IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141181

CROP RECOMMENDATION SYSTEM USING MACHINE LEARNING

Arpita Yogendra Patil[1], Prof. Shivam Limbare[2], Manoj V. Nikum[3]

Student Of MCA, Shri Jaykumar Rawal Institute of Technology Dondaicha, KBC NMU Jalgaon, Maharashtra, India[1]

Assistant Professor, MCA Department, SJRIT DONDAICHA, KBC NMU JALGAON, Maharashtra, India[2]

Assistant Professor & HOD, MCA Department, SJRIT DONDAICHA, KBC NMU JALGAON, Maharashtra, India[3]

Abstract: Agriculture is a vital sector that supports the livelihood of millions of people, particularly in developing countries like India, where farming remains a primary source of income. However, the sector faces persistent challenges such as improper crop selection, soil nutrient imbalances, climate variability, water scarcity, and lack of scientific decision-making. Traditional farming practices rely heavily on farmers' intuition, experience, or generalized recommendations, which often result in poor crop yield, excessive fertilizer usage, and increased vulnerability to environmental fluctuations. To address these issues and support precision agriculture, this research proposes a machine-learning-based Crop Recommendation System that utilizes soil nutrient values—Nitrogen (N), Phosphorus (P), and Potassium (K)—along with environmental factors such as temperature, humidity, pH level, and rainfall to recommend the most suitable crop for cultivation in a given region.

The proposed system uses supervised machine learning algorithms, including Random Forest, Decision Tree, Naive Bayes, and Support Vector Machine (SVM), to analyze large agricultural datasets and learn complex patterns between soil–climate features and crop suitability. The dataset undergoes extensive preprocessing, which includes handling missing values, normalizing numeric attributes, removing noise and outliers, and encoding categorical labels. Feature engineering further enhances the prediction quality by identifying the most influential variables such as the N:P:K ratio, soil fertility index, and climate–soil interactions. These features help the model better distinguish crop requirements under varying environmental conditions.

During the experimental phase, each algorithm was trained and evaluated using standard performance metrics, including accuracy, precision, recall, and F1-score. Results show that the Random Forest classifier outperformed other models, achieving an accuracy of 97%, largely due to its ensemble nature, robustness, and ability to handle high-dimensional data. The findings highlight that machine learning can significantly improve agricultural decision-making by offering farmers scientific guidance tailored to their land conditions. This system has the potential to enhance crop productivity, minimize risks associated with crop failure, optimize fertilizer usage, and promote long-term soil health.

This research demonstrates that integrating machine learning into agriculture provides a practical and scalable solution to modern farming challenges. Future advancements may include IoT-enabled soil sensors, satellite-based remote sensing, real-time data analysis, and deep learning models for yield prediction and dynamic crop recommendation. Overall, the study emphasizes the transformative potential of machine learning in supporting sustainable and smart agriculture. This research aims to develop and evaluate a machine-learning-based crop recommendation system using key soil and climatic features. By comparing multiple ML algorithms and identifying the most accurate model, the study contributes to the growing field of smart agriculture and demonstrates how technology can transform traditional farming practices.

I. INTRODUCTION

Agriculture plays a crucial role in the economic development and food security of many nations, especially in countries like India where a large proportion of the population depends on farming for their livelihood. Despite technological advancements, many farmers continue to rely on traditional practices, intuition, and prior experience when selecting crops for cultivation. However, crop selection is a complex decision that depends on multiple factors such as soil nutrient composition, climatic conditions, geographical characteristics, and seasonal variations. In the absence of scientific guidance, improper crop selection often leads to reduced productivity, soil deterioration, excessive use of chemical fertilizers, and increased financial risk for farmers.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⸎ Peer-reviewed & Refereed journal ⸎ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141181

One of the most persistent challenges in agriculture is the lack of awareness and accessibility to accurate soil and environmental data. Different crops have distinct requirements for soil nutrients such as Nitrogen (N), Phosphorus (P), and Potassium (K), as well as specific thresholds for pH, temperature, rainfall, and humidity. For instance, rice requires high moisture and rainfall, while crops like chickpea grow best in drier conditions with moderate nutrient levels. When the crop chosen does not align with the natural properties of the soil, the result is poor yield and economic loss. Therefore, there is a strong need for a scientific, data-driven method that assists farmers in making informed decisions regarding crop selection.

Machine learning (ML) has emerged as a powerful technology capable of addressing many challenges in modern agriculture. By analyzing large datasets containing soil parameters, climate data, and historical crop performance, machine learning algorithms can identify complex relationships that may not be apparent through traditional analysis. ML models can learn patterns from existing data and predict the most suitable crop for cultivation in specific conditions. This not only simplifies decision-making but also reduces human error and provides farmers with real-time, personalized recommendations.

The integration of machine learning into agriculture marks a major step toward precision farming—a methodology focused on optimizing resources, reducing environmental impact, and increasing agricultural efficiency. A crop recommendation system powered by machine learning can help farmers maximize yield, reduce fertilizer waste, and improve soil health. Additionally, it promotes sustainable agriculture by encouraging the cultivation of crops that are naturally compatible with local environmental conditions.

This research aims to develop and evaluate a machine-learning-based crop recommendation system using key soil and climatic features. By comparing multiple ML algorithms and identifying the most accurate model, the study contributes to the growing field of smart agriculture and demonstrates how technology can transform traditional farming practices.

II. LITERATURE SURVEY

A considerable amount of research has been conducted in the field of agricultural prediction, crop suitability analysis, and precision farming using machine learning techniques. The increasing availability of agricultural datasets and advancements in artificial intelligence have motivated researchers to develop automated tools that provide accurate and personalized recommendations to farmers. This literature review summarizes key contributions from previous studies and highlights the research gaps addressed in the present work.

Early studies in crop prediction relied heavily on statistical models and rule-based systems. These systems, however, lacked adaptability and often failed to handle complex, nonlinear relationships between soil parameters, climatic variables, and crop characteristics. With the emergence of machine learning, researchers began applying classification algorithms such as Decision Trees, Support Vector Machines (SVM), and Naive Bayes to predict crop types based on soil nutrient levels and environmental conditions.

Patel et al. demonstrated that soil nutrient concentrations—Nitrogen (N), Phosphorus (P), and Potassium (K)—play a crucial role in determining crop suitability. Their work used Decision Tree classifiers to establish crop–soil relationships and achieved moderate accuracy. However, the model struggled with high-dimensional features and was susceptible to overfitting. Kumar and Singh extended this work by implementing Random Forest classifiers for crop recommendation. Their findings showed that ensemble techniques outperform single-classifier models by reducing variance and improving generalization across diverse soil profiles.

Huang et al. introduced Support Vector Machines for agricultural classification tasks and observed improved performance in datasets with complex patterns. Their model effectively handled multiple crop categories but required extensive parameter tuning, making it computationally expensive for large-scale deployment. Jain and Khandare explored Naive Bayes classifiers for crop prediction, emphasizing the importance of preprocessing and feature normalization. However, their model was limited when dealing with continuous data and nonlinear interactions.

Recent advancements in deep learning have inspired researchers to experiment with neural network-based approaches. LSTM and CNN models have been used to predict crop yield and classify soil fertility levels. While these methods offer high accuracy, they demand large training datasets and significant computational resources, which may not be feasible for small-scale agricultural systems.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⧘ Peer-reviewed & Refereed journal ⧘ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.141181

IoT-enabled smart farming solutions have also been proposed, integrating real-time soil sensor data with machine learning models for dynamic crop recommendation. Although effective, such systems are costly and difficult to implement in rural farming communities.

The literature indicates that Random Forest and SVM consistently perform well for crop recommendation tasks due to their robustness and ability to handle nonlinear relationships. However, there is still a need for models that are accurate, scalable, computationally efficient, and suitable for real-world agricultural conditions. This research addresses these gaps by developing a machine-learning-based crop recommendation system that is both accurate and practical for farmers.

III. RESEARCH METHODOLOGY

The methodology adopted for developing the Crop Recommendation System using machine learning is structured into multiple phases, beginning with data collection and followed by preprocessing, feature engineering, model selection, training, evaluation, and final deployment. Each stage is designed to ensure that the machine learning model effectively learns patterns from soil and climatic data to provide accurate crop recommendations.

1. Data Collection
The foundation of the system is a comprehensive agricultural dataset containing soil nutrient values (Nitrogen, Phosphorus, Potassium), environmental parameters (temperature, humidity), soil pH level, and rainfall. These attributes significantly influence crop growth and productivity. The dataset may be sourced from agricultural research organizations, government portals, or open-source repositories commonly used in precision farming research. Collecting diverse and high-quality data ensures that the system can generalize well across different regions.

2. Data Preprocessing
Preprocessing is essential to enhance data quality and remove inconsistencies. This stage includes handling missing values, eliminating duplicate records, and correcting noisy entries. Numerical attributes undergo normalization or scaling to ensure uniformity and prevent algorithmic bias during training. Outliers are detected and removed using statistical techniques such as Z-score or IQR methods. Categorical labels representing crop names are transformed into numerical form using label encoding to make them suitable for model training.

3. Feature Engineering
Feature engineering improves model performance by creating new, meaningful attributes. In this study, several derived features such as the N:P:K ratio, soil fertility index, ideal pH range classification, and temperature–rainfall interaction were used. These engineered features help machine learning models better understand complex relationships between soil properties and crop requirements.

4. Model Selection and Training
Four supervised learning algorithms—Decision Tree, Random Forest, Naive Bayes, and Support Vector Machine (SVM)—were selected for experimentation due to their effectiveness in classification tasks. The dataset is divided into training and testing subsets (typically 80:20). Each model is trained on the training data to learn crop suitability patterns. Hyperparameters are tuned to optimize model performance.

5. Model Evaluation
The trained models are evaluated using accuracy, precision, recall, F1-score, and confusion matrix metrics. These performance measures help compare each model's effectiveness. Random Forest achieved the highest accuracy because of its ensemble nature and robustness in handling nonlinear data.

6. System Deployment
The final model is integrated into a user-friendly interface where farmers can input soil and climate values to receive instant crop recommendations. The system is designed to be scalable, efficient, and easy to use.

IV. RESULTS AND DISCUSSION

The performance of the machine learning models used in the Crop Recommendation System was thoroughly evaluated to determine their effectiveness in accurately predicting the most suitable crop based on soil and environmental parameters. Four supervised learning algorithms—Decision Tree, Random Forest, Naive Bayes, and Support Vector Machine (SVM)—were tested using the prepared dataset. The evaluation metrics included accuracy, precision, recall,

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⫶ Peer-reviewed & Refereed journal ⫶ Vol. 14, Issue 11, November 2025
DOI: 10.17148/IJARCCE.2025.141181

F1-score, and analysis through confusion matrices. These metrics provide a comprehensive understanding of the strengths and limitations of each model.

1. Model Performance Comparison

The results show that the Random Forest classifier achieved the highest overall performance with an accuracy of **97%**, followed by SVM with **94%**, Decision Tree with **92%**, and Naive Bayes with **85%**. The superior performance of Random Forest is attributed to its ensemble learning approach, which combines multiple decision trees to reduce overfitting and enhance generalization. The model effectively captures complex nonlinear relationships between soil nutrients, climatic factors, and crop suitability.

2. Analysis of Accuracy Metrics

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	0.92	0.89	0.88	0.88
Naive Bayes	0.85	0.82	0.80	0.81
SVM	0.94	0.91	0.92	0.91
Random Forest	**0.97**	**0.95**	**0.96**	**0.95**

These findings clearly indicate that Random Forest delivers the best balance between precision and recall, making it a reliable choice for real-world applications. On the other hand, Naive Bayes underperformed due to its assumption of feature independence, which does not hold well for continuous and correlated agricultural features such as NPK levels.

3. Confusion Matrix Interpretation

The confusion matrices reveal that Random Forest produced the least number of false predictions, demonstrating strong classification capability across all crop categories. The model correctly identified crops even when soil nutrient values were closely aligned between multiple crop options, a scenario where simpler models such as Naive Bayes struggled.

4. Insights from Feature Importance Analysis

Feature importance analysis further revealed that nitrogen content, pH value, rainfall, and temperature were the most influential attributes in determining crop suitability. This aligns with agricultural principles that emphasize the role of balanced soil nutrients and favorable climatic conditions.

5. Discussion

Overall, the results highlight the effectiveness of the proposed machine learning-based system in providing accurate crop recommendations. The high performance of the Random Forest classifier shows its suitability for deployment in real-world agricultural advisory systems. The study also demonstrates that machine learning can significantly support farmers by providing data-driven, scientific recommendations.

V. CONCLUSION AND FUTURE SCOPE

The rapid advancement of technology in recent years has opened new opportunities for enhancing agricultural productivity through data-driven decision-making. This research focused on developing a machine-learning-based Crop Recommendation System capable of assisting farmers in selecting the most suitable crop for cultivation based on soil nutrients and environmental factors. By analyzing parameters such as Nitrogen (N), Phosphorus (P), Potassium (K), temperature, humidity, pH level, and rainfall, the system provides scientific recommendations that can significantly improve crop yield and reduce the risk of crop failure.

The experimental results demonstrated that among the four machine learning models evaluated—Decision Tree, Random Forest, Naive Bayes, and Support Vector Machine (SVM)—the Random Forest classifier performed the best, achieving an accuracy of **97%**. Its ensemble nature, robustness to noise, ability to handle nonlinear relationships, and resistance to overfitting make it the most suitable model for agricultural applications. The performance metrics, confusion matrix analysis, and feature importance results indicate that machine learning can effectively identify complex soil–climate–crop relationships and provide reliable recommendations to farmers.

This research successfully highlights the potential of machine learning in precision agriculture. The proposed system can aid in optimizing the use of fertilizers, improving soil health, promoting sustainable farming, and reducing dependency on traditional, experience-based crop selection practices. The system's scientific approach to crop recommendation represents a significant step toward modernizing agriculture and ensuring food security.

❖ **Future Scope**

While the current system demonstrates high accuracy and practical usability, several enhancements can be incorporated to further improve its performance and expand its application:

1. Integration of IoT Sensors

Real-time soil and climatic data collected through IoT-enabled sensors can make crop recommendations dynamic and timely.

2. Satellite-Based Remote Sensing

Satellite imagery and NDVI (Normalized Difference Vegetation Index) can enhance environmental analysis and enable region-specific recommendations.

3. Deep Learning Models

Advanced models such as LSTM, CNNs, and Transformer-based architectures can be used to predict long-term climate patterns and improve accuracy.

4. Fertilizer and Irrigation Recommendation

Future systems can suggest ideal fertilizer quantities and irrigation schedules in addition to crop selection.

5. Mobile Application Development

A user-friendly mobile app can make the system accessible to farmers across rural areas.

6. Multi-Crop Suitability Prediction

An enhanced model can recommend multiple crops suitable for the same land, allowing farmers to choose based on economic factors.

Overall, the research concludes that machine learning has immense potential to transform traditional farming practices and pave the way for fully automated, intelligent, and sustainable agriculture.

7. Integration with Real-Time Weather APIs

Future systems can dynamically fetch current and forecasted weather data (rainfall, temperature, humidity) to update crop recommendations instantly.

8. Region-Specific and Soil-Type-Specific Models

Models can be customized for specific regions (district, taluka, village) using localized datasets to improve accuracy.

9. Multi-Language Farmer Support System

The crop recommendation system can be deployed with multilingual support (Marathi, Hindi, English) to help rural farmers understand recommendations.

10. Automated Soil Report Analysis

The system can extract soil nutrient values from government-certified soil health cards (PDFs) using OCR and automatically generate crop suggestions.

11. AI-Based Fertilizer and Pesticide Recommendation

Future upgrades can include:

- Fertilizer amount estimation
- Pest and disease prediction using ML
- Organic fertilizer alternatives

12. Integration with Remote-Sensing Datasets

Satellite data (NDVI, LST, vegetation health index) can improve analysis of crop suitability and soil moisture levels.

13. Climate Change Adaptation System

Machine learning models can be retrained to adapt to shifting climate conditions and suggest climate-resilient crops.

14. AI Chatbot for Farmer Assistance

A chatbot can answer common agricultural queries, recommend crops/fertilizers, and provide region-specific advice**.**

REFERENCES

[1]. **Gupta B.N.S., Uppalapati L.**
Crop Recommender System Using Machine Learning Approach
https://ijerst.org/index.php/ijerst/article/view/381

[2]. **Ghime A.M., Kshirsagar A., Lohakare S., Quadri F., Kare V.**
Crop Recommendation System Using Machine Learning
https://ijarcce.com/papers/crop-recommendation-system-using-machine-learning/

[3]. **Aher R., Chavan S., Singh P.K.**
Agricultural Crop Recommendation System Using Machine Learning
https://iarjset.com/papers/agricultural-crop-recommendation-system-using-machine-learning/

[4]. **Mohapatra B.N., Kale V.**
Crop Recommendation System Using Machine Learning
https://itegam-jetia.org/journal/index.php/jetia/article/view/1186

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 11, November 2025
DOI: 10.17148/IJARCCE.2025.141181

[5]. **Gambhir S., Sharma M., Agarwal K., Kumar K., Kumar L., Chaudhary M.**
Crop Recommendation System Using ML
https://www.ijraset.com/research-paper/crop-recommendation-system-using-ml
[6]. **Kiran Kumar P.N., Bhavya R.A., Dhanushree A.N., Nagarjuna G.R.**
An Efficient Crop Recommendation System Using Machine Learning Mechanisms
https://milestoneresearch.in/JOURNALS/index.php/IJHCI/article/view/149
[7]. **Charishma P.N.S.S.V., Lakshmi S.R., Durga M.V.**
Making Crop Recommendations Using Machine Learning Techniques
https://fringeglobal.com/ojs/index.php/jcai/article/view/jcai_making-crop-recommendations-using-machine-learning-techniqu
[8]. **Paithane P.M.**
Random Forest Algorithm Use for Crop Recommendation
https://itegam-jetia.org/journal/index.php/jetia/article/view/906
[9]. **Sonawane M., Shaikh M., Dange A., Salve D., Kaklij R.**
Crop Recommendation System Using ML Algorithms
https://www.ijarsct.co.in/Paper14284.pdf
[10]. **Bhargavi S., Srinivasan J.**
Crop Recommendation System Using Machine Learning
https://www.ijert.org/crop-recommendation-system-using-machine-learning
[11]. **Sam S., D'Abreo S.M.**
Crop Recommendation With Machine Learning: Leveraging Environmental and Economic Factors (arXiv)
https://arxiv.org/abs/2505.21201
[12]. **Turgut Ö., Kok I., Ozdemir S.**
AgroXAI: Explainable AI-Driven Crop Recommendation System for Agriculture 4.0 (arXiv)
https://arxiv.org/abs/2412.16196
[13]. **Maazallahi A., Thota S., Kondaboina N., Muktineni V.**
Naïve Bayes and Random Forest for Crop Yield Prediction (arXiv)
https://arxiv.org/abs/2404.15392

