Turkish Journal of Agriculture - Food Science and Technology, 14(1): 262-273, 2026
DOI: https://doi.org/10.24925/turjaf.v14i1.262-273.8336

Turkish Journal of Agriculture - Food Science and Technology

Available online, ISSN: 2148-127X | www.agrifoodscience.com | Turkish Science and Technology Publishing (TURSTEP)

Machine Learning-Based Dual Agricultural Decision Support System: An Integrated Approach for Crop and Fertilizer Recommendations

Hatice Bülbül[1,a], Enes Güler[1,b], Burcu Çarklı Yavuz[1,c,*]

[1] *Department of Information Systems Engineering, Faculty of Computer and Information Sciences, Sakarya University, 54187, Sakarya, Türkiye*
[*] *Corresponding author*

A R T I C L E   I N F O

Research Article

Received : 09.09.2025
Accepted : 01.01.2026

Keywords:
Machine learning
Agricultural decision support
Crop recommendation
Fertilizer management
Precision agriculture

A B S T R A C T

This study proposes a comprehensive dual-module agricultural decision support system that integrates machine learning algorithms for both crop and fertilizer recommendations. The system addresses critical challenges in modern agriculture by providing evidence-based recommendations to optimize productivity while maintaining sustainability. Four machine learning algorithms (Random Forest, Support Vector Machine, XGBoost, and K-Nearest Neighbors) were implemented and evaluated using comprehensive agricultural datasets containing 2,200 crop samples and 3,100 fertilizer samples. The crop recommendation module achieved 99.32% accuracy using Random Forest, while the fertilizer recommendation module attained 98.75% accuracy. The system incorporates advanced techniques including SMOTE for handling class imbalance, GridSearchCV for hyperparameter optimization, and SHAP analysis for model interpretability. Comparative analysis with existing literature demonstrates competitive performance while offering enhanced explainability and dual functionality. The developed framework provides a foundation for region-specific implementations and represents a practical contribution to intelligent agricultural decision support systems.

[a] b211200058@sakarya.edu.tr https://orcid.org/0009-0003-0334-6381
[c] bcarkli@sakarya.edu.tr https://orcid.org/0000-0003-4089-024X
[b] b211200043@sakarya.edu.tr https://orcid.org/0009-0004-1744-780X

Introduction

Global agriculture faces unprecedented challenges in the 21st century, driven by rapid population growth, climate change, and diminishing arable land resources. The world population is projected to reach 9.7 billion by 2050, necessitating a 70% increase in food production to meet growing demand (FAO, 2023). Simultaneously, agricultural land is decreasing at an alarming rate of 12 million hectares annually due to urbanization, soil degradation, and desertification (UNCCD, 2022). These challenges directly threaten the achievement of United Nations Sustainable Development Goals (SDGs), particularly SDG 2 (Zero Hunger), SDG 12 (Responsible Consumption and Production), and SDG 15 (Life on Land).

Traditional farming practices, while historically successful, are increasingly inadequate to address modern agricultural challenges. Farmers worldwide struggle with suboptimal crop selection and inefficient fertilizer management, leading to significant economic losses and environmental degradation. Studies indicate that inappropriate crop selection can result in 15-30% yield losses, while excessive or improper fertilizer application contributes to soil acidification, water pollution, and greenhouse gas emissions (Tanaka et al., 2024).

The complexity of agricultural decision-making involves multiple interacting factors such as soil properties, climate conditions, and market dynamics. This complexity necessitates sophisticated decision support tools that can contribute to sustainable agricultural transformation and climate-resilient food systems. The emergence of precision agriculture and smart farming technologies has opened new avenues for addressing these challenges through data-driven approaches.

Recent advances in deep learning architectures, particularly attention-based models such as TabNet and Transformer-based approaches, have demonstrated superior performance in handling tabular agricultural data (Venkateswara & Padmanaban, 2025; Khaliq et al., 2025). These architectures effectively capture complex non-linear relationships between soil properties, environmental conditions, and crop requirements while maintaining interpretability through attention mechanisms.

Decision support systems in agriculture have evolved from simple rule-based expert systems to sophisticated machine learning platforms capable of handling multi-dimensional data and providing real-time recommendations. Modern agricultural decision support systems integrate various data sources, including soil analysis, weather patterns, satellite imagery, and historical yield data, to provide comprehensive guidance to farmers (Prity et al., 2024). The integration of explainable artificial intelligence (XAI) techniques has further enhanced the trustworthiness and adoption of these systems by providing transparent reasoning behind recommendations (Cartolano et al., 2024).

This study addresses the critical need for integrated agricultural decision support systems by developing a dual-module intelligent system that provides both crop and fertilizer recommendations. The system leverages machine learning algorithms to analyze soil properties (pH, nitrogen, phosphorus, potassium), environmental conditions (temperature, humidity, rainfall), and other relevant factors to generate scientifically-based recommendations. While the current implementation utilizes globally available datasets for model development and validation, the framework is designed to be adaptable to region-specific data, including Turkish agricultural conditions in future implementations.

Literature Review

The application of machine learning techniques in agricultural decision support has gained significant momentum over the past decade, with numerous studies demonstrating the potential of these technologies to enhance farming practices. This literature review examines the current state of research in crop recommendation systems, fertilizer management, and integrated agricultural decision support platforms.

Early work in crop recommendation focused on rule-based expert systems, but recent developments have shifted toward machine learning approaches. Shastri et al. (2025) developed a comprehensive crop recommendation system using Gradient Boosting, achieving 99.27% accuracy with 99.32% precision, 99.36% recall, and 99.32% F1 score. Their study emphasized the importance of explainable AI techniques in agricultural applications, demonstrating how LIME (Local Interpretable Model-Agnostic Explanations) can provide detailed explanations for crop recommendations.

The integration of multiple machine learning algorithms has shown promising results in recent studies. Prity et al. (2024) conducted a comprehensive evaluation of nine machine learning models including Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Decision Tree, Random Forest, Bagging, AdaBoost, Gradient Boosting, and Extra Trees for crop recommendation. Their Random Forest model achieved the highest accuracy of 99.31%, demonstrating the effectiveness of ensemble methods in agricultural applications.

Regional adaptation of machine learning models has been explored in various geographical contexts. Demir et al. (2025) conducted a comprehensive study on irrigation status prediction in Turkey's Southeastern Anatolia Region, comparing Random Forest (95%), Decision Tree (97%), Gradient Boosting (93%), and Artificial Neural Network (98%) algorithms. Their findings demonstrated that ANN and Decision Trees performed best in provinces with large agricultural areas such as Şanlıurfa and Diyarbakır, highlighting the importance of regional adaptation in agricultural decision support systems. The study utilized the "Weather Data" dataset from Kaggle and integrated Open Weather Map API for real-time data collection, demonstrating the practical applicability of machine learning algorithms in Turkish agricultural contexts.

Recent developments in explainable artificial intelligence have addressed the critical issue of model interpretability in agricultural applications. Cartolano et al. (2024) conducted a comprehensive analysis of explainable AI models for smart agriculture environments, utilizing both SHAP (SHapley Additive ExPlanations) and LIME techniques to provide transparent reasoning behind crop recommendations. This work highlighted the importance of model explainability in building farmer trust and facilitating technology adoption.

The emergence of attention-based deep learning architectures has introduced new possibilities for agricultural decision support systems. Venkateswara and Padmanaban (2025) proposed a unified smart recommendation system utilizing TabNet, a deep learning architecture specifically designed for tabular data. Their system achieved 95.24% accuracy for fertilizer recommendations and 96.21% for crop recommendations on a Western Maharashtra dataset. The study employed SMOTE for class imbalance handling, iterative imputation for missing values, and SHAP for post-hoc interpretability, demonstrating the effectiveness of combining advanced preprocessing techniques with interpretable deep learning models.

Khaliq et al. (2025) developed an integrated AI-driven smart agriculture system that combines multiple transformer-based architectures for different agricultural tasks. Their system utilized a Transformer-based Tabular Learning (TTL) model for irrigation prediction (99.13% accuracy), a Sparse Weighted Fusion Transformer (SwiFT) for crop recommendation (98.75% accuracy), and a TabNet classifier for fertilizer recommendation (99.3% accuracy). The integration of explainable AI techniques (SHAP and LIME) with these advanced architectures provided transparent decision-making processes, enhancing farmer trust and system adoption.

Parallel to crop recommendation research, fertilizer management represents a critical component of sustainable agriculture, with significant implications for both productivity and environmental impact. Tanaka et al. (2024) conducted a comprehensive study on machine learning models for fertilizer recommendations using real on-farm precision experimental data. Their research demonstrated that while machine learning models can accurately predict yield, the uncertainty in fertilizer application rate recommendations remains significant, with coefficients of variation ranging from 13.3% to 31.5% for different fertilization strategies.

Comprehensive reviews of smart fertilizer management technologies have highlighted the integration of multiple advanced technologies. Liu et al. (2025) conducted a topical review on precision fertilization using Geographic Information Systems (GIS), Global Positioning System (GPS), Internet of Things (IoT), Artificial Intelligence (AI) and Machine Learning (ML) algorithms, and Unmanned Aerial Vehicles (UAVs) and Unmanned Ground Vehicles

(UGVs) for fertilizer management. Their review emphasized the importance of multi-sensor integration, open-access datasets, and eco-friendly AI-driven solutions for sustainable fertilizer management. The study identified key challenges including data quality, scalability constraints in low-resource settings, regulatory inconsistencies, and high costs of technology adoption for smallholder farmers, while exploring applications such as data-driven recommendations, aerial monitoring, and automated fertilizer application systems.

The application of SHAP analysis in agricultural contexts has gained considerable attention. Mohan et al. (2025) demonstrated the integration of AI and XAI techniques for precision crop yield predictions, achieving R^2 scores of 0.92 with mean squared errors as low as 0.02. Their study utilized SHAP and LIME techniques to enhance interpretability, providing actionable insights into the relative importance of key features such as temperature, rainfall patterns, and macronutrient levels.

Regional adaptation of agricultural recommendation systems has been explored in various contexts. Patel and Patel (2023) developed AgriRec, a multi-criteria agriculture recommendation system that achieved 95.85% accuracy for crop recommendation and 92.11% accuracy for fertilizer recommendation. Their study highlighted the importance of incorporating multiple criteria including soil properties, water level, farm size, and minimum support price in developing effective recommendation systems.

The incorporation of explainable AI techniques in agricultural applications has been systematically reviewed by Arrighi et al. (2025), who provided a comprehensive taxonomy for classifying food quality research using XAI techniques. Their survey emphasized the importance of SHAP, LIME, and other XAI methods in enhancing transparency and interpretability in agricultural decision-making systems.

Despite significant advances in agricultural decision support systems, several research gaps remain. Most existing systems focus on either crop recommendation or fertilizer management, with limited integration between these critical agricultural decisions. Additionally, many studies utilize datasets from specific regions without addressing the adaptability of models to different geographical and climatic conditions. The current study addresses these gaps by developing a dual-module system that incorporates machine learning techniques, proper handling of data imbalance, hyperparameter optimization, and explainable AI methods. The framework is designed to be region-adaptable while maintaining high performance across different agricultural contexts.

Materials and Methods

System Architecture and Workflow

The developed dual-module intelligent agricultural decision support system follows a comprehensive workflow consisting of five main stages: data collection, preprocessing, model training, prediction, and result presentation. The system architecture comprises user interface, data processing module, machine learning engine, and result visualization components designed with modularity and scalability in mind. The workflow begins with data input from farmers, including soil test results (pH, NPK values) and environmental conditions (temperature, humidity, rainfall). The preprocessing module validates and normalizes the input data, applying feature engineering techniques to create derived variables that enhance model performance. The SMOTE technique is applied during training to handle class imbalance issues commonly found in agricultural datasets.

Dataset Description and Characteristics

The study utilized two comprehensive agricultural datasets obtained from Kaggle, each specifically designed for different aspects of agricultural decision-making. These datasets provide the foundation for developing robust machine learning models capable of addressing real-world agricultural challenges.

The crop recommendation dataset, sourced from Kaggle (Nalluri, 2024), contains 2,200 samples with 7 features designed to predict the most suitable crop for given environmental and soil conditions. This dataset encompasses 22 different crop types across multiple agricultural categories: cereal grains (rice, maize), legumes (chickpea, kidneybeans, pigeonpeas, mothbeans, mungbean, blackgram, lentil), fruits (pomegranate, banana, mango, grapes, watermelon, muskmelon, apple, orange, papaya, coconut), and cash crops (cotton, jute, coffee). The detailed feature descriptions and their statistical summary of these are presented in Table 1 and Table 2 respectively.

The fertilizer recommendation dataset, obtained from Kaggle (Nishchalchandel, 2025), contains 3,100 samples with 12 variables designed for predicting optimal fertilizer types based on soil conditions and crop requirements. The dataset includes 8 numerical variables, 3 categorical variables, and 1 descriptive text variable. The target variable is the "Fertilizer" column, which contains the recommended fertilizer class according to environmental and crop type conditions. The complete variable descriptions are detailed in Table 3.

Data Preprocessing and Feature Engineering

Data preprocessing involved several critical steps to ensure optimal model performance. Initially, both datasets were split into training and testing sets using an 80:20 ratio with stratified sampling to maintain class distribution proportions. Label encoding was applied to convert categorical target variables into numerical format suitable for machine learning algorithms. Feature standardization was performed using StandardScaler to normalize all input features to zero mean and unit variance, ensuring equal contribution from all variables during model training. The standardization process follows the mathematical formulation

$$z = (x - \mu) / \sigma \quad (1)$$

where x represents the original feature value, μ is the mean, σ is the standard deviation, and z is the standardized value. Class imbalance was addressed using SMOTE (Synthetic Minority Oversampling Technique), which generates synthetic samples for minority classes based on k-nearest neighbors.

Table 1. Crop recommendation dataset features and descriptions

Feature	Unit	Description
Nitrogen (N)	mg/kg	Soil nitrogen content - Essential for leaf development and photosynthesis
Phosphorus (P)	mg/kg	Soil phosphorus content - Supports root development and flowering process
Potassium (K)	mg/kg	Soil potassium content - Provides cellular metabolism and disease resistance
Temperature	°C	Environmental temperature - Directly affects plant growth process
Humidity	%	Atmospheric humidity ratio - Affects plant transpiration and photosynthesis
pH	-	Soil pH value - Determines nutrient availability
Rainfall	mm	Precipitation amount - Determines water requirement and soil saturation

Table 2. Statistical summary of crop recommendation dataset features

Feature	Unit	Minimum	Maximum	Mean	Std. Deviation
Nitrogen (N)	mg/kg	0	140	50.2	36.9
Phosphorus (P)	mg/kg	5	145	53.4	32.1
Potassium (K)	mg/kg	5	205	48.1	50.6
Temperature	°C	8.8	43.7	25.6	5.1
Humidity	%	14.3	99.9	71.5	22.3
pH	-	3.5	9.9	6.5	0.8
Rainfall	mm	20.2	298.6	103.5	54.2

Table 3. Fertilizer recommendation dataset variables

Variable Name	Type	Description
Temperature (°C)	Numerical	Environmental temperature - Affects fertilizer solubility and plant metabolism
Moisture	Numerical	Soil moisture ratio (0-1 normalized) - Important for nitrogen dissolution and fertilizer transport
Rainfall (mm)	Numerical	Precipitation amount - Determines soil saturation and fertilizer absorption
pH	Numerical	Soil acidity-alkalinity degree - Has direct effect on nutrient uptake
Nitrogen (N)	Numerical	Soil nitrogen amount - Required for leaf development and photosynthesis
Phosphorous (P)	Numerical	Phosphorus level - Important for root development and energy metabolism
Potassium (K)	Numerical	Potassium level - Affects water balance and cell structure
Carbon (C)	Numerical	Organic carbon content - Accepted as indicator of soil fertility
Soil	Categorical	Soil type - Clayey, Sandy, Loamy etc. contains 5 different classes
Crop	Categorical	Crop type - Rice, Maize, Cotton etc. contains 31 different crops
Fertilizer	Categorical	Target variable - Recommended fertilizer type (e.g. DAP, Urea, NPK etc.)

The SMOTE algorithm (Chawla et al., 2002) creates new samples using the following approach:

$$\text{x\_synthetic} = \text{x\_i} + \lambda(\text{x\_neighbor} - \text{x\_i}) \qquad (2)$$

where x_i is a minority class sample, x_neighbor is a randomly selected sample from its k nearest neighbors, λ is a random number between 0 and 1, and x_synthetic is the generated synthetic sample.

Machine Learning Algorithms and Mathematical Formulations

Four machine learning algorithms were implemented and compared: Random Forest, Support Vector Machine (SVM), XGBoost, and K-Nearest Neighbors (KNN). Each algorithm was selected based on its proven effectiveness in agricultural classification tasks and complementary strengths in handling different data characteristics.

Random Forest (Breiman, 2001) combines multiple decision trees using bootstrap aggregating. For a dataset with n samples and m features, the algorithm creates B bootstrap samples and builds decision trees. The final prediction is determined by aggregating predictions from all trees:

$$\hat{y} = \text{Aggregate}\{T_1(x), T_2(x), ..., T_\beta(x)\} \qquad (3)$$

where $T_\beta(x)$ represents the prediction of the b-th tree for input x, and the aggregation function uses majority voting for classification tasks and averaging for regression tasks. Support Vector Machine (Cortes & Vapnik, 1995) finds the optimal hyperplane that separates classes with maximum margin. The optimization problem is formulated as:

$$\min(w,b,\xi)\ (1/2)||w||^2 + C \sum_{i=1}^{n} \xi_i \qquad (4)$$

$$\text{subject to: } y_i(w^T\varphi(x_i) + b) \geq 1 - \xi_i,\ \xi_i \geq 0$$

where w is the weight vector, b is the bias term, C is the regularization parameter controlling the trade-off between margin maximization and classification error, ξ_i are slack variables allowing misclassification, and $\varphi(x_i)$ is the kernel function mapping input features to a higher-dimensional space.

XGBoost (Extreme Gradient Boosting) (Chen & Guestrin, 2016) builds an ensemble of decision trees sequentially, where each tree corrects errors from previous iterations. The objective function balances prediction accuracy and model complexity:

$$\text{Obj} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{k} \Omega(f_k) \qquad (5)$$

where l is the loss function, Ω represents the regularization term preventing overfitting, and f_k denotes the k-th tree in the ensemble.

K-Nearest Neighbors (Cover & Hart, 1967) classifies samples based on the majority vote of k nearest neighbors. The Euclidean distance between two data points is calculated as:

$$d(x_i, x_j) = \sqrt{(\sum^{M}_{m=1}(x_{im} - x_{jm})^2)} \quad (6)$$

where x_i and x_j are two observation vectors, M is the number of features, and $d(x_i, x_j)$ represents the Euclidean distance between the two observations.

Hyperparameter Optimization

Systematic hyperparameter optimization was performed using GridSearchCV (Pedregosa et al., 2011) with 5-fold cross-validation to identify optimal parameter combinations for each algorithm. The optimization process follows the objective function:

$$\theta^* = \text{argmin}(\theta)\ (1/k) \sum^{k}_{i=1} L(f\theta(X_i^{val}), y_i^{val}) \quad (7)$$

where θ represents the hyperparameter set, L is the loss function, and (X_i^{val}, y_i^{val}) represents the validation data in the i-th fold of k-fold cross-validation. The hyperparameter ranges tested for each algorithm are detailed in Tables 4-7.

Model Evaluation Metrics

Model performance was evaluated using comprehensive classification metrics (Sokolova & Lapalme, 2009) including accuracy, precision, recall, and F1-score. These metrics provide different perspectives on model performance and are particularly important for imbalanced datasets (Hossin & Sulaiman, 2015). Accuracy measures the overall correctness of predictions and is calculated as:

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

Precision quantifies the proportion of positive predictions that are actually correct, indicating the model's ability to avoid false positives:

$$\text{Precision} = TP / (TP + FP) \quad (9)$$

Recall measures the proportion of actual positive cases that are correctly identified, representing the model's sensitivity to positive instances:

$$\text{Recall} = TP / (TP + FN) \quad (10)$$

F1-score provides a harmonic mean of precision and recall, offering a balanced measure that is particularly valuable for imbalanced datasets:

$$\text{F1} - \text{Score} = \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (11)$$

where TP represents True Positives (correctly predicted positive cases), TN represents True Negatives (correctly predicted negative cases), FP represents False Positives (incorrectly predicted positive cases), and FN represents False Negatives (incorrectly predicted negative cases). These metrics collectively provide comprehensive insight into model performance across different aspects of classification accuracy.

Table 4. Random Forest hyperparameter optimization ranges

Parameter	Tested Values	Description
n_estimators	[100, 200, 300]	Number of trees in the forest
max_depth	[10, 20, None]	Maximum depth of trees
min_samples_split	[2, 5, 10]	Minimum samples required to split node
min_samples_leaf	[1, 2, 4]	Minimum samples required at leaf node
class_weight	[None, balanced]	Weights associated with classes

Note: Hyperparameter optimization performed using GridSearchCV with 5-fold cross-validation.

Table 5. Support Vector Machine hyperparameter optimization ranges

Parameter	Tested Values	Description
C	[1, 10, 100]	Regularization parameter
gamma	[scale, auto]	Kernel coefficient for RBF
kernel	[rbf]	Kernel type for non-linear classification

Note: Hyperparameter optimization performed using GridSearchCV with 5-fold cross-validation.

Table 6. XGBoost hyperparameter optimization ranges

Parameter	Tested Values	Description
n_estimators	[100, 200]	Number of boosting rounds
max_depth	[3, 6, 10]	Maximum depth of trees
learning_rate	[0.01, 0.1]	Step size shrinkage
subsample	[0.7, 1.0]	Subsample ratio of training instances

Note: Hyperparameter optimization performed using GridSearchCV with 5-fold cross-validation.

Table 7. K-Nearest Neighbors hyperparameter optimization ranges

Parameter	Tested Values	Description
n_neighbors	[3, 5, 7]	Number of neighbors to consider
weights	[uniform, distance]	Weight function for neighbors
p	[1, 2]	Power parameter for Minkowski metric

Note: Hyperparameter optimization performed using GridSearchCV with 5-fold cross-validation.

Table 8. Crop Recommendation Dataset Model Performance Results

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	99.32	99.35	99.32	99.33
SVM	97.73	97.89	97.73	97.78
KNN	97.05	97.12	97.05	97.08
XGBoost	95.45	95.52	95.45	95.48

Table 9. Fertilizer Recommendation Dataset Model Performance Results

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	98.75	98.82	98.75	98.78
SVM	96.25	96.38	96.25	96.31
KNN	95.00	95.15	95.00	95.07
XGBoost	93.75	93.89	93.75	93.82

Model Interpretability and SHAP Analysis

Model interpretability was ensured through SHAP (SHapley Additive exPlanations) analysis (Lundberg & Lee, 2017), which provides transparent explanations for each recommendation. SHAP values quantify how much each input feature (soil properties, environmental conditions) contributes to the model's prediction, calculated using cooperative game theory:

$$\varphi_i = \sum_{S\subseteq\{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!}\left[f(S\cup\{i\}) - f(S)\right] \qquad (12)$$

where φ_i represents the contribution of feature i, F is the complete feature set, S represents feature subsets, and f(S) is the model prediction for subset S. This interpretability enables farmers to understand which soil and environmental factors drive each crop or fertilizer recommendation, increasing trust and facilitating informed decision-making.

Results

This section presents the detailed performance results of the developed agricultural decision support system and compares them with similar studies in the literature. Four different machine learning algorithms (Random Forest, SVM, KNN, and XGBoost) were evaluated using two different datasets (crop recommendation and fertilizer recommendation) with comprehensive performance metrics including accuracy, precision, recall, and F1-score.

The model performance results obtained after hyperparameter optimization are presented in Tables 8 and 9. Both datasets were evaluated using accuracy, precision, recall, and F1-score metrics to provide a comprehensive assessment.

These results are consistent with recent findings in agricultural deep learning applications. Venkateswara and Padmanaban (2025) reported similar performance levels using TabNet architecture (95.24% for fertilizer, 96.21% for crop), while Khaliq et al. (2025) achieved 99.3% accuracy for fertilizer recommendation using TabNet with XAI integration. The comparable performance across different architectures (ensemble methods vs. attention-based deep learning) suggests that both approaches are viable for agricultural decision support, with the choice depending on specific requirements such as interpretability, computational resources, and deployment constraints.

The crop recommendation results demonstrate exceptional performance across all evaluated algorithms, with Random Forest achieving the highest accuracy of 99.32%. The consistently high precision, recall, and F1-score values (all above 99.30%) indicate balanced performance across all 22 crop categories, suggesting effective handling of the multi-class classification challenge. SVM showed competitive performance with 97.73% accuracy, demonstrating the effectiveness of kernel-based methods for capturing non-linear relationships in agricultural data. KNN achieved 97.05% accuracy, while XGBoost exhibited 95.45% accuracy, ranking fourth among the evaluated algorithms.

The fertilizer recommendation module achieved similarly strong results, with Random Forest maintaining its superior performance at 98.75% accuracy. The performance gap between Random Forest and other algorithms is consistent with the crop recommendation results, with SVM achieving 96.25%, KNN 95.00%, and XGBoost 93.75%. The slightly lower overall accuracy values compared to crop recommendation can be attributed to the increased complexity of the fertilizer dataset, which includes both numerical and categorical variables, requiring more sophisticated handling of mixed data types.

Cross-dataset performance analysis reveals that Random Forest maintains consistent superiority across both agricultural domains, with an accuracy difference of only 0.57% between crop and fertilizer recommendations. This consistency demonstrates robust generalization capabilities and validates the algorithm's suitability for diverse agricultural prediction tasks. The performance ranking (Random Forest > SVM > KNN > XGBoost) remains stable across both datasets, indicating reliable algorithmic behavior regardless of agricultural domain complexity.

The literature comparison results show that the proposed system exhibits competitive performance for both datasets. The comparative analysis with existing studies in crop and fertilizer recommendation systems is presented in Tables 10 and 11.

The comparative analysis of crop recommendation systems reveals several important insights. While Senapaty et al. (2023) reported 100% accuracy using SGDC, they also reported 0.54 specificity, which may indicate potential generalization issues. Our Random Forest approach achieves 99.32% accuracy while maintaining balanced performance across all evaluation metrics.

Table 10. Comparative Literature Review of Crop Recommendation Systems

Study	Dataset (Source, Size)	Methods & Technology	Model Performance
Senapaty et al. (2023)	246.091 sample records from the Dataworld website	SGDC, MSVM-DAG-FFO, Voting	SGDC accuracy: 100% But 0.54 specifity
Musanase et al. (2023)	10.440 samples from various sources	Neural Networks	Accuracy: 97%
Patel and Patel (2023)	5000 samples from Gujarat Region Dataset	AgriRec Algorithm	Accuracy: 95.85
Proposed System	2200 samples from Kaggle website	RF, SVM, KNN, XGBoost + GridSearchCV + SMOTE + SHAP	RF accuracy: 99.32%

Table 11. Comparative Literature Review of Fertilizer Recommendation Systems

Study	Dataset (Source, Size)	Methods & Technology	Model Performance
Patel and Patel (2023)	Gujarat Fertilizer Dataset	AgriRec Algorithm	Accuracy: 92.11
Ikhlaq and Kechadi (2023)	~3000 fields (large-scale, heterogeneous)	Random Forest	Accuracy: 74–94%
Musanase et al. (2023)	10.440 samples from various sources	Rule-based system	No quantitative performance
Hossain and Siddique (2020)	National soil database, Upazila Nirdeshika, laboratory soil analyses	Rule-based	7–22% higher crop yields
Proposed System	3100 samples from Kaggle website	RF, SVM, KNN, XGBoost + GridSearchCV + SMOTE + SHAP	RF accuracy: 98.75%

Compared to Musanase et al. (2023) with 97% accuracy and Patel and Patel (2023) with 95.85% accuracy, our approach demonstrates improved performance, which can be attributed to the comprehensive preprocessing pipeline including SMOTE application and systematic hyperparameter optimization.

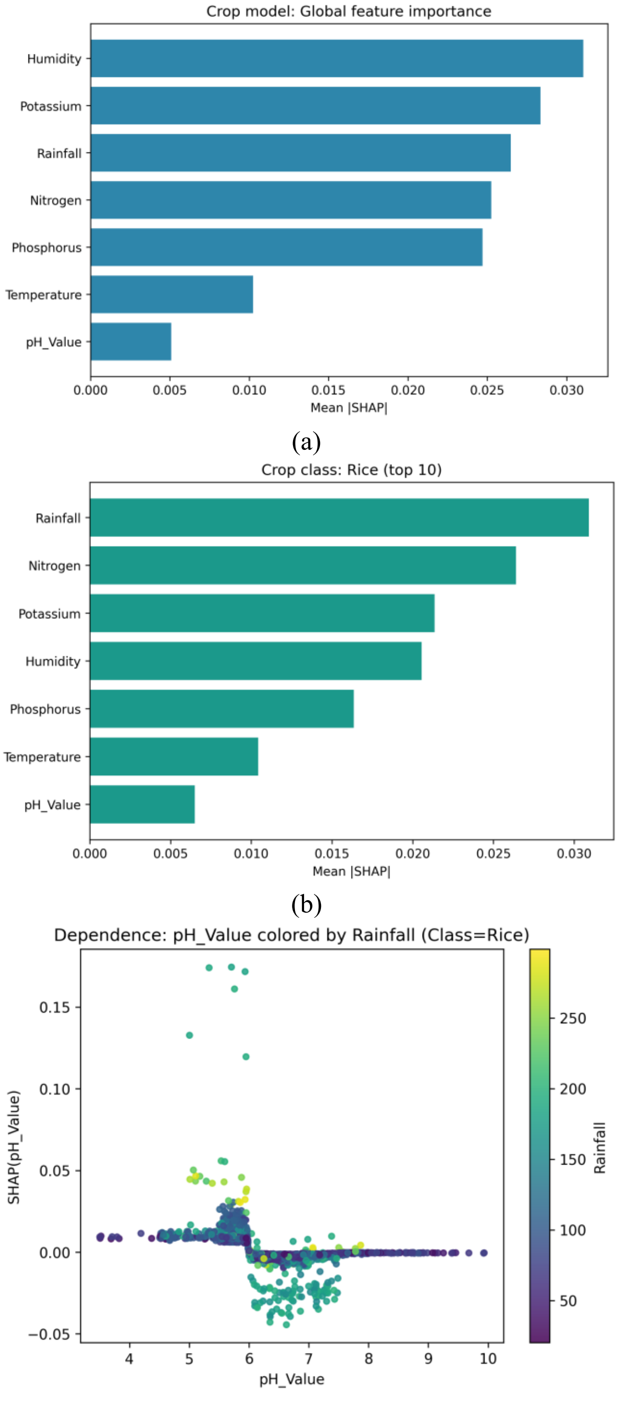
The fertilizer recommendation comparison demonstrates substantial advancement over existing approaches. Ikhlaq and Kechadi (2023) reported variable accuracy ranges of 74-94% for the first stage and 71.2-99.5% for the second stage, indicating inconsistent performance across different nutrients and application scenarios. Our integrated single-stage approach achieves consistent 98.75% accuracy across all fertilizer categories, eliminating the complexity associated with multi-stage prediction frameworks. The improvement over Patel and Patel's (2023) 92.11% accuracy and the rule-based approaches used by other studies highlights the effectiveness of our machine learning methodology.

The comparative analysis results demonstrate that the proposed system exhibits superior or competitive performance compared to existing studies in both agricultural domains. The main distinguishing advantages of the proposed system are: (1) dual-module approach providing both crop and fertilizer recommendations in a single integrated platform, (2) explainable AI implementation through SHAP analysis enabling farmers to understand the scientific reasoning behind recommendations, (3) comprehensive preprocessing including SMOTE for class imbalance handling and systematic hyperparameter optimization, (4) region-adaptable framework design that can be customized for local agricultural conditions, and (5) integration of advanced machine learning techniques with practical agricultural applications.

The system's dual-module approach represents a significant advancement over existing single-purpose systems found in the literature. Most existing studies focus on either crop recommendation or fertilizer management independently, while our integrated approach considers the interdependencies between these critical agricultural decisions. The integration of advanced preprocessing techniques, systematic hyperparameter optimization, and explainable AI methods ensures both high performance and practical applicability in real-world agricultural scenarios.

The SHAP (SHapley Additive exPlanations) analysis provides comprehensive insights into the agricultural significance of different features and their interactions within both recommendation modules. The interpretability analysis is presented through four complementary visualizations (Figures 1-4) that demonstrate global feature importance, local explanations, and model performance evaluation.

Figure 1 presents a comprehensive analysis of the crop recommendation model through three complementary perspectives. The global feature importance analysis (Figure 1a) reveals that soil pH and rainfall emerge as the most influential variables across all crop classes, with mean absolute SHAP values indicating their dominant role in crop suitability decisions. The class-specific analysis for rice cultivation (Figure 1b) demonstrates how the top 10 features specifically influence rice recommendations, providing targeted insights for this economically important crop.



(a)

(b)

(c)

Figure 1. Crop recommendation model interpretability analysis.

(a) Global feature importance ranking based on mean absolute SHAP values across all crop classes. (b) Class-specific SHAP analysis for rice showing the top 10 most influential features. (c) SHAP dependence plot illustrating pH-rainfall interaction for rice classification

The dependence plot (Figure 1c) illustrates the complex interaction between pH levels and rainfall patterns, where near-neutral pH conditions (6.0-7.0) combined with adequate rainfall (>100mm) jointly create optimal conditions for rice cultivation. This finding aligns with established agronomic principles regarding nutrient availability and water management requirements for rice production.

Figure 2 examines the fertilizer recommendation model through critical environmental dependencies that govern nutrient management decisions. The rainfall-temperature interaction (Figure 2a) reveals how precipitation patterns and thermal conditions work synergistically to modulate nutrient mobility and fertilizer effectiveness in soil systems. The pH-moisture relationship (Figure 2b) demonstrates how soil acidity levels and water content jointly determine nutrient availability and uptake efficiency. These SHAP dependence plots effectively capture the non-linear relationships and complex interactions between environmental factors that traditional linear approaches often fail to represent, enabling more precise fertilizer recommendations based on site-specific conditions.

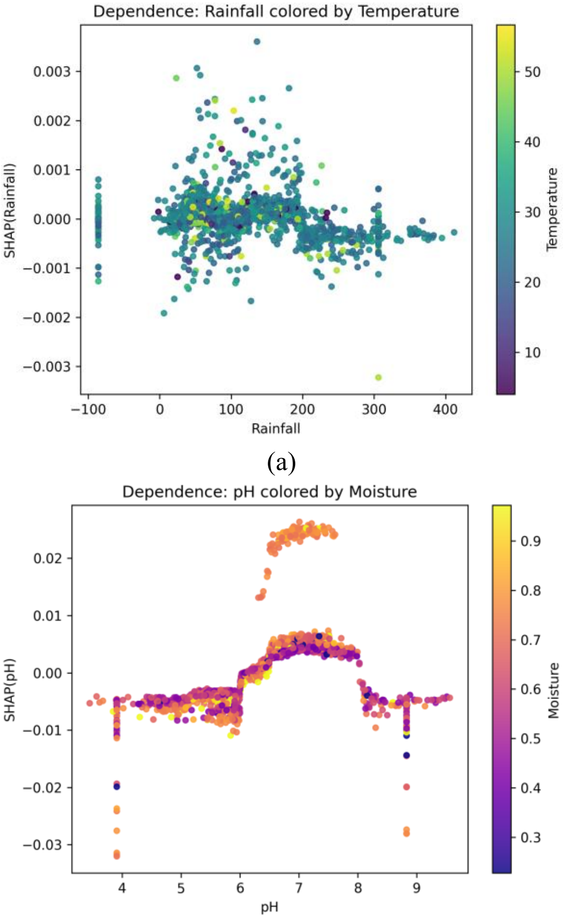
Figure 3 provides a detailed local explanation through a waterfall plot for a representative rice cultivation instance, demonstrating the transparent decision-making process of the crop recommendation model. This visualization breaks down how individual environmental and soil parameters contribute to the final prediction, with each feature's contribution clearly quantified and directionally indicated. The local explanation shows that pH levels (6.2) and rainfall patterns (202mm) serve as primary explanatory factors that determine both the direction and confidence level of the rice recommendation. This instance-level explanation framework is crucial for building farmer trust and facilitating technology adoption by providing scientifically transparent reasoning that can be validated against field observations.

Figure 4 presents the confusion matrices for both recommendation modules, providing comprehensive performance evaluation across all crop and fertilizer categories. The crop recommendation model (Figure 4a) demonstrates exceptional classification accuracy with strong diagonal patterns indicating minimal misclassification errors across 22 crop types. The fertilizer recommendation model (Figure 4b) shows similarly robust performance in distinguishing between various fertilizer requirements. The balanced performance across all classes confirms that the SMOTE technique effectively addressed class imbalance issues, ensuring reliable recommendations for both common and specialized crop-fertilizer combinations.

Discussion

The results of this study demonstrate the effectiveness of machine learning approaches in agricultural decision support systems, with particular emphasis on the benefits of ensemble methods and explainable AI techniques. The superior performance of Random Forest across both datasets can be attributed to its ability to handle complex, non-linear relationships in agricultural data while maintaining robustness against overfitting through bootstrap aggregation.

The integration of SHAP analysis provides crucial insights into the agricultural relevance of model predictions. The identification of pH and rainfall as key factors for rice recommendation aligns with established agricultural science, where soil pH affects nutrient availability and water management is critical for rice cultivation. This concordance between model outputs and agronomic knowledge enhances the credibility and practical applicability of the system for agricultural practitioners.



(a)

(b)

Figure 2. Fertilizer recommendation model feature dependencies.

(a) SHAP dependence plot showing rainfall-temperature interaction effects on fertilizer recommendations. (b) SHAP dependence plot illustrating pH-moisture relationship in fertilizer decision-making.

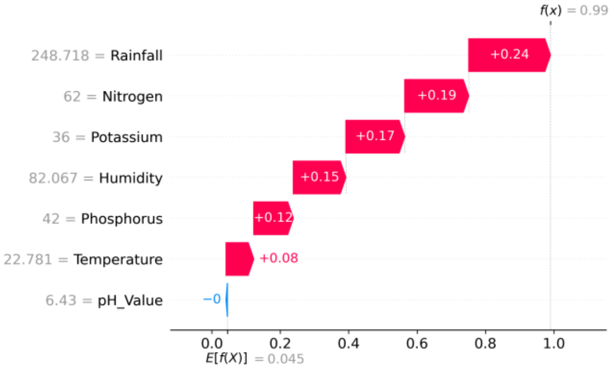


Figure 3. Local explanation for crop recommendation.

SHAP waterfall plot demonstrating individual feature contributions for a representative rice cultivation instance, showing how each parameter influences the final recommendation with quantified impact values.

Agricultural Significance of Feature Relationships

The SHAP analysis reveals critical agricultural relationships that align with established agronomic principles. The identification of pH as a primary factor for rice recommendation reflects the crop's physiological requirements for nutrient uptake. Rice cultivation thrives in slightly acidic to neutral pH conditions (6.0- 7.0) because this range optimizes the availability of essential nutrients, particularly nitrogen, phosphorus, and micronutrients such as iron and zinc. Outside this optimal pH range, nutrient availability decreases significantly: in highly acidic soils (pH < 5.5), aluminum and manganese toxicity can occur, while in alkaline soils (pH > 8.0), phosphorus becomes less available due to precipitation with calcium.

The strong influence of rainfall patterns on crop recommendations demonstrates the model's understanding of water-nutrient interactions. Adequate rainfall (>100mm) not only provides necessary moisture for rice cultivation but also facilitates nutrient dissolution and transport in the soil solution, enabling efficient root uptake. The interaction between pH and rainfall is particularly significant: higher rainfall in acidic soils can lead to nutrient leaching, particularly of nitrogen and potassium, necessitating adjusted fertilizer recommendations.

Crop-Fertilizer Relationship Analysis

The dual-module system's recommendations demonstrate scientifically sound crop- fertilizer relationships. For nitrogen-demanding crops such as rice and maize, the system appropriately recommends higher nitrogen fertilizer rates when soil nitrogen levels are deficient. The NPK balance in fertilizer recommendations reflects crop- specific nutrient uptake patterns: rice requires higher nitrogen during vegetative growth stages, while leguminous crops (chickpea, lentil) show reduced nitrogen requirements due to biological nitrogen fixation capabilities.

The model's sensitivity to phosphorus levels is particularly important for root development and flowering. Crops such as cotton and legumes, which have high phosphorus requirements during reproductive stages, receive appropriately adjusted fertilizer recommendations when soil phosphorus levels are suboptimal. Similarly, potassium recommendations align with crop requirements for water regulation and disease resistance, with higher rates suggested for crops like banana and potato that have elevated potassium demands.

Environmental Context and Nutrient Dynamics

Temperature and humidity interactions captured by the model reflect their influence on nutrient mineralization and availability. Higher temperatures (25-30°C) accelerate organic matter decomposition and nitrogen mineralization, potentially reducing fertilizer requirements. Conversely, lower temperatures slow these processes, necessitating adjusted fertilizer application rates. Humidity affects nutrient mobility and uptake efficiency: optimal humidity levels (60-80%) facilitate nutrient transport through the soil-plant continuum, while excessive humidity can lead to waterlogging and reduced oxygen availability, affecting root function and nutrient uptake.

These agricultural interpretations demonstrate that the machine learning models have successfully captured the complex, multi-factorial relationships governing crop selection and fertilizer management. The system's recommendations are not merely statistical correlations but reflect fundamental agronomic principles related to plant physiology, soil chemistry, and nutrient dynamics. This agricultural validity enhances the system's practical applicability and supports its adoption by farming communities seeking scientifically-grounded decision support.

The dual-module approach represents a significant advancement over existing single-purpose systems in the literature. While most studies focus on either crop recommendation or fertilizer management, the integrated approach provides farmers with comprehensive guidance that considers the interdependencies between crop selection and nutrient management. This holistic perspective is particularly valuable for sustainable agricultural practices where crop and fertilizer decisions must be coordinated to optimize both productivity and environmental impact.

The successful implementation of SMOTE for handling class imbalance demonstrates the importance of addressing data quality issues in agricultural machine learning applications. Agricultural datasets often exhibit imbalanced distributions due to varying crop popularity and regional preferences, which can lead to biased model performance. The consistent results across all crop and fertilizer categories indicate that the preprocessing approach effectively mitigated these challenges.

The system's contribution to precision agriculture extends beyond individual farm-level decisions to support broader agricultural sustainability objectives. By providing evidence-based recommendations that optimize resource use efficiency, the system can contribute to reducing agricultural environmental impact while maintaining productivity. This approach supports sustainable agricultural transformation and climate-resilient farming practices.

Limitations and Future Work

This study has several limitations that should be acknowledged. The datasets used reflect general agricultural conditions and are not specific to Turkish geographical and climatic conditions. While the framework is designed to be region-adaptable, validation with local Turkish agricultural data would enhance the system's practical applicability for domestic farming practices. Additionally, the current implementation does not incorporate temporal dynamics such as seasonal variations, crop rotation patterns, or long-term soil health changes, which are important factors in sustainable agricultural planning.

The system currently focuses on soil and environmental parameters but does not integrate economic factors such as market prices, input costs, or profitability analysis, which are crucial considerations for farmers' decision-making processes. Furthermore, while the SHAP analysis provides valuable interpretability, the system would benefit from integration with expert agricultural knowledge systems to provide more comprehensive guidance.

Future research directions include several promising avenues for enhancement. Region-specific adaptations for Turkish agricultural conditions should be developed through collaboration with local agricultural institutions, incorporating temporal dynamics and seasonal variations in recommendations. The integration of economic factors and market conditions into the decision-making process would provide more comprehensive guidance for farmers. Expanding the system to include pest and disease management recommendations, developing mobile applications for field deployment, and investigating the integration of climate change projections for long-term agricultural planning represent important extensions of this work. Additionally, exploring advanced deep learning architectures such as TabNet and Transformer-based models (Venkateswara & Padmanaban, 2025; Khaliq et al., 2025) could potentially enhance performance further while maintaining interpretability. The comparison between ensemble methods and attention-based architectures in Turkish agricultural contexts would provide valuable insights for optimal model selection. Implementing federated learning approaches could enable collaborative model training across different regions while preserving data privacy and regional specificity.

Conclusions

This study presented a dual-module agricultural decision support system that effectively integrates machine learning algorithms for both crop and fertilizer recommendations. The Random Forest algorithm demonstrated superior performance in both modules, achieving 99.32% accuracy for crop recommendation and 98.75% accuracy for fertilizer recommendation. The integration of SMOTE for class imbalance handling, GridSearchCV for hyperparameter optimization, and SHAP analysis for model interpretability ensures both high performance and practical applicability in agricultural contexts.

The comparative analysis with existing literature demonstrates that the proposed system achieves competitive accuracy rates while offering a dual-module approach that provides integrated crop and fertilizer recommendations within a single platform. The explainable AI implementation through SHAP analysis addresses the need for transparency in agricultural decision support systems, enabling farmers to understand the scientific reasoning behind recommendations.

Comparison with Advanced Deep Learning Architectures

While the proposed Random Forest-based system achieves competitive performance, recent advances in attention-based deep learning architectures offer complementary advantages. Venkateswara and Padmanaban (2025) demonstrated that TabNet's attention mechanism enables sequential feature selection, providing built-in interpretability alongside high accuracy (95.24% for fertilizer, 96.21% for crop). Similarly, Khaliq et al. (2025) showed that transformer-based architectures can achieve superior performance (99.3% for fertilizer) while maintaining explainability through integrated XAI techniques.

The choice between ensemble methods (Random Forest, XGBoost) and attention-based deep learning architectures depends on several factors: (1) Dataset size and complexity: Deep learning models typically require larger datasets to fully leverage their capacity, while ensemble methods perform well with moderate-sized datasets; (2) Computational resources: Transformer-based models require more computational power during training, while Random Forest offers faster training and prediction; (3) Interpretability requirements: While both approaches can be made interpretable through post-hoc methods (SHAP, LIME), attention-based models provide built-in interpretability through attention weights; (4) Deployment constraints: Ensemble methods are generally easier to deploy in resource-constrained environments.

The proposed system's strength lies in its balanced approach: achieving high accuracy (99.32% for crop, 98.75% for fertilizer) with Random Forest while maintaining computational efficiency and providing comprehensive explainability through SHAP analysis. This makes the system particularly suitable for practical deployment in diverse agricultural contexts, including regions with limited computational infrastructure.

The system's region-adaptable framework design represents a practical contribution to agricultural informatics, providing a foundation that can be customized for different geographical and climatic conditions. This adaptability is important for addressing diverse agricultural contexts and supporting context-specific sustainable farming practices.

The findings indicate that ensemble learning methods, particularly Random Forest, are well-suited for agricultural classification tasks due to their ability to handle complex, non-linear relationships in agricultural data. The integration of explainable AI techniques demonstrates the potential for developing trustworthy agricultural decision support systems that can facilitate technology adoption among farming communities.

This research contributes to the digital transformation of agriculture and supports the development of sustainable agricultural practices. The dual-module approach provides a framework that can be extended to include additional agricultural decision support functions, creating integrated platforms for sustainable farm management. Future work should focus on regional adaptation, temporal modeling, integration with real-world agricultural practices, and incorporation of economic factors, climate projections, and real-time data streams to maximize the system's impact on sustainable farming and food security.

Declarations

Ethical Approval Certificate

Not applicable for this study as it involves only publicly available datasets and does not involve human subjects or animal experiments.

Source of the Study

This article is derived from the undergraduate thesis titled "Akıllı Tarım Tavsiye Sistemi" by Hatice Bülbül and Enes Güler, Department of Information Systems Engineering, Faculty of Computer and Information Sciences, Sakarya University, 2025 (Bülbül & Güler, 2025).

Author Contribution Statement

Hatice Bülbül: Data collection, investigation, formal analysis, methodology implementation, and writing the original draft

Enes Güler: Data collection, investigation, model development, validation, and writing the original draft

Burcu Çarklı Yavuz: Project administration, supervision, conceptualization, methodology design, review and editing, and final manuscript preparation

Fund Statement

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments

The authors utilized artificial intelligence for language editing and improvement purposes in this work.

References

Arrighi, L., de Moraes, I. A., Zullich, M., Simonato, M., Barbin, D. F., & Barbon Junior, S. (2025). Explainable Artificial Intelligence techniques for interpretation of food datasets: a review. *arXiv preprint*. https://arxiv.org/html/2504.10527v1

Breiman, L. (2001). *Random Forests*. Machine Learning 45, 5–32. https://doi.org/10.1023/A:1010933404324

Bülbül, H., & Güler, E. (2025). *Akıllı Tarım Tavsiye Sistemi* [Smart Agricultural Recommendation System]. Undergraduate thesis, Department of Information Systems Engineering, Faculty of Computer and Information Sciences, Sakarya University.

Cartolano, A., Cuzzocrea, A., & Pilato, G. (2024). Analyzing and assessing explainable AI models for smart agriculture environments. *Multimedia Tools and Applications*, 83, 37225–37246. https://doi.org/10.1007/s11042-023-17978-z

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321-357. https://doi.org/10.1613/jair.953

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794). https://doi.org/10.1145/2939672.2939785

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297. https://doi.org/10.1007/BF00994018

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21-27. https://doi.org/10.1109/TIT.1967.1053964

Demir, B., Dokuz, Y., & Şen, B. (2025). Comparative Analysis of Machine Learning Algorithms for Irrigation Status Prediction. Turkish Journal of Agriculture - Food Science and Technology, 13(2), 497-503. https://doi.org/10.24925/turjaf.v13i2.497-503.7497

Essenfelder, A. H., Toreti, A., & Seguini, L. (2025). Expert-driven explainable artificial intelligence models can detect multiple climate hazards relevant for agriculture. *Communications Earth & Environment*, 6, 207. https://doi.org/10.1038/s43247-024-01987-3

Food and Agriculture Organization of the United Nations. (2023). *The State of Food Security and Nutrition in the World 2023*. FAO.

Hossain, M. A., & Siddique, M. N. A. (2020). Online Fertilizer Recommendation System (OFRS): A step towards precision agriculture and optimized fertilizer usage by smallholder farmers in Bangladesh. *European Journal of Environment and Earth Sciences*, *1*(4), 1–9. https://doi.org/10.24018/ejgeo.2020.1.4.47

Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. International Journal of Data Mining & Knowledge Management Process, 5(2), 1-11. https://doi.org/10.5121/ijdkp.2015.5201

Ikhlaq, U., & Kechadi, T. (2023, April 28–30). Machine learning-based nutrient application's timeline recommendation for smart agriculture: A large-scale data mining approach. In *Proceedings of the International Conference on Intelligent Systems and New Applications (ICISNA'23)* (pp. 42–47). Liverpool, United Kingdom.

Khaliq, A., et al. (2025). AI-Driven Smart Agriculture: An Integrated Approach for Soil Analysis, Irrigation, and Crop-Fertilizer Recommendations. IEEE Access, 13, 141124-141138. https://doi.org/10.1109/ACCESS.2025.3594162

Liu, J.-J., Wu, H., & Riaz, I. (2025). Advanced Technologies for Smart Fertilizer Management in Agriculture: A Review. IEEE Access, 13, 139766-139786. https://doi.org/10.1109/ACCESS.2025.3594361

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774. https://doi.org/10.48550/arXiv.1705.07874

Mohan, R. N. V. J., Rayanoothala, P. S., & Sree, R. P. (2025). Next-gen agriculture: integrating AI and XAI for precision crop yield predictions. *Frontiers in Plant Science*, 15, 1451607. https://doi.org/10.3389/fpls.2024.1451607

Musanase, R., Niyitegeka, D., & Tuyishimire, E. (2023). Data-driven analysis and machine learning-based crop and fertilizer recommendation system for sustainable agriculture. *Agriculture*, *13*(11), 2141. https://doi.org/10.3390/agriculture13112141

Nalluri, V. (2024). Crop recommendation dataset (Version 1) [Data set]. Kaggle. https://www.kaggle.com/datasets/varshitanalluri/crop-recommendation-dataset/data (Accessed: 1 April 2025)

Nishchalchandel. (2025). Fertilizer recommendation (Version 1) [Data set]. Kaggle. https://www.kaggle.com/datasets/nishchalchandel/fertilizer-recommendation/data (Accessed: 1 April 2025)

Patel, K., & Patel, H. B. (2023). Multi-criteria Agriculture Recommendation System using Machine Learning for Crop and Fertilizers Prediction. *Current Agriculture Research Journal*, 11(1), 137-149. https://doi.org/10.12944/CARJ.11.1.12

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825-2830.

Prity, F. S., Hasan, M. M., Saif, S. H., Hossain, M. M., Bhuiyan, S. H., Islam, M. A., & Lavlu, M. T. H. (2024). Enhancing Agricultural Productivity: A Machine Learning Approach to Crop Recommendations. *Human-Centric Intelligent Systems*, 4, 497–510. https://doi.org/10.1007/s44230-024-00081-3

Senapaty, M. K., Ray, A., & Padhy, N. (2023). IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture. *Computers*, 12(3), 61. https://doi.org/10.3390/computers12030061

Shastri, S., Kumar, S., Mansotra, V., & Salgotra, R. (2025). Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence. *Scientific Reports*, 15, 25498. https://doi.org/10.1038/s41598-025-07003-8

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information Processing & Management, 45(4), 427-437. https://doi.org/10.1016/j.ipm.2009.03.002

Tanaka, T. S. T., Heuvelink, G. B. M., Mieno, T., & Bullock, D. S. (2024). Can machine learning models provide accurate fertilizer recommendations? *Precision Agriculture*, 25, 1839–1856. https://doi.org/10.1007/s11119-024-10136-x

Thorat, B., Ranadive, I., Rajpurohit, N., Mehare, Y., & Singh, R. (2025). Fertilizer Prediction Using Machine Learning. *International Journal for Research in Applied Science and Engineering Technology*, 13(4). https://doi.org/10.22214/ijraset.2025.70432

United Nations Convention to Combat Desertification. (2022). *Global Land Outlook, Second Edition*. UNCCD.

Venkateswara, S. M., & Padmanaban, J. (2025). Interpretable deep learning models for independent fertilizer and crop recommendation. Scientific Reports, 15, 41721. https://doi.org/10.1038/s41598-025-26910-4