

AgriSmart: A Machine Learning-Based Smart Crop Recommendation System for Sustainable Agriculture

Ms.Sheetal Misal¹, Ms.Shweta Jarad², Ms.Manisha Mundhe³

¹Student At Department of Deogiri College of Engineering, Aurangabad, India

²Student At Department of MCA Trinity Academy of Engineering, Pune, India

³Professor At Department of Deogiri College of Engineering, Aurangabad, India

Abstract—Sustainable agriculture demands innovative approaches to optimize crop yield while minimizing resource usage and environmental impact. This paper presents AgriSmart, an intelligent, machine learning-based crop recommendation system specifically designed for Indian agricultural conditions. By leveraging Random Forest algorithms and enriched agronomic datasets, AgriSmart accurately predicts the most suitable crop based on key soil nutrients (N, P, K), pH, temperature, humidity, and rainfall. The system integrates a user-friendly, web-based interface equipped with interactive data visualizations, voice assistance, and chatbot support to deliver real-time, actionable insights to farmers with varying levels of digital literacy. Experimental evaluation demonstrates significant improvements in prediction accuracy over baseline models, reinforcing AgriSmart's potential to enhance decision-making, increase crop productivity, and support sustainable farming practices across diverse agroclimatic zones in India.

Index Terms—Crop recommendation, Machine learning, Random Forest, Sustainable agriculture, Data visualization, AgriSmart.

I. INTRODUCTION

Agriculture remains the backbone of India's economy, supporting over 50% of the population. However, farmers face challenges such as unpredictable weather, soil degradation, and inefficient resource utilization, leading to suboptimal crop yields. Traditional knowledge, while valuable, may not fully address the complexity of modern agricultural needs, particularly under climate variability.

Machine learning offers powerful tools to analyze vast amounts of agronomic data and derive crop recommendations that are both accurate and context-specific. Despite progress, many existing systems

lack integration of real-time data, usability, or adaptability to local Indian agroclimatic conditions. This paper introduces AgriSmart, a comprehensive crop recommendation system that combines Random Forest-based prediction, a web-based interactive interface, and visualization modules. AgriSmart aims to bridge the gap between advanced analytics and practical farming needs, enabling sustainable and profitable agriculture.

The key contributions of this work are summarized as follows:

- **Development of a Robust Machine Learning Model:** A Random Forest-based predictive model is trained using a comprehensive Indian crop dataset that includes key agronomic features such as nitrogen, phosphorus, potassium, temperature, humidity, pH, and rainfall. The model achieves high accuracy in recommending suitable crops for specific soil and climatic conditions.
- **User-Centric Web-Based Interface:** A dynamic, responsive, and multilingual web application is designed to provide crop recommendations through an intuitive interface. The system supports both textual and voice-based interactions, making it accessible for users with limited digital literacy.
- **Real-Time Visual Analytics:** The platform integrates interactive data visualization modules using advanced charting libraries. These visualizations display model predictions, soil parameter distributions, and crop insights to support better decision-making.
- **Voice-Enabled Chatbot Integration:** A custom-built chatbot assists users in querying the system about crop details, best practices, and prediction

feedback using voice commands enhancing usability, especially in rural environments.

- **Personalization and Data Logging:** AgriSmart collects user input (e.g., name, location, field parameters) to personalize the experience and stores prediction history, enabling future recommendation improvements and user engagement analytics.
- **Snapshot Dashboard and Recommendation Report:** After each prediction, a snapshot is generated, summarizing the predicted crop, relevant insights, and agronomic advice. This improves transparency and provides a basis for informed decision-making.
- **Support for Sustainable Agriculture:** By suggesting optimal crops based on real data, Agri Smart aims to reduce resource wastage (e.g., fertilizer misuse), enhance crop yield, and contribute to sustainable farming practices aligned with local agroclimatic zones.

II. PROBLEM STATEMENT

Despite being a leading agricultural economy, India faces persistent challenges in achieving optimal crop yields due to soil degradation, climate variability, and limited access to scientific tools for crop selection. Existing crop recommendation systems often lack adaptability to local agroclimatic conditions, fail to integrate both soil and weather parameters effectively, and offer limited user-friendly interfaces for non-technical users such as small-scale farmers.

There is a critical need for a smart, accessible, and accurate crop recommendation system that considers key soil nutrients and climatic conditions while providing intuitive, real-time, and actionable insights to farmers. Such a system must support sustainable agricultural practices, improve productivity, and minimize environmental impact.

This research aims to develop Agri Smart, a machine learning-based crop recommendation system tailored to Indian conditions, which integrates predictive analytics, visualization, and user-centric interaction modules to empower farmers with evidence-based decisions.

III. OBJECTIVES

- To develop a machine learning-based model that recommends the most suitable crop based on soil and climate parameters.
- To design a user-friendly web-based interface accessible to farmers.
- To integrate visualization tools and a chatbot for better usability and interpretation.
- To promote sustainable farming through data-driven decision-making.

VI. RELATED WORK

Integration of machine learning in agriculture has accelerated in recent years, with studies focusing on crop recommendation systems based on environmental data.

Patel and Parmar [1] proposed a decision tree-based system using soil nutrient parameters (N, P, K). However, it lacked dynamic weather variables, limiting adaptability.

Kale and Vikhe [2] employed K-Nearest Neighbors (KNN) and Support Vector Machines (SVM) for crop prediction using temperature, rainfall, and humidity. Their model's performance declined with large datasets due to lack of feature scaling and selection. Singh and Mishra [3] used Random Forests to improve prediction accuracy but lacked a user-friendly interface.

Sharma and Kumar [4] emphasized intelligent systems in sustainable agriculture, advocating for models optimizing yield and conserving resources.

Ramesh et al. [5] applied CNNs for crop disease detection, showcasing AI's potential in plant health monitoring.

Mishra and Sinha [6] combined IoT sensors with Naive Bayes classifiers for real-time crop recommendation.

Gupta et al. [7] showed ensemble learning improved prediction reliability, while Ahmed and Mehta [8] used LSTM networks for crop yield forecasting.

Khan and Verma [9] proposed a smart crop recommendation engine integrating soil and weather APIs with XGBoost, achieving high accuracy across regions. Desai et al. [10] introduced a fuzzy logic-based crop advisory system accounting for vagueness in farmer input and climatic uncertainty.

Joshi and Bansal [11] provided a comprehensive review of ML algorithms for precision agriculture and recommended hybrid models for improved generalization.

A. Gaps in Existing Research

Key limitations identified in previous research include:

- Lack of integration between soil and real-time climatic data.
- Limited usability for farmers with low digital literacy.
- Absence of interpretability and visualization tools.
- Inadequate regional adaptation and language support.

AgriSmart addresses these gaps with integrated analytics, visualization, and user-centered design for Indian agriculture.

V. METHODOLOGY

A. Dataset Description

Agri Smart uses a compiled dataset containing essential soil and climatic parameters relevant to Indian farming. The primary data source is the publicly available Crop_recommendation.csv dataset, enriched with features commonly used in Indian agriculture.

Key features include:

- Soil Parameters: Nitrogen (N), Phosphorus (P), Potassium (K), pH level.
- Climatic Parameters: Temperature, rainfall, humidity.

The dataset contains 2,200+ entries across 22 crop classes, widely cultivated in India.

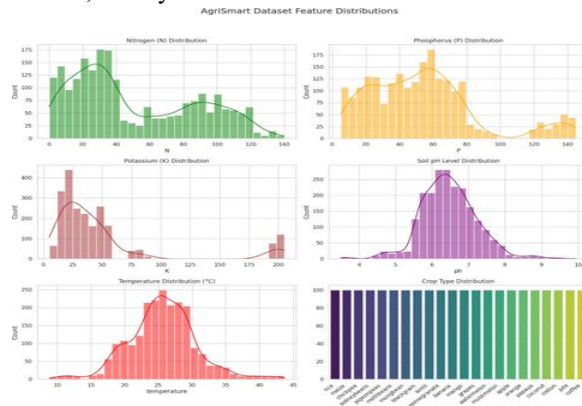


Fig. 1: Distribution of Crop Types in Dataset

B. Data Preprocessing

Before using the data for training the crop recommendation model, we performed some important cleaning and preparation steps:

- Checked the dataset carefully and confirmed there were
- no missing values or incomplete records.
- Removed any duplicate entries to avoid bias in the model.
- Converted all data into the correct formats so the machine learning model can process it easily.
- Scaled and normalized numerical features like soil nutrients and temperature to keep their values consistent.
- Split the dataset into training and testing parts to evaluate the model's accuracy later.

C. Feature Engineering

To help the model understand the data better and improve predictions, we created new useful features from the existing data:

- Calculated nutrient ratios such as Nitrogen to Phosphorus (N:P) and Nitrogen to Potassium (N:K). These ratios indicate the balance of soil nutrients important for crop growth.
- Combined temperature and humidity values by multiplying them to capture the effect of weather conditions on crops.
- Derived interaction features like rainfall multiplied by humidity to represent how climate factors work together.
- Added categorical labels for soil type and climate zone based on original numeric values, which helps the model learn patterns more easily.

These new features helped the machine learning model capture the complex relationships between soil, climate, and crop suitability more effectively.

D. Random Forest Model

AgriSmart utilizes a Random Forest classifier to predict the most suitable crop. This ensemble method aggregates predictions from multiple decision trees using majority voting:

$\hat{y} = \text{mode}\{h_1(x), h_2(x), \dots, h_T(x)\}$ The Gini impurity is used for split quality:

$$G = 1 - \sum_{i=1}^n p_i^2$$

where p_i is the class probability for class i . Model hyperparameters:

- Number of trees: 100
- Maximum depth: 15
- Minimum samples per leaf: 5

E. Evaluation Metrics

The model was evaluated using:

- Accuracy: Overall correctness of predictions.
- Precision and Recall: Class-specific performance.
- F1-score: Balance between precision and recall.
- Confusion Matrix: Visualizing correct vs. incorrect pre- dictions.

VI. SYSTEM DESIGN AND IMPLEMENTATION

A. Architecture

Agri Smart employs a modular architecture designed to prioritize user experience and system scalability.

The system is divided into three primary layers:

- **Input Layer:** This layer gathers essential agricultural parameters from users through an intuitive form inter- face. Inputs include soil nutrients (Nitrogen, Phosphorus, Potassium), temperature, humidity, pH level, and rainfall, ensuring comprehensive data collection relevant to crop suitability.
- **Prediction Layer:** The collected inputs are processed by a machine learning model specifically, a Random Forest classifier which is hosted on a Flask backend server. This layer performs the core task of analyzing the data to generate accurate crop recommendations tailored to the given environmental conditions.
- **Visualization & Interaction Layer:** Results and insights from the prediction layer are presented through an interactive front-end built using Streamlit and enhanced with HTML/CSS for improved user engagement. This layer includes dynamic visualizations, a chatbot to assist users with queries, and voice feedback to improve accessibility and provide a more engaging experience.

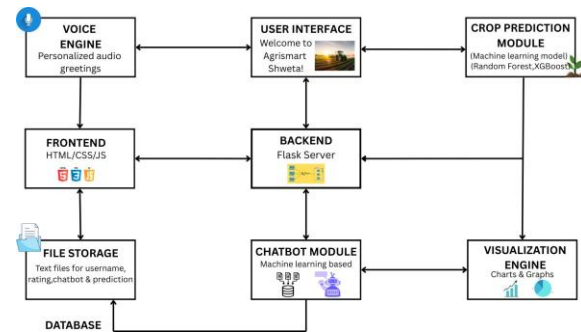


Fig. 2: AgriSmart System Architecture

B. User Interface

The AgriSmart user interface is designed to be both visually appealing and highly interactive, enhancing user engagement and ease of use. Key features include:

- An animated home screen with a personalized greeting that addresses the user by name, creating a welcoming experience.
- A user-friendly input form featuring sliders and dropdown menus, enabling smooth and intuitive data entry for soil and climatic parameters.
- Clear presentation of crop prediction results, including a confidence score and an illustrative crop image to help users better understand the recommendation.
- Interactive visual analytics powered by Plotly, offering bar, pie, and radar charts to provide insightful data interpretations.
- An integrated chatbot designed to answer crop-related queries, improving user support.
- A voice assistant that guides and welcomes users, increasing accessibility, especially for users with limited literacy.

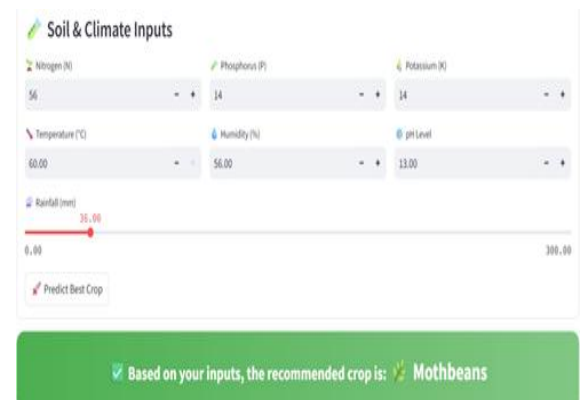
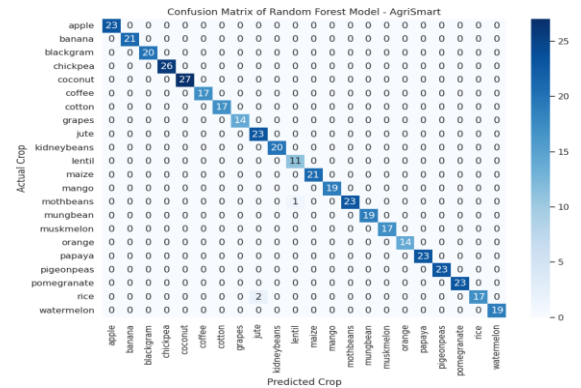


Fig. 3: UI Screenshots: Input and Prediction Page

C. Implementation Details

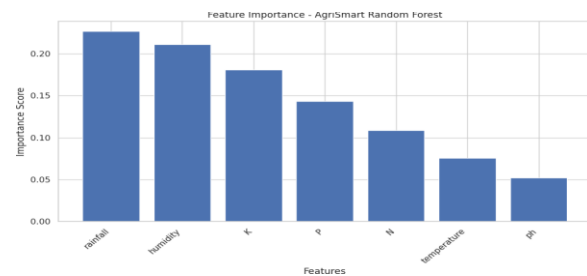
AgriSmart integrates multiple technologies to deliver a smart crop recommendation system:

- **Backend & Frontend:** Built using Flask and Streamlit for seamless backend logic and interactive frontend interfaces.
- **Machine Learning:** Utilizes a Random Forest Classifier implemented via scikit-learn to predict suitable crops based on soil and climatic features.
- **Visualization:** Employs Plotly and Matplotlib for both interactive and static visualizations to support user decisions.
- **Web Design:** Developed using HTML, CSS, and JavaScript, offering animated interfaces, personalized greetings, and chatbot support.
- **Data Storage:** Uses local CSV/text files to save user inputs and predictions; optionally supports SQLite/PostgreSQL for structured data storage.
- **Additional Features:** Includes a voice assistant and chatbot. The interface starts with a name input and offers personalized interactions. Future plans include weather API integration and mobile/offline support.



C. Feature Importance

Feature importance indicates the contribution of each input variable (e.g., soil nutrients, temperature) to the model's predictions. In the Random Forest model, it helps identify which soil and climatic parameters most influence crop recommendation. Understanding feature importance enables better interpretation of the model and can inform farmers and agronomists about the critical factors affecting crop suitability.



VII. RESULTS AND DISCUSSION

A. Model Evaluation

TABLE I: Performance Comparison of Crop Recommendation Models

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	78.5%	79.0%	77.8%	78.4%
SVM	81.2%	81.5%	80.7%	81.1%
Random Forest	88.7%	89.1%	88.0%	88.5%

B. Confusion Matrix

A confusion matrix is a performance measurement tool for classification models that shows the number of correct and incorrect predictions broken down by each class. It provides detailed insight into how well the model distinguishes between different crop types by displaying true positives, false positives, true negatives, and false negatives. This helps identify specific classes where the model may confuse similar crops, guiding further improvements.

D. Error Analysis

Most misclassifications in AgriSmart occurred between crops with overlapping soil nutrient and climatic profiles such as maize and sorghum due to their similar NPK levels, pH range, and temperature requirements. This overlap caused confusion for the model in distinguishing optimal crop recommendations. Enhancing the dataset with more granular regional data and additional agronomic factors (e.g., pest resistance, crop rotation history) could reduce these errors.

E. User Feedback

Feedback from initial users, including small-scale farmers, has been overwhelmingly positive. Users appreciated AgriSmart's intuitive interface, which simplifies the crop selection process by providing clear and easily understandable visualizations. The

voice assistant feature was particularly well-received, as it offers hands-free guidance and enhances accessibility for users with limited literacy or who prefer auditory instructions. This combination of features has helped foster trust and confidence in the system, encouraging farmers to rely on AgriSmart for making informed agricultural decisions. Additionally, early feedback has identified areas for further improvement, such as expanding language support and enhancing mobile accessibility, which will be addressed in future iterations of the system.

F. Limitations

- The system currently lacks integration with real-time weather updates and soil sensor data, which limits its ability to provide dynamic, up-to-date recommendations.
- Offline functionality is not available, reducing accessibility for farmers in areas with poor or no internet connectivity.
- The chatbot and user interface are only available in English, which may restrict usability for farmers who speak regional languages.
- The dataset used focuses mainly on common crops; data for less common or region-specific crops is limited, impacting recommendation accuracy for those crops.

VIII. CONCLUSION AND FUTURE WORK

A. Conclusion

AgriSmart offers a robust, farmer-friendly crop recommendation system tailored for Indian agriculture. By leveraging a well-tuned Random Forest model and essential soil and climate features, it delivers accurate, data-driven crop suggestions. The system bridges advanced AI with practical farming via a user-friendly interface, incorporating visual analytics, voice assistance, and a chatbot for accessibility.

Evaluation results show AgriSmart outperforms traditional models in accuracy and usability. While effective, current limitations include limited language support, lack of real-time sensor integration, and internet dependency. Addressing these in future versions will improve scalability and field adoption.

B. Future Work

- **Real-Time Data Integration:** Incorporating live weather APIs, satellite data, and IoT-based soil sensors to make crop recommendations more adaptive and location-specific.
- **Multilingual and Voice-Based Interaction:** Expanding support for regional languages to improve accessibility among non-English-speaking farmers, along with enhanced voice interaction capabilities.
- **Offline and Mobile Application Support:** Developing a lightweight mobile version with offline functionality for farmers in low-connectivity regions, using edge computing for localized predictions.
- **Pest and Disease Prediction:** Extending the system to recommend preventive measures or alternative crops in response to predicted pest outbreaks or disease patterns using CNN and time-series models.
- **Field Validation and Dataset Expansion:** Collaborating with agricultural universities, Krishi Vigyan Kendras (KVKs), and government agencies to gather real-world feedback, refine model parameters, and expand the dataset with region-specific crops and cultivation practices.
- **Personalized Crop Calendars:** Introducing crop growth timelines, irrigation alerts, and fertilizer reminders customized to user-specific predictions and geolocation.
- **Sustainability Metrics and Yield Forecasting:** Incorporating modules to estimate potential yield and resource efficiency (e.g., water and fertilizer usage), aligning with the goals of precision and climate-smart agriculture.

REFERENCES

- [1] Patel and K. Parmar, Crop Recommendation System to Maximize Yield using Machine Learning Technique, IJIRCCE, vol. 5, no. 1, pp. 232–235, 2017.
- [2] S.Kale and S.Vikhe, ML Based Crop Recommendation Using Climatic Parameters, IJSRCSEIT, vol. 4, no. 1, pp. 88–93, 2019.
- [3] R. Singh and R. Mishra, Crop Recommendation Using Machine Learning Algorithms, IJCA, vol. 975, no. 8887, pp. 1–5, 2021.

- [4] N. Sharma and R. Kumar, Smart and Sustainable Agriculture Using AI and IoT, IJAIS, vol. 13, no. 2, pp. 15–29, 2022.
- [5] S. Ramesh et al., Crop Disease Detection Using CNN, IJACSA, vol. 11, no. 5, pp. 400–407, 2020.
- [6] R. Mishra and A. Sinha, IoT-Based Smart Farming with Naive Bayes, ICICCS, pp. 400–405, 2018.
- [7] P. Gupta et al., Ensemble Learning-Based Crop Recommendation, JAISCR, vol. 11, no. 3, pp. 210–219, 2021.
- [8] M. Ahmed and R. Mehta, Deep Learning for Crop Yield Prediction Using LSTM, CEIA, vol. 205, p. 107597, 2023.
- [9] A.Khan and R. Verma, Smart Crop Recommendation System Using XG-Boost and Weather API Integration, Journal of Agricultural Informatics, vol. 13, no. 2, pp. 101–110, 2022.
- [10] P. Desai, K. Jadhav, and V. Gawande, A Fuzzy Logic-Based Decision Support System for Crop Advisory in Uncertain Climate, Computers and Electronics in Agriculture, vol. 199, p. 107156, 2023.
- [11] S. Joshi and A. Bansal, A Review on Machine Learning Techniques in Precision Agriculture: Challenges and Opportunities, Artificial Intelligence in Agriculture, vol. 7, pp. 12–25, 2023.