AGROSMART: A MACHINE LEARNING-DRIVEN SYSTEM FOR PREDICTIVE AGRICULTURE

Karuna Girase[1], Divya Khairnar[2], Ashwini Deore[3], Ashwini Borse[4], Pradnya Patil[5]

[1] *Engineering Student, Department of Computer Science and Engineering, R. C. Patel Institute of Technology Shirpur-425405,*

Dist-Dhule, Maharashtra, India

[2] *Engineering Student, Department of Computer Science and Engineering, R. C. Patel Institute of Technology Shirpur-425405,*

Dist-Dhule, Maharashtra, India

[3] *Engineering Student, Department of Computer Science and Engineering, R. C. Patel Institute of Technology Shirpur-425405,*

Dist-Dhule, Maharashtra, India

[4] *Engineering Student, Department of Computer Science and Engineering, R. C. Patel Institute of Technology Shirpur-425405,*

Dist-Dhule, Maharashtra, India

[5] *Engineering Student, Department of Computer Science and Engineering, R. C. Patel Institute of Technology Shirpur-425405,*

Dist-Dhule, Maharashtra, India

Email: karunagirase2005@gmail.com[1], divyakhairnar7502@gmail.com[2], deoreashwini1913@gmail.com[3], borseashwini81@gmail.com[4], patilpradnya.8031@gmail.com[5]

ABSTRACT

*Agriculture increasingly depends on data-driven methods to improve productivity and resource utilization. This paper presents **AgroSmart**, a machine learning-based decision support system developed to assist farmers in crop selection, fertilizer planning, and crop yield prediction. The proposed system analyzes soil and environmental parameters such as nitrogen, phosphorus, potassium content, soil pH, rainfall, temperature, and humidity.*

*For crop recommendation, a **Random Forest** classifier is used to identify the most suitable crop based on input conditions. For yield prediction, a **Linear Regression** model estimates expected production from historical patterns and current parameters. The system is implemented in Python using a Flask backend with a user-friendly web interface that allows users to enter field data and receive recommendations in real time.*

The proposed platform can help reduce inefficient fertilizer use, support informed farm decisions, and improve planning at the field level. AgroSmart demonstrates how machine learning can be applied in practical farming environments to support more efficient and sustainable agriculture.

Keywords: *Predictive Agriculture, Machine Learning, Crop Recommendation, Yield Prediction, Random Forest, Linear Regression, Smart Farming.*

1. INTRODUCTION

Agriculture plays an essential role in economic growth and food security, especially in developing countries where a large portion of the population depends on farming for income and livelihood. However, many farmers still make cultivation decisions based on traditional practices, local experience, or generalized advice. While such methods are useful in some situations, they may not always match current soil conditions, weather patterns, or crop requirements. As a result, farmers may face lower productivity, higher input costs, and uncertain yields.

With the growth of digital technologies, agriculture can now benefit from data-driven decision making. Machine Learning (ML) techniques are capable of analyzing agricultural datasets and identifying relationships between soil properties, weather conditions, and crop performance. These methods can be used to improve decisions related to crop recommendation, fertilizer management, and yield prediction.

One common challenge in farming is choosing the right crop for a specific field. If the selected crop is not suitable for the soil or climate, production may decrease and resources such as water and fertilizer may be wasted. Similarly, excessive or incorrect fertilizer application can increase cost and negatively affect soil health. Climate variability also adds further uncertainty to agricultural planning.

To address these issues, this research proposes **AgroSmart**, an integrated machine learning-based agricultural support system. The platform accepts inputs such as nitrogen, phosphorus, potassium levels, soil pH, rainfall, temperature, and humidity. Based on these parameters, it recommends suitable crops, estimates crop yield, and provides fertilizer-related guidance. The objective of the system is to help farmers make informed decisions, reduce resource wastage, and improve farm productivity through practical use of intelligent technology.

2. LITERATURE REVIEW

The use of machine learning in agriculture has increased significantly in recent years due to the availability of agricultural data and the need for better farm management practices. Researchers have applied data-driven techniques to improve crop planning, estimate production, and optimize the use of inputs such as fertilizer and water. These studies show that intelligent systems can support faster and more accurate agricultural decisions.

Several machine learning algorithms have been used for agricultural prediction tasks. Classification models such as **Random Forest** and **Support Vector Machine (SVM)** are commonly used for crop recommendation based on soil nutrients, pH, rainfall, and climatic conditions. Regression techniques are widely used for yield prediction by analyzing historical production data and environmental factors. In many cases, these models provide better accuracy than traditional statistical approaches.

Although previous work has produced useful results, many available systems are designed for only one specific task. Some applications focus only on crop recommendation, while others are limited to yield estimation or fertilizer advice. Because of this separation, farmers may need to use different tools to obtain complete guidance, which reduces convenience and practical usability.

Another limitation of conventional agricultural planning is the dependence on manual judgment or fixed recommendations. Such methods may not respond effectively to changing weather conditions, regional soil differences, or variations in crop demand. This creates a need for systems that can adapt recommendations using current data.

To address these limitations, the proposed **AgroSmart** system combines multiple agricultural services within a single platform. It integrates crop recommendation, yield prediction, and fertilizer guidance using machine

learning techniques. This unified approach is intended to improve usability, reduce decision complexity, and provide farmers with complete support through one application.

3. METHODOLOGY

3.1 System Overview

The proposed AgroSmart system is developed as a step-by-step pipeline that accepts agricultural input data, processes it, applies trained machine learning models, and returns useful recommendations to the user. The complete workflow begins with dataset collection and ends with prediction results displayed through a web application.

The main stages of the system are:

- Collection of agricultural dataset
- Data cleaning and preprocessing
- Selection of useful input features
- Training and testing of machine learning models
- Prediction generation based on user inputs
- Display of results through the web interface

This structured workflow helps in maintaining prediction accuracy and smooth system operation.

3.2 Data Processing and Feature Selection

The dataset used for this project contains important agricultural parameters such as nitrogen (N), phosphorus (P), potassium (K), soil pH, rainfall, temperature, and humidity. These factors directly influence crop growth and productivity.

Before model training, the dataset is preprocessed to improve quality and consistency. The preprocessing stage includes:

- Removing missing or duplicate records
- Correcting inconsistent values, if present
- Scaling or normalizing numerical features
- Converting categorical data into machine-readable format when required
- Selecting relevant features for prediction tasks

Proper preprocessing improves model performance and helps generate more reliable outputs.

3.3 Machine Learning Models

Two machine learning models are used in the AgroSmart system for different prediction tasks.

3.3.1 Crop Recommendation using Random Forest

For crop recommendation, the Random Forest algorithm is used as a classification model. Random Forest creates multiple decision trees and combines their outputs to improve prediction accuracy. It performs well on structured datasets and can handle nonlinear relationships among features.

Based on the values of N, P, K, pH, rainfall, temperature, and humidity, the model predicts the most suitable crop for cultivation under the given conditions.

3.3.2 Yield Prediction using Linear Regression

For yield estimation, Linear Regression is used. This model predicts crop yield as a continuous numerical value by identifying relationships between input parameters and historical output data.

Linear Regression is simple, easy to interpret, and useful as a baseline model for production forecasting.

3.4 System Implementation

The AgroSmart platform is implemented using Python for data processing and model development. The backend of the application is developed using Flask, which handles user requests and connects the trained models with the web interface. The frontend is created using HTML and CSS to provide a simple and user-friendly interface.

Users enter agricultural parameters through an input form. After submission, the system processes the values and returns crop recommendations and yield predictions in real time.

4. RESULTS

The proposed system was evaluated using sample agricultural datasets to assess its prediction performance. The results show that the model is capable of producing relevant outputs based on user-provided input parameters.

The output includes:

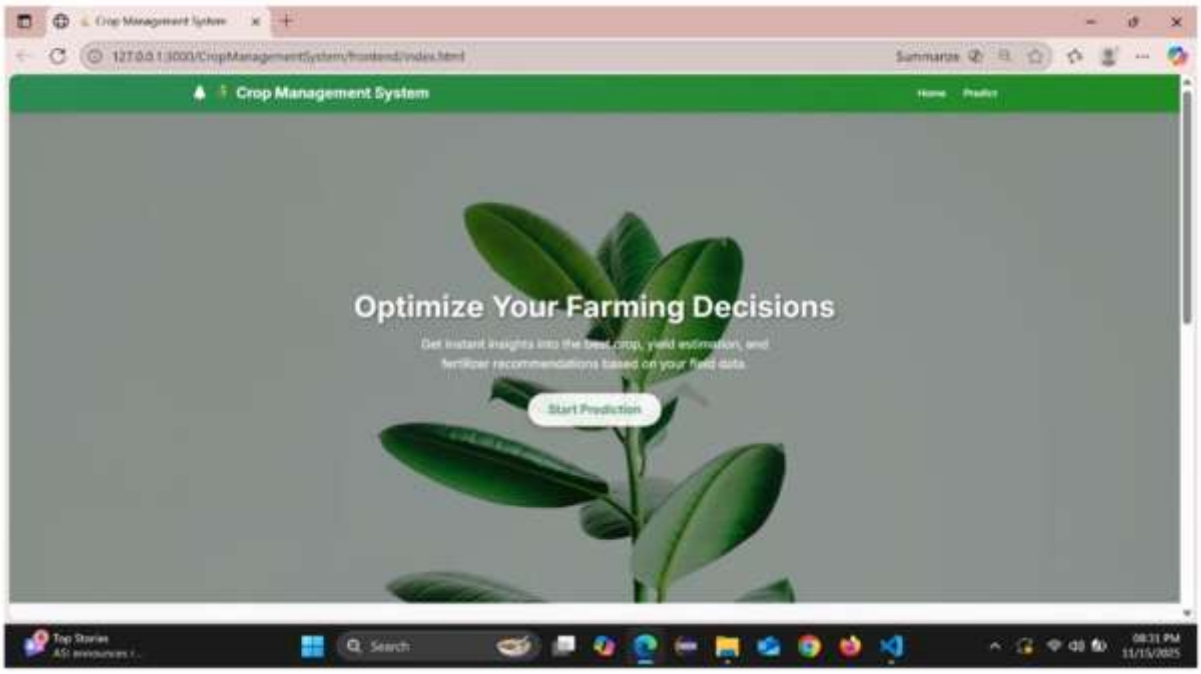

Fig 1.1 Home page

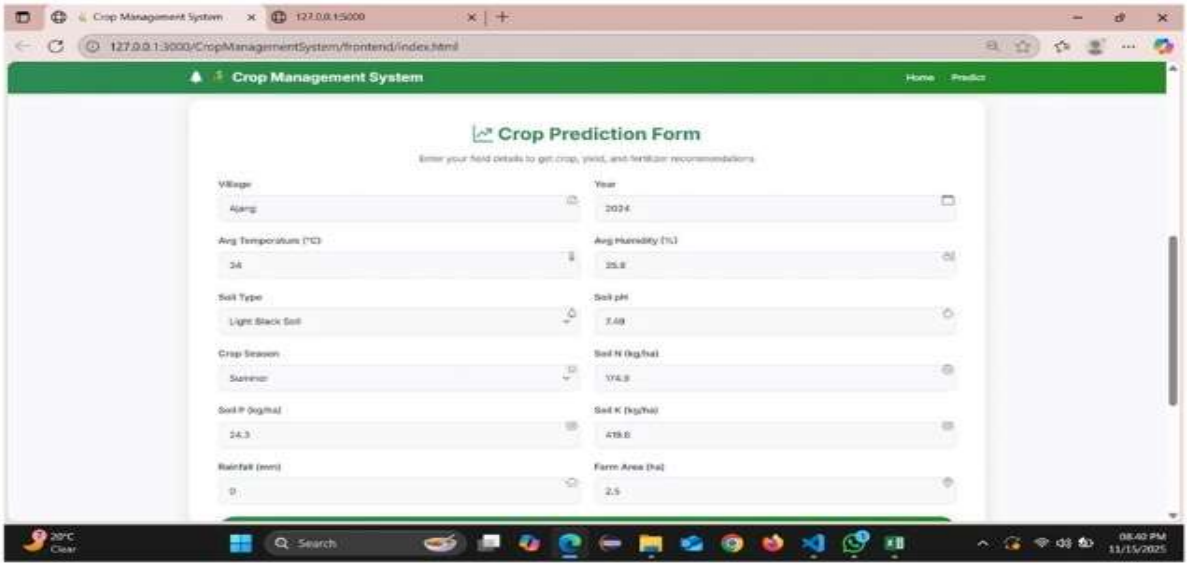

Fig 1.2 Prediction Form

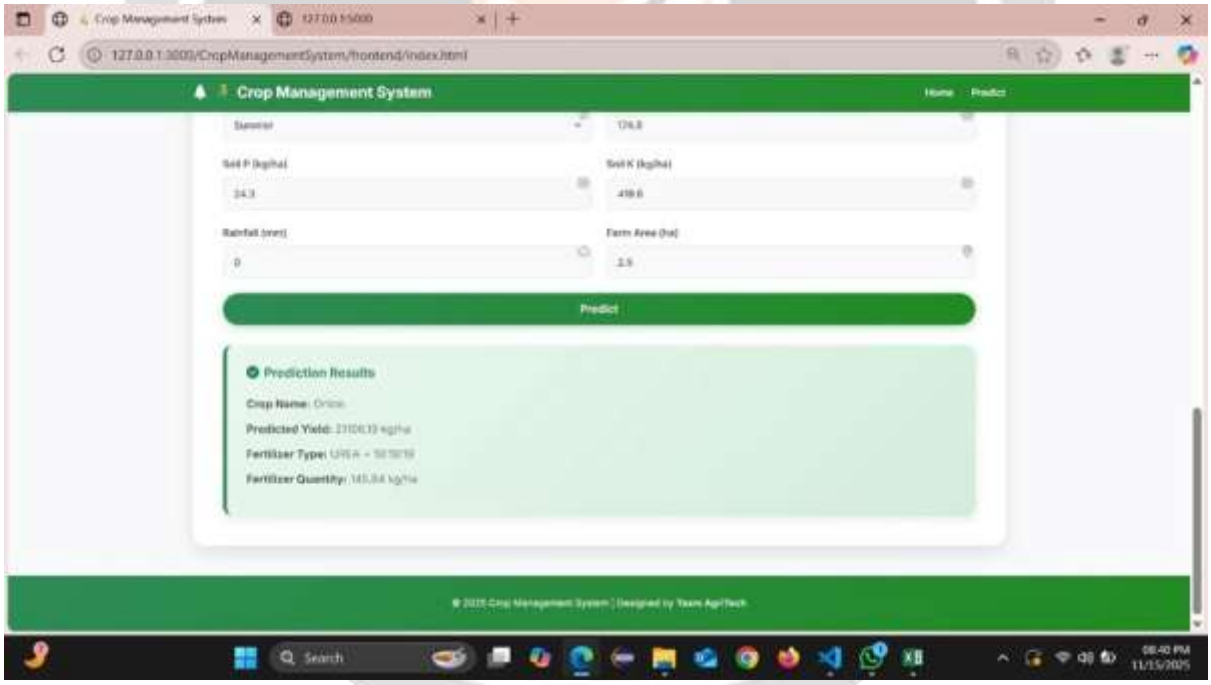

Fig 1.3 Output

Recommended crop suitable for given conditions

Estimated crop yield

Suggested fertilizer type and quantity

The interface is designed to be simple and intuitive, enabling users to interact with the system without technical complexity. The results indicate that the integration of machine learning models can effectively support agricultural decision-making.

5. ADVANTAGES AND APPLICATIONS

5.1 Advantages

The proposed system offers several benefits, including improved accuracy in crop selection, efficient use of fertilizers, and better estimation of crop yield. It reduces dependency on manual decision-making and promotes scientific farming practices.

5.2 Applications

The system can be applied in smart farming environments, agricultural advisory services, and educational platforms. It can also support government initiatives aimed at improving agricultural productivity and sustainability.

6. CONCLUSION

This research presents AgroSmart, an intelligent system that leverages machine learning techniques to enhance agricultural decision-making. By combining crop recommendation, yield prediction, and fertilizer optimization into a single platform, the system provides a comprehensive solution to common farming challenges.

The implementation demonstrates that machine learning can play a significant role in transforming traditional agriculture into a more efficient and data-driven process. The system not only improves productivity but also contributes to sustainable resource management.

7. FUTURE SCOPE

Future improvements may include integration with real-time data sources such as IoT-based soil sensors and weather APIs. Developing a mobile application and expanding the dataset to cover multiple regions can further enhance the usability and scalability of the system. Advanced techniques such as deep learning can also be incorporated for disease detection and precision farming.

8. REFERENCES

1. Liakos, K. et al., “Machine Learning in Agriculture: A Review,” Sensors Journal.
2. Scikit-learn Documentation, https://scikit-learn.org
3. Python Software Foundation, https://www.python.org
4. ICAR Agricultural Reports
5. Kaggle Agricultural Datasets
6. W3Schools, HTML/CSS Tutorials