



International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Smart Crop Selection and Yield Prediction Through Machine Learning and Explainable Artificial Intelligence

Prof. Sankareswari¹, Prof. Rashmi More¹, Prof. Atiya Kazi¹, Prof. Madhura Zagade¹

¹Department of Information Technology, Finolex Academy of Management and Technology, Ratnagiri, Maharashtra, India.

Abstract

Effective agricultural decision making requires both the identification of a suitable crop under prevailing soil and climatic conditions and the estimation of its expected productivity. This study proposes an integrated dual model machine learning framework that combines crop recommendation with crop yield prediction. The crop recommendation component was developed using 2,200 observations representing 22 crop classes and seven soil and climatic attributes, namely nitrogen, phosphorus, potassium, temperature, humidity, soil pH, and rainfall. The yield prediction component utilized a historical crop production dataset containing 246,091 records, which was reduced to 240,505 observations after preprocessing. The processed dataset covered 90 crops, 33 states, 646 districts, and agricultural records from 1997 to 2015. Decision Tree, Gaussian Naive Bayes, Support Vector Machine, Logistic Regression, and Random Forest were evaluated for crop recommendation, while Linear Regression, Decision Tree Regression, and Random Forest Regression were compared for yield estimation. The selected Random Forest classifier achieved an accuracy of 99.55 percent, precision of 99.57 percent, recall of 99.55 percent, and macro F1 score of 99.55 percent. For crop yield prediction, Random Forest achieved a coefficient of determination of 0.9475, a mean absolute error of 0.7033, and a root mean squared error of 2.6432. Statistical analysis confirmed significant overall performance differences among the evaluated classification and regression models. Explainability analysis further identified humidity as the most influential variable for crop recommendation and season as the dominant contributor to yield prediction. The proposed framework integrates crop selection, crop mapping, yield estimation, production estimation, statistical validation, and interpretable machine learning within a unified agricultural decision support system.

Keywords: Crop recommendation; Crop yield prediction; Random Forest; Explainable artificial intelligence; Precision agriculture

1. Introduction

Agriculture is increasingly being transformed by data-driven technologies that support crop selection, yield forecasting, resource management, and precision farming. Variations in soil characteristics, nutrient availability, weather conditions, and regional agricultural practices make crop selection and production planning complex decision-making tasks. Conventional crop selection practices often depend on historical experience and generalized recommendations, which may not adequately account for the interaction among soil nutrients and climatic conditions at a particular location. The increasing availability of agricultural datasets has therefore encouraged the application of machine learning (ML) techniques for developing data-driven agricultural decision-support systems. Recent studies have demonstrated that ML models can utilize soil nutrients such as nitrogen, phosphorus, and potassium together with temperature, humidity, soil pH, and rainfall to identify suitable crops under varying agro-climatic conditions. Crop recommendation is an important component of precision agriculture because the selection of an appropriate crop can influence subsequent resource requirements and production outcomes. Several ML-based studies have investigated crop classification using environmental and soil-related parameters. For example, a comparative study using NPK, soil pH, temperature, rainfall, and humidity demonstrated the applicability of Support Vector Machine (SVM), Random Forest (RF), Decision Tree, K-Nearest Neighbors, and XGBoost for crop recommendation, with ensemble-based approaches achieving high predictive performance. Similarly, studies using the commonly adopted 2,200-record crop recommendation dataset have investigated the relationships between seven agronomic and climatic variables and 22 crop classes. These studies establish the feasibility of ML-based crop recommendation; however, predictive performance alone does not necessarily explain why a particular crop is recommended for a given combination of field conditions. In parallel with crop recommendation, crop yield prediction provides an important quantitative component of agricultural decision support. Yield prediction can assist in production planning, resource allocation, and estimation of expected agricultural output. Unlike crop recommendation, which is generally formulated as a classification problem, yield prediction is commonly

treated as a regression problem in which historical production, cultivated area, crop type, geographical location, season, and temporal information can be used to estimate yield. Machine learning models such as Linear Regression, Decision Tree Regression, and Random Forest Regression have been widely considered for agricultural prediction tasks because agricultural relationships are often nonlinear and may involve interactions among multiple categorical and numerical variables. Consequently, combining crop recommendation with yield prediction can provide a more comprehensive decision-support process than treating crop selection and production estimation as independent analytical tasks. Despite the growing use of ML in agriculture, the interpretability of predictive models remains an important concern. Ensemble models such as Random Forest can capture nonlinear relationships and interactions between input variables, but their individual prediction mechanisms are not readily apparent to end users. In agricultural applications, this limitation is particularly relevant because a prediction is expected to support a real-world decision involving soil, climate, crop selection, and production planning. A model that only provides a predicted crop or yield without indicating the contribution of the input variables provides limited information about the reasoning underlying that prediction. Recent research on explainable artificial intelligence (XAI) in agriculture has consequently emphasized the importance of transparency and interpretable model outputs for agricultural decision-support systems. Studies specifically examining XAI for crop recommendation have demonstrated that feature-attribution methods can be used to identify the variables contributing to model predictions. Explainable Artificial Intelligence provides a mechanism for examining the relationship between model inputs and predictions without necessarily replacing the underlying predictive model. Among the available XAI techniques, SHapley Additive exPlanations (SHAP) provides a theoretically grounded feature-attribution framework based on Shapley values from cooperative game theory. Lundberg and Lee introduced SHAP as a unified approach for assigning an importance value to each feature for an individual prediction [6]. Subsequent work developed efficient tree-based SHAP methods for explaining predictions from tree ensembles such as Random Forest and gradient-boosted decision trees. SHAP explanations can be examined at both global and local levels. Global explanations characterize the overall contribution of features across a collection of observations, whereas local explanations describe how individual features contribute to a particular prediction. This distinction is useful in agricultural applications because it allows researchers to identify general patterns in model behaviour while simultaneously explaining individual crop or yield predictions. Recent agricultural XAI research has demonstrated increasing interest in SHAP-based explanations. For example, XAI-based crop recommendation research has used feature-attribution techniques to identify the influence of soil and environmental variables on crop recommendations. More recent work has also explored SHAP-based interpretation of crop recommendation models and emphasized the importance of transparent decision-making in agricultural AI systems. Similarly, XAI approaches have been investigated for crop-yield prediction to address the interpretability limitations of conventional ML models. Nevertheless, much of the existing work focuses either on crop recommendation or yield prediction as an individual prediction task. There remains scope for an integrated framework in which a recommended crop is subsequently incorporated into a yield prediction process and the resulting decisions are explained using a consistent XAI mechanism. Another important consideration in agricultural datasets is the treatment of categorical variables. Yield-related datasets commonly contain attributes such as State, District, Season, and Crop. Label encoding can convert these attributes into numerical representations for model training; however, the numerical codes assigned to categories should not be interpreted as inherently ordered or continuous quantities. For an interpretable agricultural decision-support system, explanations should therefore retain the semantic identity of the original attributes rather than suggesting that a larger category code represents a greater agricultural effect. This issue becomes particularly relevant when SHAP explanations are generated for tree-based models because the resulting feature contributions must be communicated in a manner that is meaningful to agricultural stakeholders. To address these challenges, this study develops an integrated machine-learning framework for crop recommendation and yield prediction with explainable artificial intelligence. The proposed framework operates in two sequential stages.

In the first stage, a crop recommendation model uses seven agronomic and climatic attributes—nitrogen, phosphorus, potassium, temperature, humidity, pH, and rainfall—to identify a suitable crop. A dataset containing 2,200 observations is processed using stratified training and testing, and five classification algorithms namely Decision Tree, Naive Bayes, Support Vector Machine, Logistic Regression, and Random Forest are comparatively evaluated using Accuracy and Macro F1-score. Based on the experimental evaluation, Random Forest is selected as the crop recommendation model. In the second stage, a crop production dataset containing 246,091 records is processed to develop a yield prediction model. Records with missing production values, invalid or zero cultivated area, extreme yield values, and crops with fewer than 50 observations are removed during preprocessing. Yield is derived from production and cultivated area, while State, District, Season, and Crop are encoded for model development. Linear Regression, Decision Tree Regression, and Random Forest Regression are comparatively evaluated using R^2 , Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). The selected Random Forest regression model subsequently receives the recommended crop together with State, District, Crop Year, Season, and Area to estimate crop yield. Estimated production is then calculated from the predicted yield and cultivated area. The framework further incorporates

SHAP-based XAI without modifying or retraining the selected Random Forest models. For crop recommendation, TreeSHAP is applied to the seven agronomic and climatic features. Global interpretability is obtained using the mean absolute SHAP value across a representative subset of test observations, thereby identifying the overall contribution of each feature to model predictions. Local interpretability is obtained for an individual crop recommendation by examining the SHAP values associated with the predicted crop class and presenting the positive and negative contributions of the input variables through a waterfall visualization. The same interpretability strategy is extended to the yield prediction model using State, District, Crop Year, Season, Crop, and Area. For computational efficiency, a representative subset of test observations and approximate TreeSHAP computation are used for the larger yield dataset.

The main contribution of this study is therefore not limited to predictive modelling but extends to the integration of crop recommendation, yield estimation, production estimation, and model explanation within a sequential decision-support framework. The framework connects environmental and soil conditions to crop selection and subsequently links the selected crop with geographical, temporal, and cultivated-area information for yield estimation. The incorporation of global and local SHAP explanations provides an additional layer of transparency by showing both the general importance of agricultural variables and the contribution of individual variables to a specific prediction. This integrated approach is intended to provide a more transparent and practically interpretable ML-based framework for agricultural decision support.

2. Literature Review

Decision support systems for agriculture have transitioned from being single-function to integrated Agriculture 4.0 platforms that integrate diverse data elements and provide decision support for multiple aspects of farm management [1]. The studies examined in this section are organized around the two tasks: recommending which crop is to be grown and predicting the amount of yield that will be obtained and the methodological advances that have been developed in support of these tasks. Machine learning that is used for crop recommendation is discussed in Section 2.1, while the benchmark datasets are dominant and ensemble classifiers. The problem of predicting the yield of crops is tackled at various scales, from country to field level models in Section 2.2. In Sections 2.3 and 2.4, the two methodological approaches to the drawbacks of the single-source, single-task model, namely multimodal data fusion and spatio-temporal/graph deep learning are reviewed. Section 2.5 will examine how these recommendations can be easily understood by farmers and agronomists using explainable AI. The review concludes that there is a gap, throughout all five themes, in the absence of an integrated framework that is able to carry out the crop recommendation and yield forecasting at the same time on heterogeneous agricultural data.

2.1 Machine learning for crop recommendation

Most crop recommendation research used the 2200 records, 22 crops soil-and-climate data and compared the common classifiers. Five models were compared by Cheema and Pires [2] on a 20-crop subset and it was concluded that the best model was a decision tree with AdaBoost (accuracy of about 98%). In the 70:30 split, Kaur et al. [3] compared the performance of SVM, KNN, gradient boosting, AdaBoost, LightGBM and CatBoost. The results were 99.70% for AdaBoost and between about 97.6% and 98.5% for the other models. Alam et al. [4] used soft voting on six learners (LR, NB, KNN, DT, RF, XGBoost) with a performance of 99.54%, followed by Naïve Bayes and Random Forest with 99.31%. Dey et al. [5] took a different approach to the problem. They applied a Kaggle dataset consisting of 2100 records which includes applied NPK fertilizer and divided the crops into agricultural and horticultural. When they were divided into two groups, XGBoost achieved an accuracy of 99.09% and 99.3%, respectively. When they were combined, XGBoost had an accuracy of 98.51%. The authors have suggested that separate modelling of crop categories is better than the combined modelling of crop categories. What emerges from the above studies are two things. The model rankings are unstable, with accuracies close to ceiling, and only a few tenths of a percent apart between methods. Almost all of them conduct random hold-out splits on a small set of data. A number of authors recognize that this data is restricted geographically [3, 5].

2.2 Machine learning for crop yield prediction

The inputs vary depending on the scale of yield prediction studies; ranging from the country level to the individual field. Aworka et al. [6] worked at the coarsest scale. They used FAO records of pesticide application, agricultural emission of N_2O , rainfall and temperature to model annual yields of four crops in 14 East African countries. On a random 70:30 split, Random Forest reached $R^2 = 92.3\%$ (RMSE 0.343 on log yield), ahead of gradient boosting (90.2%) and SVR (86.4%). Their models were based on default values, and they removed Country and Year from features. Razavi et al. [7] modeled five Senegalese crops collectively (pearl millet, groundnut, rice, maize, sorghum) with the help of remotely sensed and soil features. Grid search resulted in improvement of Random Forest and LightGBM by CatBoost and XGBoost, the CatBoost model showing the best performance with lowest RMSE (0.294 on standardized yield). Due to the limited amount of data, they generated synthetic samples with VAE and computed the permutation importance to identify the most important drivers. Omary et al. [8] processed data at the level of the farm, using 9,897 records of smallholders

from three regions of Tanzania (2022–2024). Extra Trees was the top of four teams. It had a pooled multi-crop R^2 value of 95%, but only an R^2 value of 79–81% for maize, rice and beans alone. The authors indicate that pooling can actually obscure the problem of difficulty of a crop. They report that their yields are farmer-reported, a limitation. Gharakhanlou and Perez [9] used climate data, vegetation indices, soil characteristics and honeybee census information to predict canola and soybean yields at census-subdivision level for all of Canada. The best one was XGBoost (R^2 0.95 and 0.96) and boosting outpaced bagging in general. Honeybees were the most important species in both crops, accounting for 52.34% of the canola system and 57.18% of the soybean system. They also made projections for yield under RCP2.6, RCP4.5 and RCP8.5 scenarios.

2.3 Remote sensing, multimodal and hybrid approaches

Multimodal data fusion enables agricultural prediction models to combine heterogeneous information sources, such as remote sensing imagery, weather, soil, topographic, and field data, to capture complementary characteristics of crop growth. Fatholouloumi et al. [10] employed the maximum and minimum of Sentinel-2 vegetation indices of maize instead of fixed-date observations. Han et al. [11] integrated a process-based crop model to data assimilation and extreme-weather indicators and ML in the summer maize. Random Forest was found to work the best and predictions were reliable 82 days prior to maturity. Mena et al. [12] introduced a gated fusion network for sub field yield using Sentinel-2 imagery and weather, soil and topography. The model had an R^2 of 0.68 at sub-field scale for Argentina, and of approximately 0.80 at field scale. Yewle et al. [13] constructed a deep ensemble model which combines SAR, optical and meteorological data for rice in the Mekong Delta with MAE of 336 kg/ha. The contribution of weather to UAV multispectral imagery by Mia et al. [14] showed that the best models could be obtained when incorporating weekly weather data, but there was still some variation in the spatial patterns between models with a similar accuracy. Maimaitjiang et al [15] used RGB, multispectral and thermal UAV data for soybean, and they obtained $R^2 = 0.720$ (rRMSE 15.9%) after intermediate fusion for DNN. Li et al. [16] estimated the wheat's yield by a two-step fusion model with a R^2 of 0.8536 (rRMSE 16.12%) and 0.8370 in the following year, followed by screening of drought-tolerant varieties.

2.4 Spatio-temporal and graph deep learning

A second line explicitly models spatial and temporal dependence. Qiao et al. [17] used a knowledge-guided spatial-temporal attention graph network (KSTAGE) and tested the model at county level in China and the contiguous United States. Yang et al. [18] proposed KGML-Graph for corn yield in 627 U.S. counties (2000–2020). It decreased the RMSE by at least 10.8% in temporal extrapolation and maintained its accuracy in regions without data. For the EPFL multimodal dataset, Sarkar et al. [19] obtained an accuracy of around 97% by applying a Meta-Transformer with temporal graph neural networks to the task of yield classification. Gamidell and Chaurasia [20] developed the explainable multi-season CNN-BiLSTM model using Sentinel-2 data that obtained a RMSE of 0.273 t/ha and R^2 of 0.894.

2.5 Explainable AI for Agricultural Decision Support

A new series of studies explores why a crop is recommended and how effective the recommendation is. Alam et al. [4] introduced an entropy measure of uncertainty to the soft voting ensemble. They tested for robustness by introducing Gaussian noise to the test inputs, and looked at the SHAP patterns for models of different uncertainty. Shastri et al. [21] used gradient boosting (99.27% accuracy) and explained individual recommendations with LIME. Shams et al. [22] created a model, called XAI-CROP, based on decision-tree classification and its explanations using LIME, and evaluated its performance against GB, DT, RF and Naïve Bayes models. They obtain regression metrics MSE 0.9412, MAE 0.9874 and R^2 0.94152 for a recommendation task. Yaganteeswarudu et al. [23] developed an explanatory system using the LIME that provides explanations such as the reasons why a person is recommended to buy coffee when if it is hot, cold or rainy outside. They correlate explanations with transparency and user trust. Badshah et al. [24] used a combination of feature importance, LIME and k-fold cross validation, and the Random Forest recorded a 99.7% accuracy score. The works are all agreed that explanation is appreciated, but the explanations are typically descriptive. Very few consider their requirements to agronomic needs, and the uncertainty work of seems to be there only. All the reviewed studies show the effectiveness of developing crop recommendations and yield prediction using machine learning, spatio-temporal modelling, explainable AI, and multimodal data fusion. But there are various approaches with various goals, data modalities, and level of integration and many studies focus on a specific prediction task or selected combinations of these. This suggests a research need of developing a single framework for agricultural decision support system that simultaneously uses heterogeneous agricultural data for crop recommendation and yield forecasting, and provides meaningful output for agricultural decision making.

3. Proposed Methodology

As illustrated in Fig. 1, the proposed methodology implements a dual model machine learning framework in which crop recommendation is formulated as a multiclass classification problem and crop yield prediction is formulated as a regression problem. The crop recommendation branch uses soil and climatic attributes,

including nitrogen, phosphorus, potassium, temperature, humidity, pH, and rainfall, while the yield prediction branch uses geographical, temporal, seasonal, crop, and cultivated area information. Each dataset is preprocessed independently to preserve its original structure and prediction objective, including data quality assessment, removal of invalid records, categorical encoding, yield computation, outlier trimming, and hold out train test partitioning. Multiple candidate classifiers and regressors are then evaluated, after which Random Forest models are selected for both tasks. The classifier aggregates the decisions of 200 trees using majority voting, whereas the regressor estimates yield from the mean output of 200 regression trees with controlled tree depth. Integration is achieved through crop name harmonization, which maps the crop predicted by the classification model to the corresponding crop representation used by the yield dataset before applying the same encoding transformations used during training. The harmonized crop information is combined with contextual variables and passed to the regression model to estimate yield and production. Model reliability is assessed using independent test sets together with nonparametric statistical tests, namely Cochran Q and pairwise McNemar tests for classification and Friedman and pairwise Wilcoxon signed rank tests for regression, followed by Holm correction for multiple comparisons. Finally, SHAP based global and local interpretability is applied to both Random Forest models to quantify feature contributions and provide transparent explanations for the resulting crop recommendations and yield estimates. diagram.

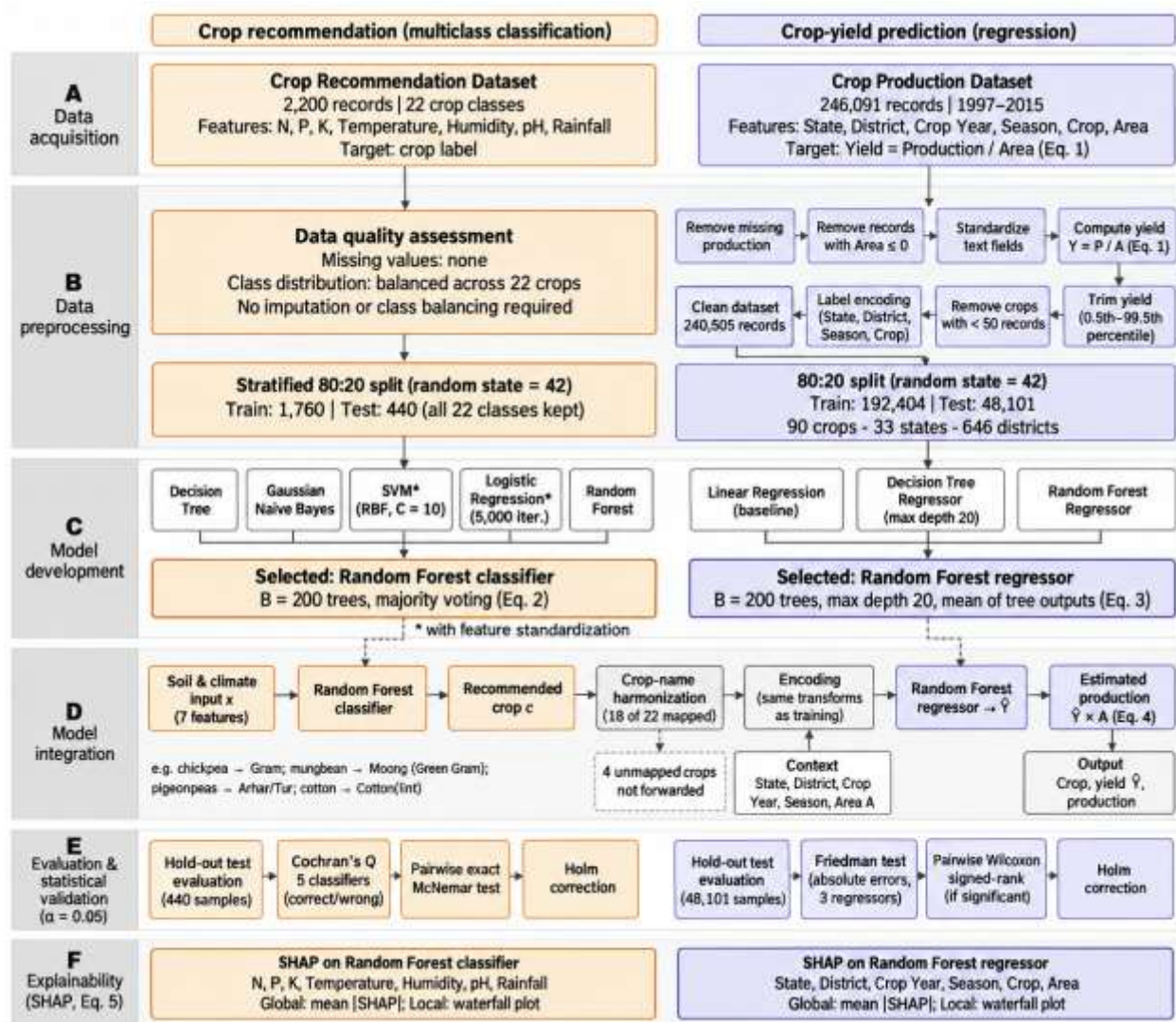


Fig. 1. Overview of the proposed dual-model framework. (A) Input datasets; (B) dataset-specific preprocessing; (C) comparison of candidate classifiers and regressors and selection of Random Forest models; (D) sequential integration through crop-name harmonization; (E) hold-out evaluation and non-parametric statistical validation; (F) SHAP-based explainability.

3.1 Datasets and Problem Formulation

Two complementary agricultural datasets were used. The Crop Recommendation Dataset represents the suitability problem using soil nutrient and climatic measurements, whereas the Crop Production Dataset provides historical agricultural records for modelling yield. Their principal characteristics are summarized in Table 1.

Table 1. Input Datasets Used in the Proposed Dual-Model Framework

Dataset	Purpose	Records	Input Features	Target
Crop Recommendation Dataset	Crop recommendation	2,200	N, P, K, Temperature, Humidity, pH, Rainfall	Crop Label
Crop Production Dataset	Crop yield prediction	246,091 (240,505 after preprocessing)	State, District, Crop Year, Season, Crop, Area	Yield = Production / Area

The Crop Recommendation Dataset contains 2,200 observations distributed across 22 crop classes. Each observation is described by nitrogen (N), phosphorus (P), potassium (K), temperature, humidity, soil pH, and rainfall. These seven variables define the environmental feature space used to learn the crop recommendation function. The Crop Production Dataset initially contains 246,091 historical records. After preprocessing, 240,505 valid observations remain, covering 90 crops, 33 states, and 646 districts over the period 1997-2015. Because the historical production data provide total production and cultivated area rather than yield directly, crop yield was derived for every valid record. For the i -th record, yield is defined by Eq. (1), where P_i denotes recorded production and A_i denotes cultivated area. Equation (1) converts absolute production into productivity per unit area and supplies the regression target used in the yield-prediction stage

$$Y_i = \frac{P_i}{A_i} \quad (1)$$

3.2 Data Preprocessing

The two datasets were preprocessed independently because their structures and prediction objectives differ. The crop recommendation data were examined for missing values, class distribution, and consistency of the seven numerical attributes. The dataset contains complete observations and balanced representation across the 22 crop categories; therefore, no missing-value imputation or class-balancing procedure was introduced. For the crop production data, records with unavailable production values were discarded because yield could not be calculated for those observations. Records with zero or negative cultivated area were also removed to avoid undefined or physically invalid yield values in Eq. (1). The State, District, Season, and Crop attributes were standardized as text fields and redundant leading and trailing spaces were removed before encoding. After computing yield, extreme observations were restricted using the 0.5th and 99.5th percentiles of the empirical yield distribution. This trimming step reduces the influence of exceptionally large or small yield values while retaining the dominant structure of the historical data. Crop categories represented by fewer than 50 records were subsequently excluded so that the regression model was not trained on categories with inadequate historical support. The categorical variables State, District, Season, and Crop were transformed using label encoding. The resulting yield-prediction feature vector therefore comprised encoded state, encoded district, crop year, encoded season, encoded crop, and cultivated area. Both learning tasks used an 80:20 training-test split with random state 42. Stratified splitting was applied to crop recommendation, resulting in 1,760 training and 440 test samples while preserving all 22 crop classes. The preprocessed yield data produced 192,404 training and 48,101 test observations.

3.3 Crop Recommendation Model

Crop recommendation was formulated as a multiclass classification problem in which the seven soil and climatic variables were used to identify one of 22 crop classes. Five classifiers were examined under the same hold-out evaluation protocol: Decision Tree, Gaussian Naive Bayes, Support Vector Machine (SVM), Logistic Regression, and Random Forest. The set provides complementary modelling characteristics, including probabilistic classification, linear decision functions, nonlinear kernel separation, single-tree partitioning, and ensemble learning. Feature standardization was included in the SVM and Logistic Regression pipelines because both methods are sensitive to differences in input scale. The SVM employed a radial basis function kernel with penalty parameter $C = 10$, whereas Logistic Regression used a maximum of 5,000 optimization iterations. The tree-based models were fitted directly to the original numerical variables. The final crop recommendation component was implemented using a Random Forest containing 200 decision trees. Each tree independently produces a crop class, and the ensemble combines these decisions through majority voting. For environmental input vector x , the final crop prediction is expressed by Eq. (2). In Eq. (2), the function associated with each tree denotes the class predicted by that tree, while B is the total number of estimators. The use of an ensemble reduces dependence on any single tree structure and allows nonlinear interactions among nutrient and climatic variables to contribute to the recommendation.

$$\hat{c} = \text{mode}\{h_1(x), h_2(x), \dots, h_B(x)\}, \quad B = 200 \quad (2)$$

3.4 Crop-Yield Prediction Model

The second stage estimates crop yield from historical contextual variables. Linear Regression, Decision Tree Regression, and Random Forest Regression were compared. Linear Regression was used as a baseline, while the two tree-based approaches were included to model nonlinear interactions among location, crop year, season, crop type, and area. The Decision Tree Regressor used a maximum depth of 20. The selected Random Forest Regressor consisted of 200 regression trees with maximum depth 20. Unlike a single regression tree, the

forest combines estimates produced by multiple trees. For a yield-prediction feature vector z , the final estimate is obtained according to Eq. (3). Here, each tree contributes an individual yield estimate. As shown in Eq. (3), the ensemble output is the mean of the individual tree estimates, thereby reducing the variance associated with a single partitioning structure.

$$\hat{Y} = \frac{1}{B} \sum_{b=1}^B h_b(z), \quad B = 200 \quad (3)$$

3.5 Integration of Crop Recommendation and Yield Prediction

The two trained models were integrated sequentially to transform soil and climatic measurements into a crop recommendation and a corresponding productivity estimate. The recommendation model first predicts the crop from the seven environmental variables. Because the naming conventions of the two datasets are not identical, a deterministic crop-name harmonization layer was introduced before the predicted crop was passed to the yield model. The mapping resolves semantically equivalent labels such as chickpea to Gram, mungbean to Moong (Green Gram), pigeonpeas to Arhar/Tur, and cotton to Cotton(lint). Using this procedure, 18 of the 22 recommendation crops could be associated with corresponding categories in the historical production data. Crops without a reliable equivalent were not forwarded to the yield stage, preventing forced mapping to unrelated categories. After harmonization, the mapped crop and contextual information are encoded using the same transformations applied during model development. The resulting input, consisting of state, district, crop year, season, crop, and area, is supplied to the Random Forest Regressor and the expected yield is obtained using Eq. (3). The predicted yield can subsequently be converted into estimated total production using Eq. (4). In Eq. (4), the left-hand term represents estimated production, while the predicted yield is multiplied by cultivated area A . Thus, the integrated framework provides a sequential output comprising a recommended crop, expected yield, and corresponding production estimate.

$$\hat{P} = \hat{Y} \times A \quad (4)$$

3.6 Statistical Validation of Model Performance

Metric-based model comparison was complemented by non-parametric statistical testing at a significance level of 0.05. For yield prediction, the absolute errors generated by Linear Regression, Decision Tree Regression, and Random Forest Regression were compared using the Friedman test. When the global test indicated a statistically significant difference, pairwise Wilcoxon signed-rank tests were applied to the corresponding error vectors. Holm adjustment was used for the resulting multiple comparisons to control the family-wise error rate. For crop recommendation, each classifier prediction was converted into a sample-wise correct or incorrect outcome. Cochran's Q test was used to examine whether the five classifiers differed in their probability of correct prediction over the same test observations. Pairwise comparisons were subsequently carried out using the exact McNemar test, again followed by Holm correction. This statistical procedure was used to determine whether observed differences among competing algorithms were supported beyond descriptive performance measures alone.

3.7 Model Explainability Using SHAP

To improve interpretability of the final tree-based models, SHapley Additive exPlanations (SHAP) were employed. SHAP decomposes a model output into a baseline prediction and additive feature-level contributions. For an input x , the explanation is represented by Eq. (5).

$$f(x) = \phi_0 + \sum_{j=1}^d \phi_j \quad (5)$$

In Eq. (5), the baseline term denotes the expected model output and each feature-specific SHAP term represents the contribution of the corresponding predictor to the prediction. For crop recommendation, SHAP values were computed for a reproducibly selected subset of test observations to quantify the influence of N, P, K, temperature, humidity, pH, and rainfall. Global importance was obtained from the mean absolute SHAP magnitude, while a local waterfall explanation was used to identify the variables supporting or opposing an individual predicted crop. The same explainability strategy was applied to the Random Forest yield model to evaluate the contribution of State, District, Crop Year, Season, Crop, and Area. The SHAP analysis therefore provides both global insight into model behaviour and local explanations for individual predictions without changing the underlying prediction models.

3.8 Performance Evaluation Metrics

The two components of the proposed framework were evaluated using measures appropriate to their respective learning tasks. Since crop recommendation is a multiclass classification problem, performance was assessed using accuracy, macro-precision, macro-recall, macro-F1-score, area under the receiver operating characteristic curve (AUC), and the confusion matrix. Accuracy measures the proportion of correctly classified samples, whereas precision and recall quantify prediction reliability and class-detection capability. The F1-score summarizes precision and recall through their harmonic mean. The classification measures of accuracy, precision, recall, and F1-score are summarized in Eq. (6).

$$\begin{aligned} \text{Accuracy} &= \frac{1}{n} \sum_{i=1}^n I(\hat{c}_i = c_i), & \text{Precision}_k &= \frac{TP_k}{TP_k + FP_k} \\ \text{Recall}_k &= \frac{TP_k}{TP_k + FN_k}, & F1_k &= \frac{2 \text{Precision}_k \text{Recall}_k}{\text{Precision}_k + \text{Recall}_k} \end{aligned} \quad (6)$$

In Eq. (6), n is the number of test observations and TP_k , FP_k , and FN_k denote the true-positive, false-positive, and false-negative counts for crop class k . Because the recommendation task contains 22 classes, precision, recall, and F1-score were macro-averaged so that every crop contributed equally to the aggregate performance. AUC was additionally reported to assess the overall discriminatory capability of the classification models, while a multiclass confusion matrix was used to inspect class-wise correct predictions and misclassification patterns. For yield prediction, the coefficient of determination (R-squared), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) were used. MAE measures the mean absolute deviation between observed and predicted yield values, while RMSE assigns greater influence to larger errors. R-squared measures the proportion of variation in the observed yield explained by the regression model. The regression criteria are given in Eq. (7).

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|, \quad RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (7)$$

In Eq. (7), the numerator contains the residual sum of squares between observed and predicted yields, the denominator contains the total sum of squares around the mean observed yield, and n denotes the number of test samples used in the accompanying MAE and RMSE expressions. Better regression performance is indicated by lower MAE and RMSE values together with a higher R-squared value. These metrics were used to compare Linear Regression, Decision Tree Regression, and Random Forest Regression, while the classification measures in Eq. (6), together with AUC and confusion matrix analysis, were used to compare the five crop recommendation algorithms. The statistical tests described in Section 3.6 were used as a complementary assessment of whether the measured model differences were statistically supported.

4. Results and Validation

4.1 Crop recommendation performance

The comparative classification results in Table 2 show that the seven soil and climatic attributes provide strong discriminatory information for separating the 22 crop classes, since all five classifiers achieved more than 97% accuracy. Random Forest and Naive Bayes achieved the highest accuracy of 99.55%, with only 2 errors among 440 test samples, indicating that the crop classes in the feature space are highly separable. However, the behaviour of the models provides additional insight beyond accuracy alone. The strong performance of Naive Bayes suggests that individual crop classes form relatively compact distributions with respect to the selected soil and climatic variables, whereas the comparable performance of Random Forest indicates that nonlinear feature interactions are also important. Support Vector Machine achieved 98.86% accuracy, while the lower performance of Logistic Regression at 97.27% suggests that purely linear decision boundaries are insufficient to describe all crop class relationships. Random Forest achieved the highest macro F1 score of 99.55% and an area under the curve of 1.000, supporting its selection for the integrated framework. The class wise results in Table 3 further show that 18 of the 22 crops were classified perfectly, while only blackgram and rice had reduced recall of 95.00% because one sample from each class was misclassified. The confusion of blackgram with maize and rice with jute indicates limited overlap in their environmental feature distributions rather than a broad failure of the classifier. Thus, the results suggest that the recommendation problem is well represented by the selected seven variables, with only a small number of crops exhibiting overlapping environmental characteristics.

Table 2. Performance comparison of classification models for crop recommendation

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	AUC	Misclassified (of 440)
Decision Tree	97.95	98.06	97.95	97.94	0.989	9
Naive Bayes	99.55	99.59	99.55	99.54	1.000	2
Support Vector Machine	98.86	98.96	98.86	98.87	1.000	5

Logistic Regression	97.27	97.40	97.27	97.25	1.000	12
Random Forest	99.55	99.57	99.55	99.55	1.000	2

Table 3. Class wise performance of the Random Forest model for crops with misclassifications

Crop	Precision (%)	Recall (%)	F1 Score (%)	Misclassified as
Blackgram	100.00	95.00	97.44	Maize (1)
Rice	100.00	95.00	97.44	Jute (1)
Jute	95.24	100.00	97.56	None
Maize	95.24	100.00	97.56	None
Remaining 18 crops	100.00	100.00	100.00	None
Macro average	99.57	99.55	99.55	

4.2 Yield prediction performance

The yield prediction stage was evaluated to determine how effectively the selected models could estimate crop yield from state, district, crop year, season, crop type, and cultivated area. The results presented in Table 4 show a clear difference between linear and tree-based approaches. Random Forest Regression produced the best overall performance, with a coefficient of determination of 0.948, meaning that approximately 94.8% of the variation in crop yield was explained by the model. It also produced the lowest mean absolute error of 0.703, and root mean squared error of 2.643, indicating that its predictions remained closer to the actual yield values than those of the other models. Decision Tree Regression also performed well, explaining 91.5% of the yield variation, but its higher mean absolute error of 0.823 and root mean squared error of 3.371 indicate less stable predictions, particularly for records with larger deviations. Random Forest improves on a single tree by averaging the outputs of multiple trees, which reduces model variance and limits the effect of individual splits. In contrast, Linear Regression explained only 9.4% of the observed yield variation and produced much larger errors, with a mean absolute error of 5.268 and root mean squared error of 10.986. This weak performance indicates that yield does not follow a simple linear relationship with the selected predictors. Crop yield varies according to interacting factors such as crop type, season, geographical location, year, and cultivated area; therefore, a nonlinear model is required to capture these dependencies. The comparison in Table 4 was performed to identify the model that not only explains the largest proportion of yield variation but also maintains low prediction error, and the results support the selection of Random Forest Regression for the yield prediction component of the proposed framework.

Table 4. Performance comparison of regression models for crop yield prediction

Model	R ²	MAE (t/ha)	RMSE (t/ha)
Linear Regression	0.094	5.268	10.986

Decision Tree	0.915	0.823	3.371
Random Forest	0.948	0.703	2.643

4.3 Feature importance and global explanations

Feature importance analysis was performed to identify the variables that contributed most strongly to the decisions of the selected Random Forest models and to determine whether the learned relationships were meaningful from an agricultural perspective. The results are summarized in Table 5 using two complementary approaches: impurity based importance and mean absolute Shapley values. For crop recommendation, impurity based importance ranked rainfall first with a value of 0.220, followed by humidity at 0.217 and potassium at 0.181. Shapley analysis produced a slightly different order, ranking humidity first with a mean absolute contribution of 0.0318, followed by potassium at 0.0278 and rainfall at 0.0262. The difference in ranking is expected because impurity based importance measures how effectively a feature reduces uncertainty while constructing the trees, whereas Shapley values measure how strongly a feature changes the final prediction. Despite this difference, both methods consistently identify humidity, potassium, and rainfall as the dominant variables for crop selection. This indicates that moisture related conditions and soil nutrient status provide the strongest separation among the 22 crop classes, while temperature and pH contribute comparatively less once the other environmental variables are considered. The agreement between the two explanation methods strengthens the interpretation that the Random Forest model is using agriculturally relevant patterns rather than relying on a single dominant variable.

For yield prediction, Table 5 shows a stronger contrast between the two importance measures and provides further insight into how the regression model operates. Crop type received the highest impurity-based importance of 0.634, considerably higher than state at 0.145 and season at 0.106. This indicates that crop identity forms the main structural basis of the Random Forest model, since different crops naturally exhibit different productivity ranges and therefore provide highly informative splits during tree construction. However, Shapley analysis ranked season first with a mean absolute contribution of 2.9525, followed by crop at 2.3212 and state at 1.0296. This means that although crop type is used most strongly to organize the decision structure of the forest, season produces the largest changes in individual yield predictions relative to the model baseline. Such behaviour is technically meaningful because the same crop may produce substantially different yields under different seasonal conditions due to changes in rainfall availability, temperature, cultivation period, and associated growing conditions. State also remained important under both measures, indicating that regional differences contribute significantly to yield variation. In comparison, crop year, cultivated areas, and district showed lower importance. The relatively limited contribution of cultivated area is expected because the target is yield rather than total production and increasing cultivated area does not necessarily increase productivity per unit area. Similarly, the lower importance of district suggests that much of the spatial variation is already captured by the broader state level together with crop and season. Therefore, the difference between impurity based and Shapley rankings should not be considered contradictory; instead, it shows that crop type is central to how the trees partition the data, while season has the strongest effect on the magnitude of individual yield predictions. Taken together, the results in Table 5 indicate that the yield model is primarily driven by the combined influence of crop type, season, and state, which also helps explain the strong performance of the nonlinear Random Forest model.

Table 5. Feature importance of the Random Forest models using impurity based (MDI) and SHAP measures

Feature	MDI Importance	MDI Rank	Mean SHAP	SHAP Rank
(a) Crop recommendation				
Rainfall	0.220	1	0.0262	3
Humidity	0.217	2	0.0318	1

Potassium (K)	0.181	3	0.0278	2
Phosphorus (P)	0.151	4	0.0244	5
Nitrogen (N)	0.103	5	0.0258	4
Temperature	0.075	6	0.0105	6
pH	0.052	7	0.0049	7
(b) Yield prediction				
Crop	0.634	1	2.3212	2
State	0.145	2	1.0296	3
Season	0.106	3	2.9525	1
Area	0.056	4	0.2807	5
Crop Year	0.033	5	0.4735	4
District	0.027	6	0.1688	6

4.4 Local explanation of an individual recommendation

A local Shapley explanation was generated to determine how each input feature influenced an individual crop recommendation rather than relying only on the final predicted class. The representative case presented in Table 6 was predicted as rice with a probability of 0.935, indicating strong model confidence for the given soil and climatic conditions. Rainfall of 202.94 mm produced the largest positive contribution of +0.2133, making it the strongest factor supporting the rice prediction. Nitrogen at 90 contributed +0.1724, followed by humidity of 82.00% with +0.1603 and potassium at 43 with +0.1472. Temperature and phosphorus also supported the prediction with contributions of +0.1070 and +0.0962, respectively. In contrast, pH of 6.50 produced a small negative contribution of -0.0075, indicating that it slightly reduced the predicted probability but had little influence on the final decision. The distribution of contributions in Table 6 shows that the recommendation was not generated from a single dominant attribute; instead, the model combined several favourable soil and climatic conditions to arrive at the rice prediction. The substantially larger positive contributions of rainfall, nitrogen, humidity, and potassium also explain why the small opposing effect of pH did not change the predicted crop. This local analysis was performed to make an individual recommendation interpretable by showing both the direction and magnitude of each feature contribution, allowing the final crop prediction to be traced back to the specific input conditions responsible for it.

Table 6. Local SHAP explanation for a representative input predicted as rice (probability 0.935)

Feature	Input Value	SHAP Value	Effect on Prediction
---------	-------------	------------	----------------------

Rainfall (mm)	202.94	+0.2133	Supports
Nitrogen (N)	90.00	+0.1724	Supports
Humidity (%)	82.00	+0.1603	Supports
Potassium (K)	43.00	+0.1472	Supports
Temperature (°C)	20.88	+0.1070	Supports
Phosphorus (P)	42.00	+0.0962	Supports
pH	6.50	-0.0075	Opposes

4.5 Integrated Crop Recommendation and Yield Prediction Using Random Forest

The integrated framework was designed to connect crop recommendation directly with yield and production estimation so that the final output provides more information than a crop label alone. The recommendation stage first uses nitrogen, phosphorus, potassium, temperature, humidity, pH, and rainfall to identify the most suitable crop. Because the crop names used in the recommendation and production datasets are not always identical, a crop harmonization step was introduced before yield estimation. Eighteen of the 22 recommendation crops could be matched with corresponding categories in the production dataset, including mappings such as chickpea to Gram, lentil to Masoor, mungbean to Moong(Green Gram), pigeonpeas to Arhar/Tur, and cotton to Cotton(lint). Crops without a reliable equivalent were not passed to the yield prediction stage, which avoids generating unsupported estimates. After a valid crop match is obtained, the recommended crop is combined with state, district, crop year, season, and cultivated area to estimate yield. The representative case presented in Table 7 illustrates the complete operation of the framework. For soil and climatic conditions consisting of nitrogen 90, phosphorus 42, potassium 43, temperature 20.88 degrees Celsius, humidity 82.00%, pH 6.50, and rainfall 202.94 mm, the model recommended rice. When this crop was considered for Nicobars district in the Andaman and Nicobar Islands during the Kharif season of 2000 with a cultivated area of 102 ha, the framework predicted a yield of 2.696 and an estimated production of 274.996 t. The corresponding historical record reported a production of 321 t, equivalent to a yield of approximately 3.147, giving a difference of about 14.3% between estimated and recorded production. This result shows that the framework can transform field level soil and climatic conditions into both a crop recommendation and a quantitative estimate of expected output. It also shows that a highly suitable crop recommendation does not necessarily produce an exact yield estimate, since historical yield is additionally affected by factors such as local weather variation, irrigation, management practices, soil condition, and other field level influences that are not represented in the available production attributes. The integrated output therefore provides a more practical form of agricultural decision support by linking crop suitability with the expected productivity of that crop under a specific location, season, year, and cultivated area. Additionally, the framework demonstrates that recommendation accuracy and yield estimation should be evaluated together when assessing practical agricultural usefulness. This combined analysis provides a clearer basis for selecting a crop while also estimating the production that may be expected under the specified cultivation conditions.

Table 7. Output of the integrated dual model framework for a sample case

Stage	Input	Output
Crop recommendation	N 90, P 42, K 43, 20.88 °C, 82.00 % humidity, pH 6.50, 202.94 mm rainfall	Rice

Yield prediction	Nicobars, Andaman and Nicobar Islands, Kharif season, 2000, area 102 ha	2.696 t/ha
Production estimate	Predicted yield × area	274.996 t

4.6 Statistical significance analysis

Statistical significance analysis was performed to verify whether the observed differences among the evaluated models were consistent across the test samples. As reported in Table 8, Cochran's Q test for crop recommendation produced $Q = 18.571$ with a probability value of 0.00095, confirming a significant overall difference among the five classifiers. For yield prediction, the Friedman test produced a statistic of 54179.65 with a probability value below 0.001, indicating a significant difference among the three regression models. Pairwise Wilcoxon signed rank tests with Holm correction further confirmed significant differences between Random Forest and Linear Regression, with a statistic of 26,012,666, and between Random Forest and Decision Tree Regression, with a statistic of 514,554,460, both with probability values below 0.001. These results show that the performance differences are not due to random variation and provide statistical support for selecting Random Forest Regression as the final yield prediction model.

Table 8. Statistical significance tests for differences among models

Task	Test	Comparison	Test Statistic	p value	Significant ($\alpha = 0.05$)
Crop recommendation	Cochran's Q	All 5 classifiers	$Q = 18.571$	0.00095	Yes
Yield prediction	Friedman	All 3 regressors	$\chi^2 = 54179.65$	< 0.001	Yes
Yield prediction	Wilcoxon (Holm)	Random Forest vs Linear Regression	$W = 26,012,666$	< 0.001	Yes
Yield prediction	Wilcoxon (Holm)	Random Forest vs Decision Tree	$W = 514,554,460$	< 0.001	Yes

4.7 Comparative Analysis with State of the Art Methods

The comparison in Table 9 indicates that the proposed framework achieves crop recommendation performance comparable with the strongest existing methods while using a comparatively simpler modelling strategy. The proposed Random Forest classifier obtained 99.55% accuracy and 99.55% F1 score, which is effectively on par with BORS CRP [30] at 99.57% and the stacking and voting approach of Ahmed et al. [28] at 99.56%. On a test set of 440 samples, one additional error changes the accuracy by approximately 0.23 percentage points; therefore, differences of only 0.01 to 0.02 percentage points are not practically meaningful. More importantly, the proposed model reaches this level of performance without Bayesian optimized stacking [30], ensemble fusion [28], synthetic oversampling [29], or metaheuristic optimization [31]. It also improves upon the accuracies reported by Apat et al. [29], Shastri et al. [21], Turgut et al. [32], and Srinivasu et al. [31], showing that a single Random Forest can provide highly competitive recommendation performance when the input soil and climatic variables are sufficiently discriminative. For yield prediction, the proposed model explains 94.8% of the observed yield variation, slightly higher than the 94.2% reported by XAI CROP [22]; however, this comparison should be interpreted with caution because the datasets and evaluation protocols are different. The stronger contribution of the present work is therefore not a marginal numerical improvement, but the integration of crop recommendation, yield prediction, production estimation, and explainability within one framework. Unlike most compared studies that stop at crop selection, the proposed approach extends the recommendation into an estimate of expected agricultural output and provides both global and local Shapley based explanations, making the resulting predictions more informative for practical decision support.

Table 9. Comparison of the proposed framework with existing studies

Study	Dataset	Method	Crop Recommendation	Yield Prediction	XAI Technique	Accuracy (%)	F1 Score (%)	R ²
Apat et al. [29], <i>J. Sci. Ind. Res.</i>	Kaggle crop recommendation (22 crops)	CatBoost with SMOTE	Yes	No	None	99.51	99.16	NR
Srinivasu et al. [31], <i>Computational Intelligence</i>	Soil and environmental data (2,300 samples)	RBF network with Spider Monkey Optimization	Yes	No	SHAP	98.20	NR	NR
Turgut et al. (AgroXAI) [32], <i>IEEE BigData</i>	Crop recommendation data	Random Forest	Yes	No	SHAP, LIME, counterfactual, ELI5	99.24	99.23	NR
XAI CROP [22], <i>Neural Computing and Applications</i>	Crop yield data	XAI CROP model	Yes	Yes	Explainable model	NR	NR	0.942
Shastri et al. [21], <i>Scientific Reports</i>	Kaggle crop recommendation (22 crops)	Gradient Boosting	Yes	No	LIME	99.27	99.32	NR
Ahmed et al. [28], <i>Scientific Reports</i>	Kaggle crop recommendation (2,200 samples) and FAO yield data	Stacking and voting ensemble with feature fusion	Yes	No	Feature importance	99.56	NR	NR
BORS-CRP [30], <i>Quality & Quantity</i>	Public crop recommendation data (22 crops)	Bayesian optimized stacking	Yes	No	None	99.57	99.57	NR
Proposed	Kaggle crop recommendation (2,200 samples) and Indian district crop production (240,505 records)	Random Forest classifier and Random Forest regressor	Yes	Yes	SHAP (global and local)	99.55	99.55	0.948

5 Conclusion and Future Directions

This study developed an integrated machine learning framework for crop recommendation and crop yield prediction. Five classifiers were evaluated for crop recommendation, while three regression models were compared for yield estimation. The Random Forest classifier achieved an accuracy of 99.55 percent, precision of 99.57 percent, recall of 99.55 percent, and macro F1 score of 99.55 percent. For yield prediction, Random Forest Regression achieved a coefficient of determination of 0.9475, mean absolute error of 0.7033, and root mean squared error of 2.6432. Statistical tests further confirmed significant overall differences among the evaluated models. The proposed framework connects crop recommendation with yield estimation through crop name harmonization. Explainability analysis showed that humidity contributed most strongly to crop recommendation, while season was the most influential variable for yield prediction. Thus, the framework provides both predictive performance and interpretable feature level information.

Future work can focus on incorporating recent production records, additional soil and climatic variables, and remote sensing information to improve yield estimation. The framework should also be validated on independent regional and temporal datasets to assess generalization. Standardized crop naming across datasets can further improve integration between recommendation and yield prediction. Prediction intervals and uncertainty estimation may also be incorporated to provide more reliable decision support under varying agricultural conditions.

References

- [1] Z. Zhai, J. F. Martínez, V. Beltran, and N. L. Martínez, "Decision support systems for agriculture 4.0: Survey and challenges," *Computers and Electronics in Agriculture*, vol. 170, Art. no. 105256, 2020.
- [2] S. M. Cheema and I. M. Pires, "AloT based soil nutrient analysis and recommendation system for crops using machine learning," *Smart Agricultural Technology*, vol. 11, Art. no. 100924, 2025, doi: 10.1016/j.atech.2025.100924.
- [3] P. Kaur, D. Gupta, S. Juneja, M. Uppal, and P. Gupta, "Data-driven crop recommendation framework using machine learning models and IoT-enabled sensing," *Procedia Computer Science*, vol. 282, pp. 168–178, 2026.
- [4] M. S. B. Alam, V. Esichaikul, A. Lameesa, S. F. Ahmed, and A. H. Gandomi, "An approach for crop recommendation with uncertainty quantification based on machine learning for sustainable agricultural decision-making," *Results in Engineering*, vol. 26, Art. no. 105505, 2025, doi: 10.1016/j.rineng.2025.105505.
- [5] B. Dey, J. Ferdous, and R. Ahmed, "Machine learning based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables," *Heliyon*, vol. 10, Art. no. e25112, 2024, doi: 10.1016/j.heliyon.2024.e25112.
- [6] R. Aworka, L. S. Cedric, W. Y. H. Adoni, J. T. Zoueu, F. K. Mutombo, C. L. M. Kimpolo, T. Nahhal, and M. Krichen, "Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries," *Smart Agricultural Technology*, vol. 2, Art. no. 100048, 2022, doi: 10.1016/j.atech.2022.100048.
- [7] M. A. Razavi, A. P. Nejadhashemi, B. Majidi, H. S. Razavi, J. Kpodo, R. Eeswaran, I. Ciampitti, and P. V. V. Prasad, "Enhancing crop yield prediction in Senegal using advanced machine learning techniques and synthetic data," *Artificial Intelligence in Agriculture*, vol. 14, pp. 99–114, 2024.
- [8] B. S. Omary, M. A. Dida, and D. G. Nyambo, "Farm-level yield prediction for maize, rice, and beans in Tanzania using machine learning and multi-source agricultural data," *Smart Agricultural Technology*, vol. 13, Art. no. 101904, 2026, doi: 10.1016/j.atech.2026.101904.
- [9] N. Mahdizadeh Gharakhanlou and L. Perez, "From data to harvest: Leveraging ensemble machine learning for enhanced crop yield predictions across Canada amidst climate change," *Science of the Total Environment*, vol. 951, Art. no. 175764, 2024, doi: 10.1016/j.scitotenv.2024.175764.
- [10] S. Fatholouloumi, A. Gillespie, and A. Biswas, "Enhancing crop yield predictions through spectral-temporal integration across multiple agricultural fields," *Computers and Electronics in Agriculture*, vol. 245, Art. no. 111552, 2026, doi: 10.1016/j.compag.2026.111552.
- [11] D. Han, P. Wang, Y. Zhang, Y. Li, J. Tang, Q. Li, Q. Wang, Y. Ma, and D. Wu, "Integrating a process-based crop model with machine learning for early-season summer maize yield prediction in the Huang-Huai-Hai Plain," *Field Crops Research*, vol. 349, Art. no. 110730, 2026.
- [12] F. Mena, D. Pathak, H. Najjar, C. Sanchez, P. Helber, B. Bischke, P. Habelitz, M. Miranda, J. Siddamsetty, M. Nuske, M. Charfuelan, D. Arenas, M. Vollmer, and A. Dengel, "Adaptive fusion of multi-modal remote sensing data for optimal sub-field crop yield prediction," *Remote Sensing of Environment*, vol. 318, Art. no. 114547, 2025, doi: 10.1016/j.rse.2024.114547.
- [13] A. D. Yewle, L. Mirzayeva, and O. Karakuş, "Multi-modal data fusion and deep ensemble learning for accurate crop yield prediction," *Remote Sensing Applications: Society and Environment*, vol. 38, Art. no. 101613, 2025, doi: 10.1016/j.rsase.2025.101613.

- [14] M. S. Mia, R. Tanabe, L. N. Habibi, N. Hashimoto, K. Homma, M. Maki, T. Matsui, and T. S. T. Tanaka, "Multimodal deep learning for rice yield prediction using UAV-based multispectral imagery and weather data," *Remote Sensing*, vol. 15, no. 10, Art. no. 2511, 2023, doi: 10.3390/rs15102511.
- [15] M. Maimaitijiang, V. Sagan, P. Sidike, S. Hartling, F. Esposito, and F. B. Fritsch, "Soybean yield prediction from UAV using multimodal data fusion and deep learning," *Remote Sensing of Environment*, vol. 237, Art. no. 111599, 2020, doi: 10.1016/j.rse.2019.111599.
- [16] G. Li, B. Zhou, M. Ni, H. Chen, Y. Liu, Y. Zhang, M. Zhang, and M. Wang, "Stages-based multimodal data fusion model (S-MDFM) for wheat yield prediction and screening of drought-resistant and high-yield varieties," *Computers and Electronics in Agriculture*, vol. 239, Art. no. 111122, 2025, doi: 10.1016/j.compag.2025.111122.
- [17] M. Qiao, X. He, X. Cheng, P. Li, Q. Zhao, C. Zhao, and Z. Tian, "KSTAGE: A knowledge-guided spatial-temporal attention graph learning network for crop yield prediction," *Information Sciences*, vol. 619, pp. 19–37, 2023, doi: 10.1016/j.ins.2022.10.112.
- [18] J. Yang, L. Liu, Q. Yang, X. Jia, B. Peng, K. Guan, and Z. Jin, "Knowledge-guided graph machine learning improves corn yield mapping in the U.S. Midwest," *Remote Sensing of Environment*, vol. 335, Art. no. 115287, 2026, doi: 10.1016/j.rse.2026.115287.
- [19] S. Sarkar, A. Dey, R. Pradhan, U. M. Sarkar, C. Chatterjee, A. Mondal, and P. Mitra, "Crop yield prediction using multimodal Meta-Transformer and temporal graph neural networks," *IEEE Transactions on AgriFood Electronics*, vol. 2, no. 2, 2024.
- [20] Y. Gamidell and M. A. Chaurasia, "Explainable multi season spatio temporal deep learning framework for crop yield forecasting using Sentinel-2 remote sensing data," *Discover Computing*, vol. 29, Art. no. 452, 2026, doi: 10.1007/s10791-026-10375-8.
- [21] 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, Art. no. 25498, 2025, doi: 10.1038/s41598-025-07003-8.
- [22] 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, doi: 10.1007/s00521-023-09391-2.
- [23] A. Yaganteeswarudu, S. K. Biswas, V. Aruna, and D. Tripathi, "Enhancing transparency in smart farming: Local explanations for crop recommendations using LIME," *Procedia Computer Science*, vol. 258, pp. 1993–2005, 2025, doi: 10.1016/j.procs.2025.04.450.
- [24] A. Badshah, B. Y. Alkazemi, F. Din, K. Z. Zamli, and M. Haris, "Crop classification and yield prediction using robust machine learning models for agricultural sustainability," *IEEE Access*, vol. 12, 2024, doi: 10.1109/ACCESS.2024.3486653.
- [25] J. Madhuri, B. R. Reddy, G. Kalnoor, M. Indiramma, and N. Nagarathna, "Optimizing crop recommendations with improved deep belief networks: A multimodal approach," *IEEE Access*, vol. 13, pp. 31762–31773, 2025, doi: 10.1109/ACCESS.2025.3542284.
- [26] A. Ashwitha and C. A. Latha, "Crop recommendation and yield estimation using machine learning," *Journal of Mobile Multimedia*, vol. 18, no. 3, pp. 861–884, 2022, doi: 10.13052/jmm1550-4646.18320.
- [27] N. Piyar S and S. Baulkani, "An explainable spatio-temporal deep learning framework for crop yield prediction and recommendation," *Scientific Reports*, 2026, doi: 10.1038/s41598-026-54439-7.
- [28] R. A. Ahmed, W. El-Shafai, Z. A. Ahmed, E.-S. M. El-Rabaie, and F. E. Abd El-Samie, "High-precision crop recommendation system with stacking ensemble classifiers for optimizing agricultural productivity," *Scientific Reports*, vol. 15, Art. no. 43377, 2025, doi: 10.1038/s41598-025-09640-5.
- [29] 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*, vol. 82, no. 5, pp. 558–567, 2023, doi: 10.56042/jsir.v82i05.1092.
- [30] S. Chandrakar, M. Pandey, and R. K. Jade, "BORS-CRP: A Bayesian optimized stacking framework for crop recommendation in precision agriculture," *Quality & Quantity*, vol. 60, no. 4, pp. 13665–13689, 2026, doi: 10.1007/s11135-026-02742-w.
- [31] P. Naga Srinivasu, M. F. Ijaz, and M. Woźniak, "XAI-driven model for crop recommender system for use in precision agriculture," *Computational Intelligence*, vol. 40, no. 1, Art. no. e12629, 2024, doi: 10.1111/coin.12629.
- [32] O. Turgut, I. Kok, and S. Ozdemir, "AgroXAI: Explainable AI-driven crop recommendation system for Agriculture 4.0," in *2024 IEEE International Conference on Big Data (BigData)*, IEEE, 2024, doi: 10.1109/BigData62323.2024.10825771.