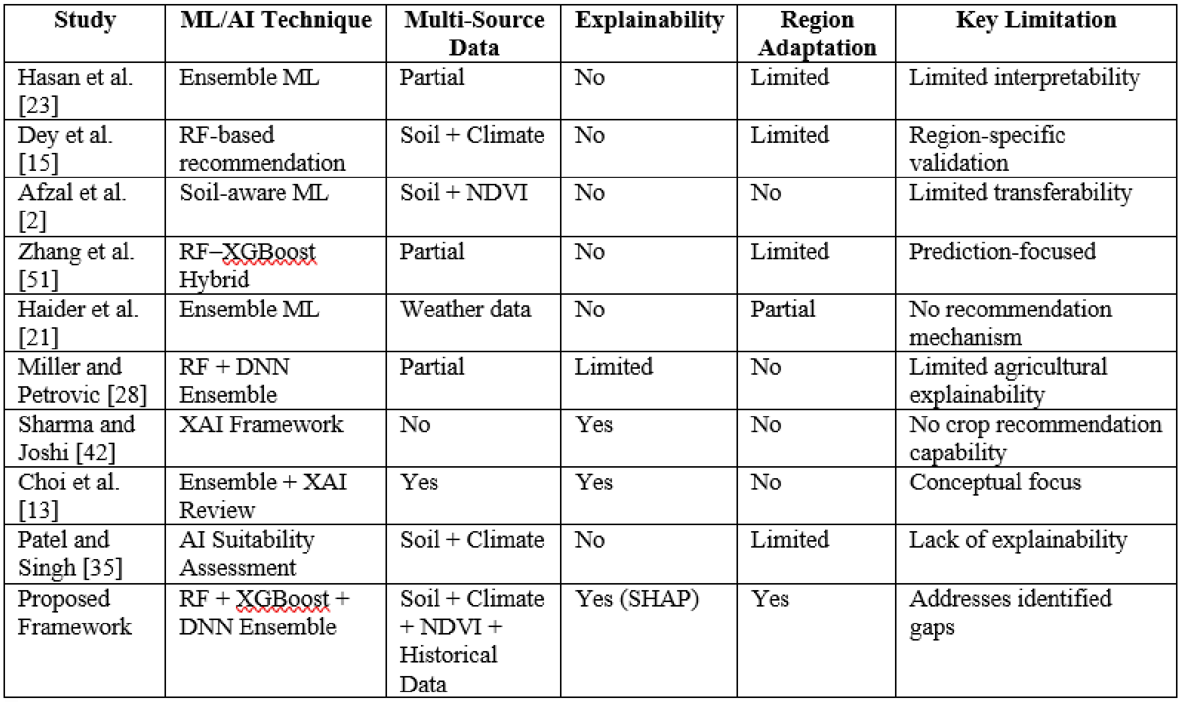
Out of the various XAI mechanisms, SHAP and LIME methods quantify the contribution of individual input features and give valuable insights into how environmental and agronomic factors influence prediction yields. Their effectiveness has been demonstrated across numerous agricultural applications, including crop recommendation, output prediction, disease detection, and resource management [13], [42].

XAI is becoming more and more popular, although its application in crop recommendation systems is still very restricted. Numerous studies that are now available mainly concentrate on predictive performance and provide minimal details regarding the rationale behind the recommendations. Lack of interpretable outputs can undermine user trust and impede real-world implementation. Crop recommendation systems that combine excellent predictive capabilities with transparent and intelligible decision-making processes are therefore clearly needed.

Analysis of Research Gaps

The application of artificial intelligence and machine learning to agricultural decision support has advanced significantly, according to a study of the literature. Nonetheless, a number of persistent issues that continue to limit the practical utility of crop recommendation systems are shown by the corpus of extant research. Table 1 compares representative contributions in terms of data integration, explainability, geographic adaptation, and recommendation capability in order to better place the current work within the current research environment.

Table 1. Comparative Analysis of Existing Studies and Research Gaps



Study	ML/AI Technique	Multi-Source Data	Explainability	Region Adaptation	Key Limitation
Hasan et al. [23]	Ensemble ML	Partial	No	Limited	Limited interpretability
Dey et al. [15]	RF-based recommendation	Soil + Climate	No	Limited	Region-specific validation
Afzal et al. [2]	Soil-aware ML	Soil + NDVI	No	No	Limited transferability
Zhang et al. [51]	RF–XGBoost Hybrid	Partial	No	Limited	Prediction-focused
Haider et al. [21]	Ensemble ML	Weather data	No	Partial	No recommendation mechanism
Miller and Petrovic [28]	RF + DNN Ensemble	Partial	Limited	No	Limited agricultural explainability
Sharma and Joshi [42]	XAI Framework	No	Yes	No	No crop recommendation capability
Choi et al. [13]	Ensemble + XAI Review	Yes	Yes	No	Conceptual focus
Patel and Singh [35]	AI Suitability Assessment	Soil + Climate	No	Limited	Lack of explainability
Proposed Framework	RF + XGBoost + DNN Ensemble	Soil + Climate + NDVI + Historical Data	Yes (SHAP)	Yes	Addresses identified gaps

Four recurrent constraints in the body of current literature are identified by the analysis:

G1: Inadequate integration of diverse agricultural datasets into a single recommendation system.

G2: Cross-regional transferability and region-adaptive learning are not given enough consideration.

G3: The absence of transparent and user-trusting explainable recommendation systems.

G4: Predictive performance is prioritized over useful recommendation generation.

To address the above limitations, the suggested framework employs mixed agricultural data including soil characteristics, weather elements, previous crop documentation and remotely sensed vegetation indices in a single crop recommendation system. The combination of agricultural information from multiple sources along with hybrid ensemble learning, XAI, and location-specific customization gives our suggested method a holistic and dependable structure for intelligent crop recommendation.

3. METHODOLOGY

The designed explainable crop recommendation system considers soil properties, weather data, remote sensing variables and historical cropping data to provide intelligent crop recommendations that can adapt to different regions. With machine learning, hybrid ensemble learning, Bayesian optimization, and XAI, we aim to enhance prediction accuracy and simultaneously ensure the interpretability, trustworthiness, and generalizability of our framework across diverse farming environments.

Figure 1. shows the overview of the proposed explainable crop recommendation framework. The proposed framework mainly includes seven modules: multi-source data acquisition, data preprocessing, feature engineering, hybrid ensemble learning, Bayesian optimization, explainability analysis, and crop recommendation as shown in Figure 1. Each module plays a critical role in processing heterogeneous agricultural data and translating them into precise, interpretable, and user-friendly crop recommendations for facilitating decision-making in precision agriculture.

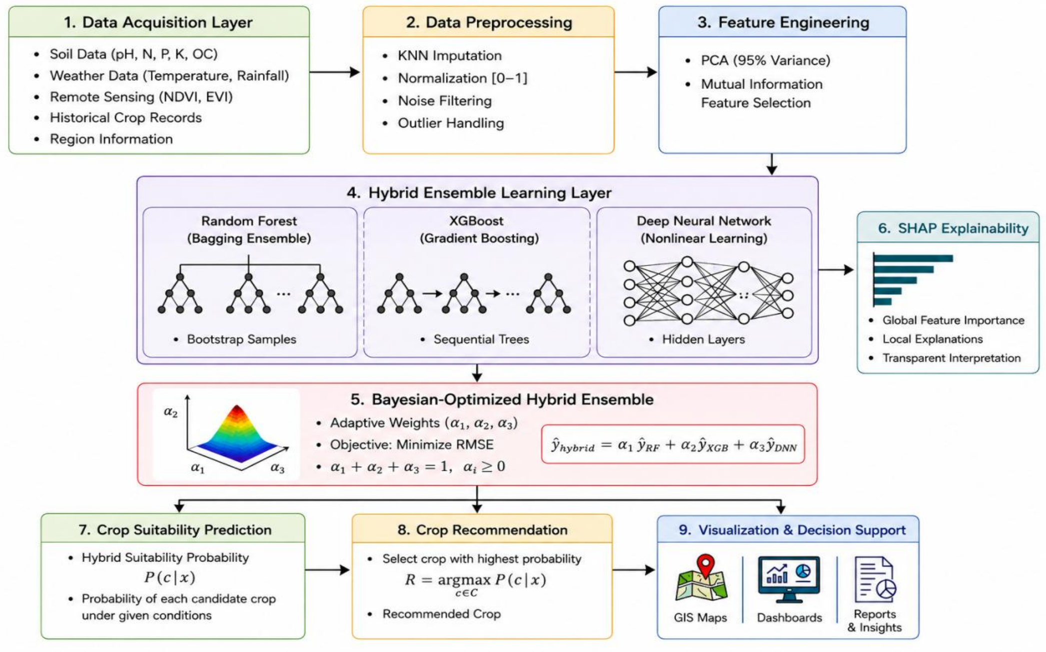


Figure 1. Architecture of the Proposed Explainable Hybrid Ensemble Framework for Region-Adaptive Crop Recommendation.

The workflow of the presented crop recommendation framework is illustrated in Figure 1. First, agricultural datasets from various sources are acquired. Second, considering the datasets might be in diverse formats and missing values, noise and outliers exist, the data preprocessing technique is applied to clean the raw data. Then the feature engineering method is employed to select informative variables by using mutual information analysis and then applying PCA to extract the most important principal components that can describe the data with minimum loss of information for crop recommendation. Next, three models including RF, XGBoost, and DNN are trained with the engineered features. Third, RF, XGBoost, and DNN models make crop predictions separately, then we employed an ensemble strategy optimized by Bayesian to fuse the three predictions with appropriate weights. Lastly, we conduct SHAP analysis to evaluate each feature's contribution and interpret each prediction and the whole model. Based on the crop suitability scorecard, the framework selects the crop that fits best with the provided conditions. Finally, the recommended crop and related information, including feature importance and explanation, are presented to the user with the help of visualization and decision-support modules.

Data Acquisition and Integration: Crop growth and yields are influenced by numerous interrelated factors, including soil characteristics, weather conditions, vegetation health, and historical land use. To comprehensively characterize these factors, we collected agricultural data from various sources, including soil databases, meteorological stations, remote sensing datasets, and historical cropping records. Integrating heterogeneous data sources allows a comprehensive overview of agricultural conditions and helps achieve robust crop suitability predictions.

Table 2 summarizes numerous categories of data considered in this work.

Table 2. Data Sources Used in the Proposed Framework

Data Type	Source	Attributes
Soil Data	ICRISAT / ICAR	pH, Nitrogen, Phosphorus, Potassium, Organic Carbon
Weather Data	IMD / NOAA	Temperature, Rainfall, Humidity
Remote Sensing Data	MODIS / Sentinel-2	NDVI, EVI
Historical Crop Data	FAO / Agricultural Departments	Crop Type, Yield, Region

The integrated dataset is mathematically represented as follows:

$$D = \{(x_{\mathrm{i}}, y_{\mathrm{i}}) \mid x_{\mathrm{i}} \in \mathbb{R}^{\mathrm{n}}, y_{\mathrm{i}} \in C\}$$

where C is set of candidate crops, yi is corresponding crop class, and xi specifies or denotes the feature vector linked to ith agricultural observation.

Each feature vector is defined as follows

$$\mathrm{xi} = [\mathrm{Si}, \mathrm{Ci}, \mathrm{Vi}, \mathrm{Hi}]$$

where Si, Ci, Vi, and Hi specifies soil characteristics, climate variables, vegetation indices, and historical agricultural data. Such structure makes it possible to process numerous agricultural data types within a single recommendation framework.

Data pre-processing: As agricultural datasets may include missing values, vague measurements, parameters registered at different numerical levels etc., preprocessing pipeline was employed prior to model building.

KNN algorithm was used to fill the missing values. It substitutes values according to nearby samples with similar behavior.

The imputed value is represented as:

$$x'_j = \frac{1}{k} \sum_{i=1} x_i$$

where is the predicted value and k is number of closest nearby samples. This mechanism minimizes information loss while preserving local data patterns.

Min–Max normalization was employed during model training to guarantee that every variable contributed equally:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

where *and* represent the minimum and maximum values of the respective feature. By scaling all features to a common range, this transformation improves numerical stability during model training.

Feature engineering was utilized to upgrade learning performance and decrease dataset redundancy. In order to solve multicollinearity among input variables and minimize dimensionality, Principal Component Analysis (PCA) was initially used. The expression for the transformation is:

$$\mathrm{Z} = \mathrm{XW}$$

where Z is the transformed feature space, W is the eigenvector matrix, and X is the normalized feature matrix. For additional analysis, components that accounted for at least 95% of the cumulative variance were kept.

The significance of individual variables to crop classes was assessed using Mutual Information (MI) analysis following dimensionality reduction. MI is computed as

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log\left(\frac{p(x,y)}{p(x)p(y)}\right)$$

where p(x) and (p(y) are the marginal probability distributions and p(x,y) is the joint probability distribution. Larger MI features were kept for model training since they were thought to be much informative.

Hybrid Ensemble Learning Framework: RF, XGBoost, and DNN are combined as they learn differently and can extract various patterns in the data. Moreover, these approaches effectively capture nonlinear relationships common in agricultural settings.

Random Forest: Random Forest builds several decision trees on bootstrapped datasets and randomly selects features. The combined forecast is calculated as follows:

$$\hat{y}_{RF} = \frac{1}{n} \sum_{i=1}^{n} T_i(x)$$

where is the forecast produced by the ith decision tree.

Through sequential learning, in which subsequent trees concentrate on fixing earlier prediction errors, XGBoost enhances predictive performance. The expression for its objective function is:

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

where Ω is the regularization term that controls model complexity and *l* is the loss function.

Highly nonlinear interactions between environmental variables are captured by the DNN component. Its result is shown as:

$$\hat{y}_{DNN} = \sigma(W_{3\emptyset}\,(W_{2\emptyset}(W_{1x} + b_1) + b_2) + b_3)$$

where σ stands for the ReLU activation function, σ for the output activation function, and Wi and bi stand for trainable weights and biases.

Hybrid Ensemble with Bayesian Optimization

A weighted ensemble approach was used to aggregate the outputs of the RF, XGBoost, and DNN models in order to enhance recommendation performance. To find each model's most appropriate contribution, Bayesian Optimization was used in place of fixed or manually chosen weights.

Different aspects of agricultural data are captured by the three learning algorithms. DNN models intricate connections between soil, climate, and vegetation variables, XGBoost enhances prediction performance by gradient boosting, and Random Forest is good at managing noisy observations and nonlinear correlations. By combining these models, the framework can lessen the constraints of individual learners while gaining from their complementing strengths.

The predictions generated by the three base models are represented as , , and respectively. These predictions are then combined to compute the final ensemble output as follows:

$$\hat{y}_{Hybrid} = \alpha_1 \hat{y}_{RF} + \alpha_2 \hat{y}_{XGB} + \alpha_3 \hat{y}_{DNN}$$

where , , and specify the weights assigned to the respective models. To ensure a valid weighted combination, the ensemble weights are constrained as follows:

$$\alpha_1 + \alpha_2 + \alpha_3 = 1$$

$$\alpha_i \geq 0 , \qquad i = 1,2,3$$

By reducing the prediction error, Bayesian Optimization yields the ideal weight vector:

$$\theta^* = \arg\min_{\theta} RMSE(\theta)$$

Where $\theta = \{\alpha_1 , \alpha_2 , \alpha_3\}$

and RMSE(θ) specifies the Root Mean Square Error related to a specific weight configuration.

Bayesian optimization works by building a probability model of the objective function and using it to select the most promising hyperparameters to evaluate in the true objective function. It often outperforms traditional search methods such as grid search and random search with fewer evaluations, and lends itself to efficient hyperparameter optimization, even for computationally expensive machine learning models.

Optimization is repeated until convergence or search budget has been reached. Maximum weight calculated from Bayesian optimization is then used to create final ensemble prediction. Adaptively combining the predictions from RF, XGBoost, and DNN models enables our ensemble to yield more accurate, robust, and stable predictions, allowing for more dependable estimates of crop suitability.

Figure 2 shows the workflow of Bayesian-optimized ensemble framework. In our proposed framework, the predictions made by each of the three base learners are combined through the process of adaptive weight optimization to create a final crop suitability prediction..

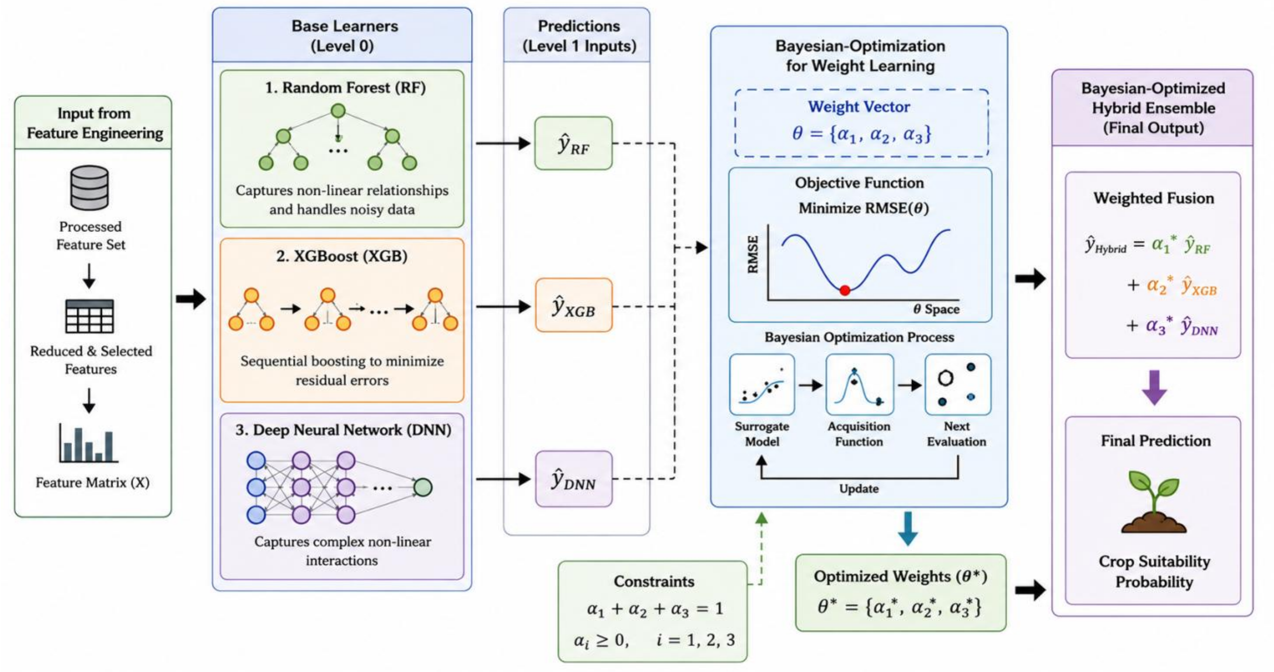


Figure 2. Bayesian-Optimized Hybrid Ensemble Learning Framework.

Figure 2 depicts the hybrid ensemble model architecture that we propose. The predictions of three complementary learners are used: RF, XGBoost and DNN. Each learner is fed with the features obtained after the preprocessing step, and each generates its own independent prediction of the suitability score. We combine the predictions of each individual learner by weighted aggregating them. Bayesian optimization is performed in order to find the optimal weights that yield the lowest error in prediction. This ensemble prediction is used as an input for the final prediction of crop suitability, which is then recommended to the user.

Explainability Analysis Using SHAP

Ensemble models can provide several advantages to crop recommendation due to the nature of combining learning algorithms leading to better quality predictions. However, for such decisionsupport tools in agriculture, recommendations made by models are helpful only if users are able to comprehend why the recommendation was made. As crop selection needs to be based on pragmatic decisions taking place in-field, simply knowing what crop to grow is not enough - the model should also provide reasons why said crop is preferred. Thus, SHAP-based explainability was adopted for the proposed framework. SHAP values estimate the impact of every input variable on the prediction and illustrate how features related to soil, climate, and vegetation play a role in computing the final suitability score. Interpretation is provided both at the model-level as well as for individual recommendations. This ensures transparency in decision process and allows users to build trust with the crop recommendation system.

For feature i, the SHAP value is calculated as follows:

$$\emptyset_i = \sum_{S \subseteq N\{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} \, [f(S \cup \{i\} - f(S)]$$

where ϕi is the contribution of feature i, N is the whole feature set, S is a subset of features that does not include i, and f(·) is the ensemble model's prediction function.

SHAP analysis was used to determine which variables most influence predictions of crop suitability. Beyond providing global feature importance, the framework also generates instance-level explanations showing how each feature affects every crop recommendation. This allows users to see how the model came to its specific prediction, allowing for greater transparency and more confident farming decisions.

Generation of Crop Recommendations

The system creates crop recommendations based on the estimated suitability probability after collecting the ensemble predictions and related explanations.

Let the collection of potential crops be represented by C={c1,c2,…,ck}. The hybrid ensemble model calculates each crop's suitability probability for a given agricultural observation x:

$$P(c|x) = f_{Hybrid}(x)$$

The suggested crop is identified as:

$$R_i = \arg\max_{c \in C} P(c|x)$$

where Ri is suggested crop for the ith observation. The final suggestion is made for the crop with the highest probability of suitability.

Algorithm 1: Framework for Explainable Hybrid Crop Recommendations

Input includes historical crop records, soil data, climate variables, and remote sensing indications.

Output: Suggested crops and explanations of their significance

Step 1: Gather and combine agricultural data from various sources.

Step 2: Use KNN-based imputation to deal with missing values.

Step 3: Use Min–Max scaling to normalize feature values.

Step 4: Use PCA to reduce dimensionality.

Step 5: Select features based on mutual information.

Step 6: Train the Random Forest model in step six.

Step 7: Train the XGBoost model in step seven.

Step 8: Train the Deep Neural Network model in step eight.

Step 9: Use the three basis models to generate predictions.

Step 10: Use Bayesian Optimization to optimize the ensemble weights (α1,α2,α3).

Step 11: To get the ensemble forecast, combine the weighted model outputs.

Step 12: Calculate SHAP values to interpret the model.

Step 13: Calculate the likelihood that each potential crop will be suitable.

Step 14: Opt crop that has best chance of being suitable.

Step 15: Return back suggested crop along with any illustrations.

Proposed system combines multi-source agricultural data with feature engineering, hybrid ensemble learning, Bayesian optimization, and XAI to support accurate and reliable crop recommendations. Its performance is evaluated via comprehensive experimental analysis and compared with existing machine learning approaches to assess its predictive accuracy, robustness, and overall effectiveness.

4. EXPERIMENTATION AND RESULTS ANALYSIS

***Dataset Description*:** Datasets containing information about soil parameters, climatic attributes, vegetation indices, and crop history were used for validating the framework. They were selected from diverse regions so that the model could be validated across multiple agro-climatic and environmental conditions. Furthermore, this approach ensured that crop suitability analysis could draw upon diverse observations when formulating region-specific recommendations.

Datasets include information collected from Food and Agriculture Organization (FAO), NASA EarthData, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Indian Meteorological Department

(IMD) and local state agencies. Each repository maintains a different category of agriculture-based information. Integrating them helped include agronomic, climatic and environmental details within the scope of a single framework. This enhanced the trustworthiness of the recommendation while helping with assessing adaptability across regions.

Table 3: Overview of the Dataset and Features

Dataset	Source	Parameters	Temporal Coverage	Region	Format
ICRISAT Soil Dataset	International Crops Research Institute for Semi-Arid Tropics	pH, Nitrogen (N), Phosphorus (P), Potassium (K), Organic Carbon (OC)	2010–2024	India (Maharashtra, Andhra Pradesh)	CSV / Shapefile
IMD Climate Data	Indian Meteorological Department	Avg. Temperature (°C), Humidity (%), Rainfall (mm/month), Solar Radiation	2010–2024	South & Central India	CSV / NetCDF
MODIS NDVI (MCD12Q2)	NASA EarthData	NDVI, EVI, LAI (Leaf Area Index)	2010–2024	500m Spatial Resolution	GeoTIFF
FAO Agroclimatic Database	FAO Global Agro-Ecological Zones	Crop yield, cultivation type, seasonal suitability	2015–2024	Global	CSV
Local Agricultural Office Records	Regional Agriculture Departments	Crop type, yield records, market data	2010–2024	Regional	Excel

The datasets in table 3 include data on vegetation status, climate, soil fertility, and past agriculture practices gathered over several years. Variations of these data sources facilitates generation and assessment of crop recommendation models in a range of environmental circumstances, as Table 4 contains a list of the study's important parameters.

Table 4. Key Features Used in the Proposed Framework

Feature	Type	Unit	Description
Soil pH	Continuous	–	Acidity/alkalinity measure of soil
Nitrogen (N)	Continuous	kg/ha	Nutrient essential for vegetative growth
Phosphorus (P)	Continuous	kg/ha	Root development nutrient
Potassium (K)	Continuous	kg/ha	Improves crop resistance
Organic Carbon	Continuous	%	Soil fertility indicator
Rainfall	Continuous	mm	Seasonal rainfall
Temperature	Continuous	°C	Average daily air temperature
NDVI	Continuous	–	Vegetation vigor index
EVI	Continuous	–	Enhanced Vegetation Index
Relative Humidity	Continuous	%	Moisture availability indicator

The chosen characteristics are significant determinants of crop suitability and growth. Soil fertility and nutrient availability can be inferred from soil characteristics including pH, nitrogen, phosphorus, potassium, and organic carbon. The environmental circumstances encountered during crop growth are described by weather-related variables, such as temperature, humidity, and rainfall. In order to use satellite data to record crop conditions and vegetation health, NDVI and EVI were included.

All studies were conducted in a cloud-based computing environment that has an NVIDIA RTX A6000 GPU with 48 GB VRAM, 128 GB DDR4 RAM, and an Intel Xeon 32-core processor running at 2.9 GHz. The models were run on Ubuntu 22.04 LTS and implemented in Python 3.10. For model creation, optimization, and explainability analysis, libraries such as Scikit-learn, XGBoost, TensorFlow, SHAP, and Optuna were utilized.

Data Partitioning and Validation Strategy: Stratified sampling was used to separate the dataset into training, validation, and testing subsets in order to assess model performance under unknown circumstances. Seventy percent of the data was used for training, fifteen percent for validation, and fifteen percent for testing. Also, a 10-fold stratified cross-validation approach was used to evaluate the generalization performance and stability of the model. To guarantee that each sample was assessed precisely once, nine folds were utilized for training and one-fold for testing in each cycle.

The trials were repeated with different random seeds to lessen the impact of random variation, and the average outcomes were shown.

Hyperparameter Optimization: The choice of hyperparameters has a significant impact on model performance. Thus, appropriate parameter settings for the RF, XGBoost, and DNN models were found using Bayesian Optimization implemented through the Optuna framework.

The RMSE was to be minimized as the optimization goal:

$$\theta^* = \arg\min_{\theta} RMSE(\theta)$$

where θ is the set of model hyperparameters. Bayesian Optimization decreases processing cost and explores the search space more effectively than traditional search techniques. All trials were then conducted using the optimized parameter values.

Metrics for Evaluation

Regression-based and classification-based metrics were employed to evaluate the performance of the suggested model. MAE, RMSE, and R^2 are regression-based metrics that were used to evaluate prediction error and fit. Accuracy and F1- score were classification-based metrics that evaluate how correct and balanced the predictions for crop were. This combined evaluation made it possible to analyse the framework from both prediction and recommendation perspectives.

$$\boldsymbol{MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |}$$

◢ $$\mathbf{RMSE} = \sqrt{\frac{\mathbf{i}}{\mathbf{n}}\sum_{\mathbf{i=1}}^{\mathbf{n}}(\mathbf{y_i} - \hat{\mathbf{y}}_\mathbf{i})^2}$$

$$\boldsymbol{R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2}}$$

$$\boldsymbol{F_1 = \frac{2\ X\ Precision\ X\ Recall}{Precision + Recall}}$$

Accuracy was utilized to assess how accurate crop recommendations were across various datasets and geographical areas.

Analysis of the Rainfall-Yield Correlation

To determine how climatic conditions, affect agricultural output, the link between rainfall and crop yield was investigated. Crop yield often rises with increasing rainfall, as seen in Figure 3, suggesting a positive correlation between productivity and water availability. A high association between the two variables is seen by the fitted regression line and the Pearson correlation coefficient (r=0.82). Furthermore, the coefficient of determination (R2=0.67) indicates that rainfall patterns account for a significant amount of yield variance. These findings validate the inclusion of rainfall in the crop recommendation model and highlight its significance as a critical component of crop growth.

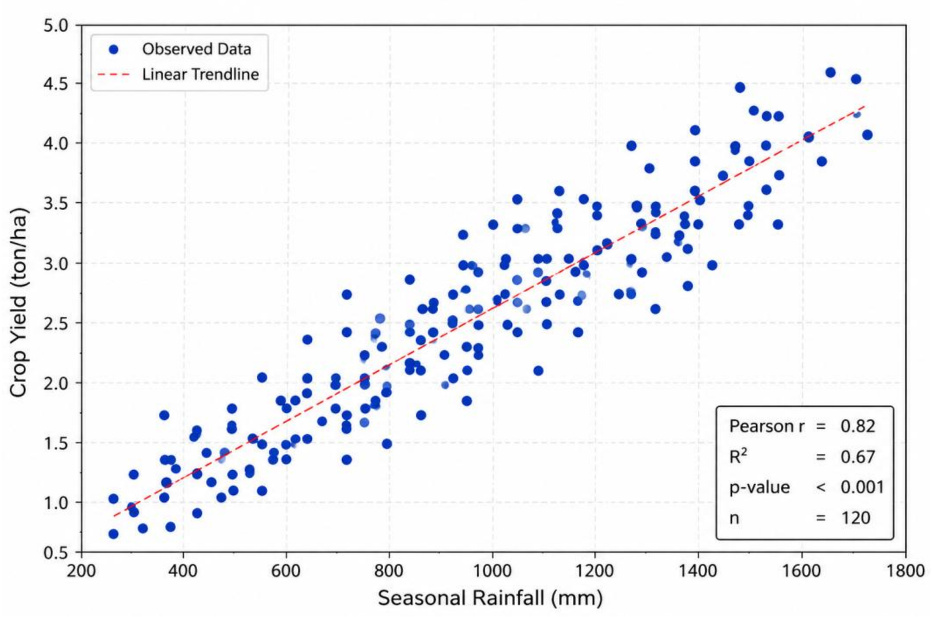


Figure 3. Correlation between Rainfall and Crop Yield

Principal Component Analysis (PCA)

The normalized dataset was subjected to PCA in order to find a compact representation of the input variables and investigate feature redundancy. Figure 4 displays the variance that the primary components account for.

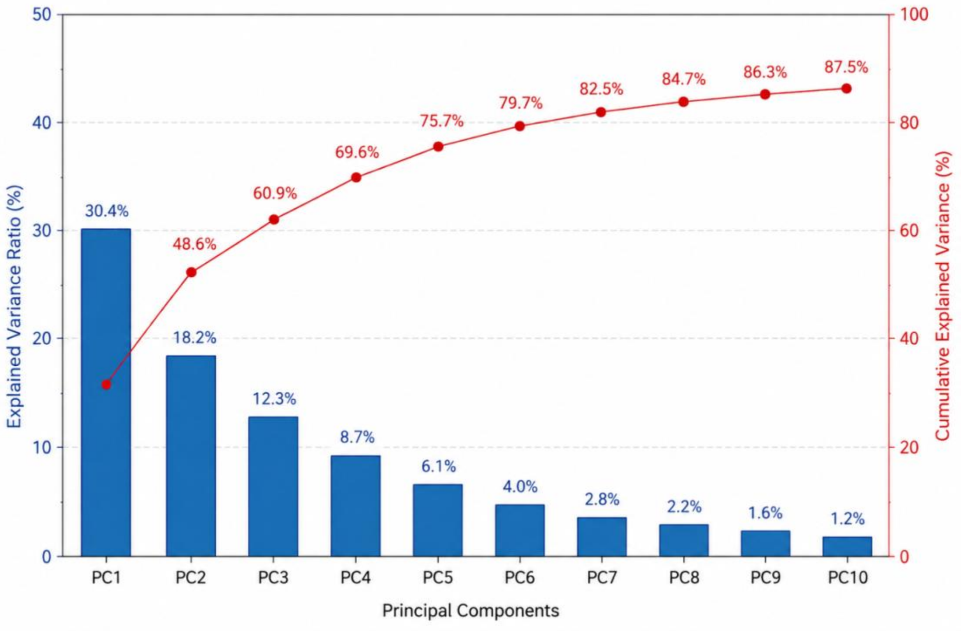


Figure 4. PCA Variance Explained by Principal Components

The cumulative variance represented by consecutive principal components is shown in Figure 4. The first few components preserve the majority of the dataset's information, whereas the contributions of subsequent components progressively diminish. This supports the use of PCA for dimensionality reduction and shows that redundancy exists in the original feature space. Computational complexity can be decreased without substantial information loss by keeping the components that account for most of the variance.

Analysis of Comparative Performance

A number of popular machine learning models, such as Decision Tree, RF, XGBoost, and DNN, were utilized to assess the suggested framework.

Table 5: Model Performance Comparison on Crop Suitability Prediction

Model	RMSE ↓	MAE ↓	R^2 ↑	F1-Score ↑	Accuracy (%) ↑
Decision Tree	9.87	7.65	0.81	0.76	82.3
Random Forest (RF)	7.35	5.21	0.89	0.83	87.5
XGBoost (XGB)	6.92	4.97	0.91	0.85	88.2
Deep Neural Network (DNN)	6.78	4.80	0.92	0.87	89.4
Proposed Hybrid Model (RF + XGB + DNN)	**5.46**	**3.92**	**0.95**	**0.92**	**93.1**

The hybrid ensemble model performed the best overall across all evaluation metrics, according to the data. RMSE and MAE decreased to 5.46 and 3.92, respectively, while an accuracy of 93.1%, an F1-score of 0.92, and a R^2 value of 0.95 were attained. These findings imply that using Bayesian optimization to combine RF, XGBoost, and DNN models yields more accurate predictions than using each learning approach alone.

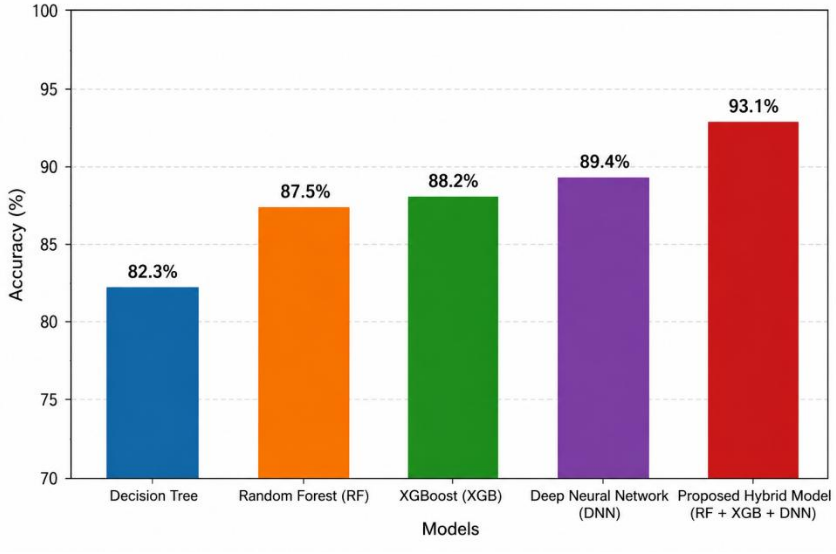


Figure 5. Comparative Accuracy of Evaluated Models

Figure 5. represents the Classification Accuracy obtained by the evaluated models. The result shows that there is a gradual increase in classification accuracy from traditional Decision Tree algorithm to complex machine learning models. Among the individual models, DNN produced the highest accuracy of 89.4%, followed by XGBoost with 88.2% and RF with 87.5%. The hybrid ensemble model achieved the best overall accuracy of 93.1%. This signifies that integrating complementary learning models improves crop recommendation better than individual models.

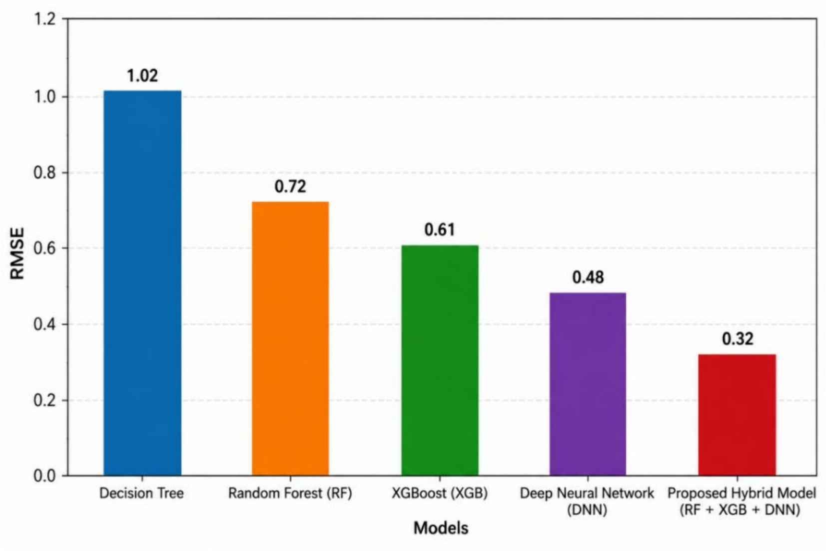


Figure 6. RMSE Comparison Across Models

The RMSE values produced by the various models are displayed in Figure 6. Smaller prediction errors and better agreement with actual crop suitability results are shown by lower RMSE values. DNN had the lowest RMSE of all the standalone models, followed by Random Forest and XGBoost. By achieving the lowest RMSE, the suggested hybrid ensemble demonstrated enhanced prediction accuracy and decreased estimate error.

Validation Across Datasets

Experiments were carried out utilizing datasets from Africa and India to assess the model's performance in various agricultural settings. Both locations have a comparable performance trend, according to the results shown in Figure 7. The models retained comparable performance on the Africa dataset, demonstrating good transferability, although achieving somewhat greater accuracy on the India dataset. With 93.8% accuracy for India and 90.2% for Africa, the suggested hybrid ensemble produced the greatest results, proving its capacity to generalize across geographically different settings.

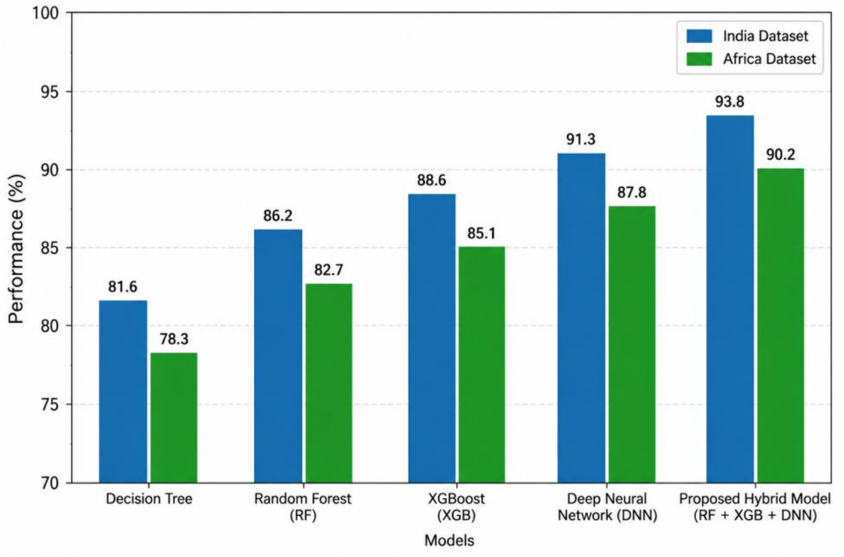


Figure 7. Cross-Dataset Performance Comparison

Explainability Analysis of SHAP

Understanding the variables impacting crop recommendations depends on the interpretability of the model. In order to assess feature contributions, SHAP analysis was carried out.

Figure 8 shows SHAP-based feature importance plot of the proposed model. It can be observed that out of the given four features rainfall has highest feature importance towards predicting the suitability of crops whereas temperature, soil ph and nitrogen follows. This observation clearly indicates that the proposed model considers the climatic conditions and soil fertility indicators while recommending crops. Also, the ranking makes intuitive sense as precipitation and temperature have direct impact on crop cultivation. Further, soil ph and nitrogen determines the availability of nutrients to plants and their growth. Therefore, it appears the model has successfully identified intuitive associations present in the data.

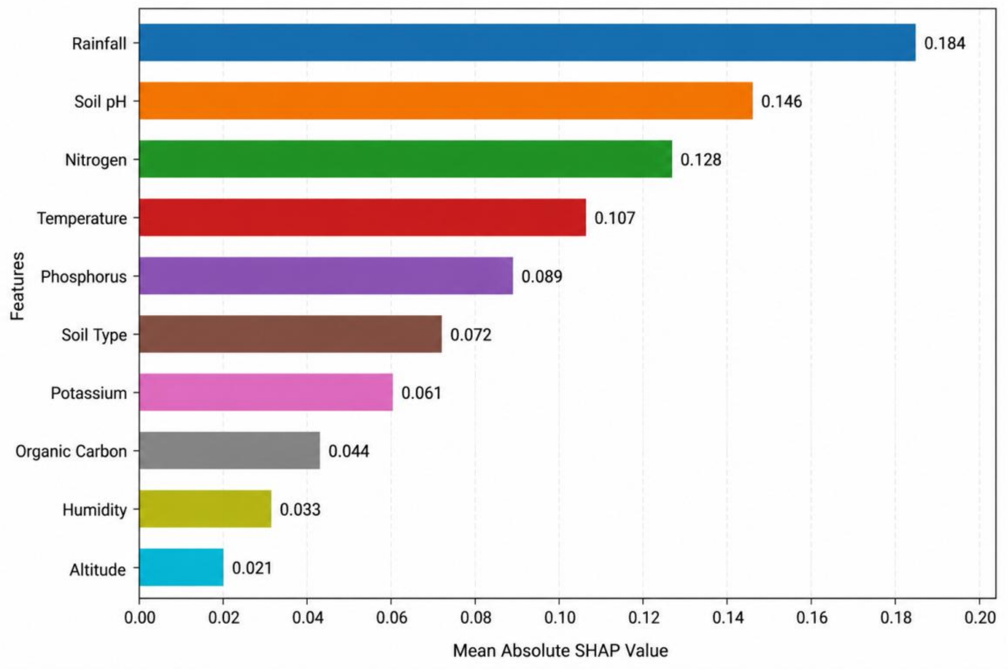
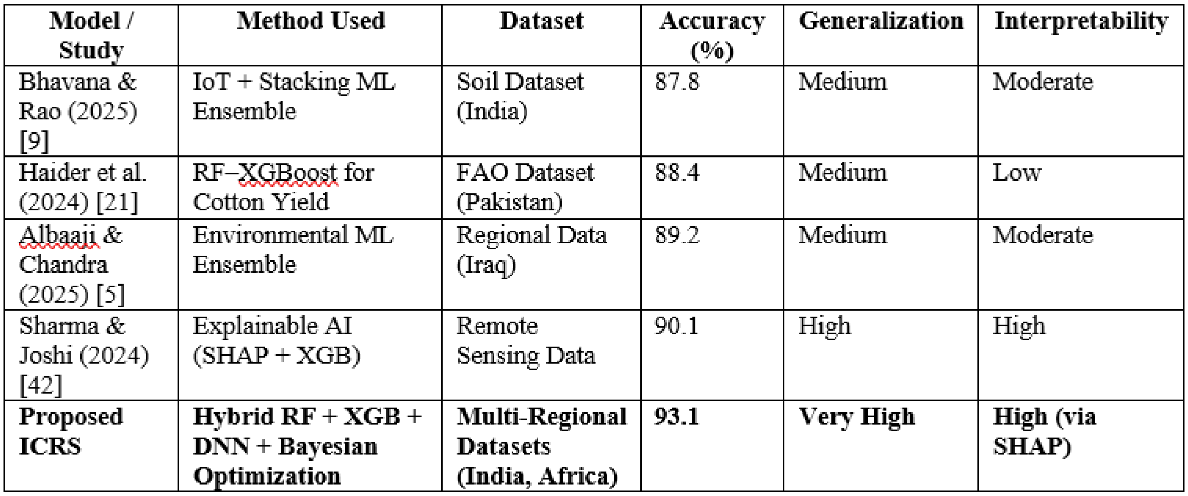


Figure 8. SHAP-Based Feature Importance Analysis

Comparing with Cutting-Edge Systems

To evaluate the effectiveness of the proposed method, its performance was compared with recently developed agricultural recommendation and prediction systems.

Table 6. Comparison with State-of-the-Art Systems



Model / Study	Method Used	Dataset	Accuracy (%)	Generalization	Interpretability
Bhavana & Rao (2025) [9]	IoT + Stacking ML Ensemble	Soil Dataset (India)	87.8	Medium	Moderate
Haider et al. (2024) [21]	RF–XGBoost for Cotton Yield	FAO Dataset (Pakistan)	88.4	Medium	Low
Albaaji & Chandra (2025) [5]	Environmental ML Ensemble	Regional Data (Iraq)	89.2	Medium	Moderate
Sharma & Joshi (2024) [42]	Explainable AI (SHAP + XGB)	Remote Sensing Data	90.1	High	High
Proposed ICRS	**Hybrid RF + XGB + DNN + Bayesian Optimization**	**Multi-Regional Datasets (India, Africa)**	**93.1**	**Very High**	**High (via SHAP)**

Comparison of our proposed framework with baseline approaches given in the evaluation is shown in Table 6. Overall, our proposed framework achieved the highest classification accuracy of 93.1%, outperforming other methods. In addition to high prediction accuracy, the proposed framework also generalizes well across datasets collected from different regions. With SHAP based explanations, it offers explainable crop recommendations. Hence, our proposed framework acts as a strong explainable baseline for crop recommendation systems that can be used in place of other machine learning and explainable AI-based approaches.

Implications for Society and Technology

The integrative framework addresses this need by bringing together heterogeneous agricultural datasets with explainable machine learning to support data-driven crop recommendation and agricultural decision-making systems. For example, it may allow farmers to determine which crops would be more suitable to grow given the soil, climate and other environmental conditions of their farms.

It can also help them better capitalize on the resources that are currently available to them and avoid the pitfalls of selecting crops that are less suitable to their land. Beyond farm-level crop recommendations, the framework could potentially be used to inform higher-level, regional agricultural planning and resource management initiatives. Incorporating explainability also allows users to see what factors contribute to a given prediction. This helps users grasp the rationale for the crop recommendations they receive.

5. CONCLUSION

The proposed hybrid ensemble method takes in soil features, climate features, remote sensing indicators and prior growing crops as inputs to provide recommendation to crops that can be grown. The framework was designed to address key challenges related to prediction accuracy, regional adaptability, and model interpretability in agricultural decision-support systems. It integrates soil characteristics, climatic variables, remote sensing indicators, and historical crop information, and combines Random Forest, XGBoost, and Deep Neural Network models through Bayesian-optimized ensemble learning. SHAP-based explainability was also included to identify the factors that influence crop suitability predictions and to make the recommendation process more transparent. Experimental results shows that the hybrid ensemble learning approach outperformed the base learning models by attaining overall classification accuracy of 93.1% and showed stable performance over datasets collected from different regions.

Results from this study show that hybrid ensemble learning integrated with explainable AI can result in accurate as well as explainable crop recommendation systems. The framework provides a scalable and interpretable approach that can support farmers, agricultural planners, and policymakers in making informed crop selection decisions. Future work may include the use of real-time IoT sensor data, high-resolution satellite imagery, and advanced climate forecasting information to improve recommendation accuracy and adaptability under changing agricultural conditions.

References

1. Abhishek, K., & Singh, D. (2020). Ensemble learning approaches for crop recommendation and disease detection. Procedia Computer Science, 167, 2401–2412. https://doi.org/10.1016/j.procs.2020.03.296
2. Afzal, H., Amjad, M., Raza, A., Munir, K., & Villar, S. G. (2025). Incorporating soil information with machine learning for crop recommendation. Scientific Reports, 15(2), 88676. https://doi.org/10.1038/s41598-025-88676-z
3. Ahmed, H. A. B., Ezzahar, J., Er-Raki, S., Khabba, S., & Elfarkh, J. (2025). Multi-objective calibration of a dual-source SVAT model using root zone soil moisture: Application to winter wheat in semi-arid Morocco. Sciforum Preprint. https://sciforum.net/paper/view/24706
4. Ahmed, S., Mahmoud, A. S., Farg, E., & Mohamed, A. M. (2023). Use of ensemble learning and big data in crop identification. Heliyon, 9(7), e01546.
5. Albaaji, G. F., & Chandra, V. S. S. (2025). Hybrid ensemble framework for integrating environmental parameters for crop yield prediction. Neural Computing and Applications. https://doi.org/10.1007/s00521-025-11221-6
6. Behera, S., Reddy, N. T. K., & Pradhan, S. (2025). Development of an intelligent agricultural decision support system for crop recommendation using machine learning techniques. Next-Generation Computing Systems and Technologies (NGCST), ICCK.
7. Bhattacharjee, R., & Sen, R. (2024). AI-driven decision support for region-specific crop planning: A hybrid model perspective. Expert Systems with Applications, 247, 122048. https://doi.org/10.1016/j.eswa.2024.122048
8. Bhattacharya, D., & Rahman, S. (2024). Data-driven crop management using integrated soil and climate analytics: A systematic review. Computers and Electronics in Agriculture, 216, 108223. https://doi.org/10.1016/j.compag.2024.108223

9. Bhavana, M., & Rao, K. S. (2025). Smart agriculture with Internet of Things for precise crop prediction using interfused machine learning and advanced stacking ensemble from soil parameters. ResearchGate Preprint.
10. Cai, S., Sun, C., & Tian, J. (2025). Towards smart agriculture: AI-driven prediction of key genes for revolutionizing crop breeding. Planta, Springer. https://doi.org/10.1007/s00425-025-04841-8
11. Candiago, S., Marsoner, T., Tscholl, S., & Fraga, H. (2025). A global database of olive oil geographical indications and their properties. Frontiers in Sustainable Food Systems, 9(1641032). https://doi.org/10.3389/fsufs.2025.1641032
12. Cao, J., Wang, H., Li, J., Tian, Q., & Niyogi, D. (2022). Hybrid subseasonal-to-seasonal ensemble prediction for winter wheat yields. Remote Sensing, 14(7), 1707. https://doi.org/10.3390/rs14071707
13. Choi, J. W., Hidayat, M. S., Cho, S. B., Hwang, W. H., & Lee, H. (2025). Recent trends in ML, DL, ensemble learning, and XAI for evaluating crop yields. Plants, 14(18), 2841. https://doi.org/10.3390/plants14182841
14. Darra, N., Espejo-Garcia, B., Kasimati, A., & Kriezi, O. (2023). Satellite-driven ensemble ML for tomato yield prediction. Sensors, 23(5), 2586.
15. Dey, B., Ferdous, J., & Ahmed, R. (2024). Machine learning-based recommendation of agricultural and horticultural crops under the regime of NPK, soil pH and climatic variables. Heliyon, 10(6), e19143. https://doi.org/10.1016/j.heliyon.2024.e19143
16. Ding, Y., Wang, M., Jin, J., Sun, Z., Zhang, J., & Zhu, L. (2025). Risk of real-time irrigation decision-making system for farmland in arid irrigation districts: Methodology and case study. Agricultural Water Management, Elsevier. https://doi.org/10.1016/j.agwat.2025.108905
17. Duisenbekova, A. (2025). Machine learning in ensuring food security: Domestic and international practices. Scientific Collection "InterConf+".
18. Fadl, M. E., Zekari, M., Labad, R., & Faqeih, K. Y. (2025). Integrating RUSLE, AHP, GIS, and cloud-based geospatial analysis for soil erosion assessment under Mediterranean conditions. Scientific Reports, 15(1), 22503. https://doi.org/10.1038/s41598-025-22503-3
19. FAO. (2023). The State of Food and Agriculture: Leveraging Data for Sustainable Agriculture. https://doi.org/10.4060/cc8013en
20. Gomez, C., & Liu, M. (2023). Integration of NDVI and climate data for yield prediction using ensemble deep learning. Remote Sensing of Environment, 298, 113878. https://doi.org/10.1016/j.rse.2023.113878
21. Haider, S. T., Ge, W., Li, J., Rehman, S. U., & Imran, A. (2024). Ensemble machine learning framework for cotton crop yield prediction using weather parameters. IEEE Access, 12, 10664030.
22. Haq, S. I. U., Wang, G., Khan, S. N., Song, C., & Ma, C. (2025). Multi-dimensional optical remote sensing in agriculture: Spectral, angular, and spatial scaling for crop stress monitoring. Smart Agricultural Systems, Elsevier. https://doi.org/10.1016/j.smartags.2025.100814
23. Hasan, M., Marjan, M. A., Uddin, M. P., & Afjal, M. I. (2023). Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation. Frontiers in Plant Science, 14, 1234555. https://doi.org/10.3389/fpls.2023.1234555
24. Afra Naaz, Mustafa Basthikodi, Comprehensive review on women safety technologies: Indian perspective, Women's Studies International Forum, Volume 114, 2026, 103241, ISSN 0277-5395, https://doi.org/10.1016/j.wsif.2025.103241. [25] Jabed, M. A., & Murad, M. A. A. (2024). Crop yield prediction in agriculture: A comprehensive review. Heliyon, 10(11), e16867.
25. Kumar, P., & Champa, H. N. (2025). Enhancing agricultural yield: A unified stacking ensemble method for crop recommendations using soil and weather attributes. Int. J. Environ. Sci., 8(2), 1189.
26. Basthikodi, M., Faizabadi, A. R. & Ahmed, W. HPC based algorithmic species extraction tool for automatic parallelization of program code. Int. J. Recent. Technol. Eng. 8, 1004–1009 (2019).
27. Miller, J. A., & Petrovic, S. (2023). Hybrid ensemble learning for sustainable crop recommendation under data uncertainty. Applied Soft Computing, 142, 110409. https://doi.org/10.1016/j.asoc.2023.110409
28. Mohamed, S. A., Maksoud, O. O. A., Fathy, A., & Mohamed, A. S. (2025). A hybrid deep learning and rule-based model for smart weather forecasting and crop recommendation using satellite imagery. Scientific Reports, 15(2), 21506. https://doi.org/10.1038/s41598-025-21506-4
29. Mythili, K., & Rangaraj, R. (2021). Deep learning with PSO-based hyperparameter tuning for crop recommendation. Indian Journal of Science and Technology, 14(17), 450.
30. Nimmagadda, S. (2025). Real-time crop prediction and recommendation using machine learning. Proc. Indian National Sci. Acad., 91(3), 1857.
31. Basthikodi, Mustafa, Ananth Prabhu, and Anush Bekal. "Performance analysis of network attack detection framework using machine learning." Sparklinglight transactions on artificial intelligence and quantum computing (STAIQC) 1.1 (2021): 11-22.
32. Pandey, A., & Misra, S. (2021). Big data analytics in agriculture: A review on applications, challenges, and future directions. Information Processing in Agriculture, 8(4), 447–464. https://doi.org/10.1016/j.inpa.2021.05.004
33. Pandey, G. (2025). Climate-aware crop recommendation using IoT and deep ensemble learning. IJRTTE, 4(2).
34. Patel, R., & Singh, A. (2024). AI-based predictive models for soil-crop suitability assessment in semi-arid regions. Agricultural Systems, 216, 103964. https://doi.org/10.1016/j.agsy.2024.103964
35. Pham, H. T., Awange, J., Kuhn, M., Nguyen, B. V., & Bui, L. K. (2022). Enhancing crop yield prediction using satellite vegetation indices. Sensors, 22(3), 719.

36. Mustafa Basthikodi, Poornima B V, Developing an explainable human action recognition system for academic environments: Enhancing educational interaction, Results in Engineering, Volume 26, 2025, 105014, ISSN 2590-1230, https://doi.org/10.1016/j.rineng.2025.105014.
37. Reddy, S., & Kalra, S. (2022). Integrating IoT and ML for precision crop recommendation: Challenges and prospects. Sensors, 22(19), 7538. https://doi.org/10.3390/s22197538
38. Sarkar, S., Leyton, J. M. O., Noa-Yarasca, E., & Adhikari, K. (2025). Integrating remote sensing and soil features for corn yield prediction. Sensors, 25(2), 543. https://doi.org/10.3390/s25020543
39. Saxena, S., & Raj, R. (2023). Adaptive machine learning frameworks for precision agriculture. IEEE Access, 11, 115234–115248. https://doi.org/10.1109/ACCESS.2023.3310849
40. Schweng, S., Bernardini, L., Keiblinger, K., & Kaul, H. P. (2026). What can artificial intelligence do for soil health in agriculture? Agriculture, Ecosystems & Environment Review, Elsevier. https://doi.org/10.1016/j.agee.2026.109847
41. Sharma, P., & Joshi, A. (2024). Explainable artificial intelligence for precision agriculture: Bridging the gap between data and decisions. Computers and Industrial Engineering, 195, 110934. https://doi.org/10.1016/j.cie.2024.110934
42. Salins, R.D., Ashwin, T.S., Prabhu, G.A. et al. Person identification from arm's hair patterns using CT-twofold Siamese network in forensic psychiatric hospitals. Complex Intell. Syst. 8, 3185–3197 (2022). https://doi.org/10.1007/s40747-022-00771-0
43. .[44] Sultana, S., Hasan, T., Tanjil, F. I., & Islam, S. (2025). IoT and hybrid ML-based web framework for crop recommendation. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5132627
44. Sun, Z., Sun, L., Liu, Y., Wuyun, D., & Wang, L. (2025). Near real-time prediction of crop vegetation index time series during the entire growth period using crop reference curves. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5653010
45. Afranaaz and M. Basthikodi, "Enhancing Women's Safety - A Comprehensive Review of Emerging Technologies and Automated Crime Detection Using ML/DL Algorithms," 2024 International Conference on Computing, Semiconductor, Mechatronics, Intelligent Systems and Communications (COSMIC), Mangalore, India, 2024, pp. 1-8, doi: 10.1109/COSMIC63293.2024.10871586.
46. Turner, G., Kerr, I., Kovats, S., & Green, R. (2025). Reviewing methods for quantifying climate risks to food security and safety in WHO EURO. International Journal of Public Health Research, ProQuest.
47. Waghaye, A. M., Kumar, M., & Singh, R. K. (2025). Spatiotemporal trend detection and modeling of reference evapotranspiration using machine learning techniques for the Marathwada region, Maharashtra, India. Water Practice and Technology, 20(3), 139. https://doi.org/10.2166/wpt.2025.139
48. Wang, X., Miao, Y., Dong, L., Mi, G., & Batchelor, W. D. (2025). Precision nitrogen management for maize based on crop modeling, remote sensing, and machine learning. Computers and Electronics in Agriculture, Elsevier. https://doi.org/10.1016/j.compag.2025.108230
49. Shruthi, Maheshappa, Basthikodi Mustafa, and G. Ananth Prabhu. "Parellel Implementation of Modified Apriori Algorithm on Multicore Systems." Proceeding of IMCICICSIT (2016).
50. Zhang, Y., Chen, J., & Wang, P. (2022). Crop yield estimation and recommendation using hybrid random forest-XGBoost models. Information Processing in Agriculture, 9(4), 452–462. https://doi.org/10.1016/j.inpa.2021.09.005
51. Wannasingha, U. H., Waqas, M., & Wangwongchai, A. (2025). Advances in artificial intelligence to model the impact of El Niño–Southern Oscillation on crop yield variability. MethodsX, 12(5), 101947. https://doi.org/10.1016/j.mex.2025.101947