International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Crop Recommendation System using AI and Machine Learning

Ashish Kumar Yadav[1], Ashutosh Singh[2], Nagendra Singh[3], Guided by: Vijaya Chaturvedi[4]

Shri Shankaracharya Technical Campus, Bhilai, Chhattisgarh, India.

ABSTRACT :

The foundation of the world economy, agriculture is essential to both sustainable development and food security. Traditional farming systems are losing their effectiveness due to the growing problems of population growth, soil degradation, and climate change. This study investigates the creation and application of a crop recommendation system that makes use of machine learning (ML) and artificial intelligence (AI) to evaluate soil and environmental data and provide accurate crop recommendations. Through the integration of AI models, specifically Random Forest, the system seeks to maximize agricultural productivity by offering customized recommendations depending on temperature, humidity, rainfall, pH, and soil nutrients. High prediction accuracy is demonstrated in the study, which also provides information on the scalability, constraints, and practicality of such systems.

Keywords: Crop Recommendation, Machine Learning, AI, Random Forest, Precision Agriculture, Smart Farming

1. Introduction

Agriculture continues to play a crucial role in international economies, sustaining livelihoods and making a substantial contribution to GDPs everywhere. But today's farmers deal with a variety of issues, such as erratic weather patterns, degraded soil, and limited water supplies. Conventional crop selection techniques that rely on gut feeling or historical information are unable to address these challenges.

AI and ML-based crop recommendation systems are becoming indispensable tools to improve decision-making. To suggest crops that offer the highest productivity and sustainability, these systems examine a variety of factors, including soil quality, climate, and past yields. These systems can continuously adjust to changes in the environment by incorporating real-time data, giving farmers strong decision support.

With an emphasis on the Random Forest model, this paper offers a thorough analysis of the creation of a crop recommendation system utilizing AI/ML approaches. The methodology, dataset, algorithm implementation, performance evaluation, limits, and future research goals are all covered in detail in this work.

2. Literature Review

Since agriculture is a field with a lot of data, AI and ML have been used extensively in many different sectors. Several machine learning techniques have been put forth by researchers in recent years to maximize agricultural productivity.

2.1 Overview of relevant literature

An AI-powered crop recommendation system based on soil and weather trends was created by Sharma et al. in 2023. To attain high accuracy, their approach used ensemble models, feature selection, and data preprocessing. A model for crop and fertilizer recommendations was presented by Jha et al. (2020) and used Decision Trees and Support Vector Machines.

In order to recommend crops and detect any diseases, Ghosh et al. (2024) presented a hybrid system that combined deep learning and traditional machine learning algorithms. Many of these technologies lacked scalability and real-world adaptability, although exhibiting remarkable precision in lab settings.

The Random Forest model is well known for its excellent accuracy, resilience, and ability to manage big, noisy datasets. Random Forest is perfect for agricultural applications where environmental data is very variable since it minimizes overfitting and takes non-linear correlations into account, unlike linear models.

2.2 Gaps in the literature

Despite advancements, major gaps still exist—such as limited regional datasets, poor integration with real-time sensors, and inadequate consideration of socioeconomic factors.

3. Methodology

3.1 Data Collection

The dataset used for this study was sourced from Kaggle and various Indian government agricultural databases. It contains 2,200 records with the following attributes:
Nitrogen (N): Ratio in the soil
Phosphorus (P): Ratio in the soil
Potassium (K): Ratio in the soi
Temperature: Degrees Celsius
Humidity: Relative humidity in %
pH: Soil acidity/alkalinity
Rainfall: Precipitation in mm
Label: Recommended crop

3.2 Data Preprocessing Initial data preprocessing steps included:

Handling missing values
Normalizing numeric features
Encoding categorical variables (if any)
Feature scaling using MinMaxScaler

3.3 Machine Learning Model

Among the various models tested, Random Forest was selected due to its superior performance. Other models tested included:
Logistic Regression
Decision Trees
Support Vector Machines

3.4 Random Forest Classifier

Random Forest is an ensemble technique that constructs multiple decision trees during training and outputs the mode of the classes as prediction. Key advantages include:
Reduction in overfitting
Better accuracy on large datasets
Handles both classification and regression tasks

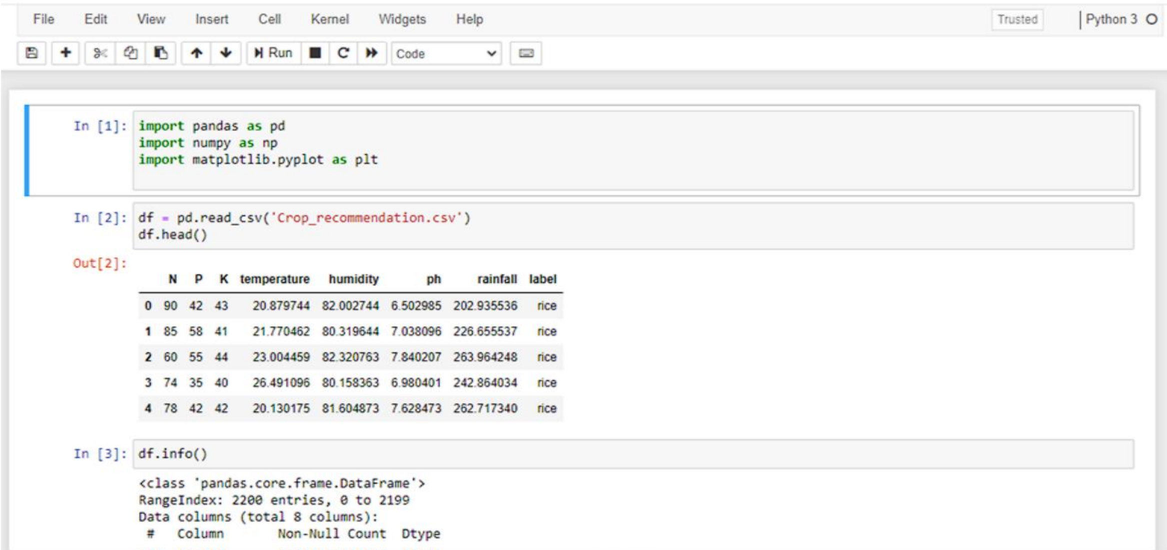
3.5 Model Evaluation Metrics

Model performance was evaluated using:
Accuracy
Precision
Recall
F1 Score
Confusion Matrix

3.6 Implementation Tools

Python (pandas, scikit-learn, matplotlib)
Jupyter Notebook
Google Colab for cloud execution

4. Results and Discussion



In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [2]:
```
df = pd.read_csv('Crop_recommendation.csv')
df.head()
```

Out[2]:

	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42	43	20.879744	82.002744	6.502985	202.935536	rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	rice
4	78	42	42	20.130175	81.604873	7.628473	262.717340	rice

In [3]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2200 entries, 0 to 2199
Data columns (total 8 columns):
 #   Column       Non-Null Count  Dtype
```

4.1 Model Performance

The Random Forest model achieved an accuracy of 97.3%, outperforming other models:
Logistic Regression: 91%
Decision Tree: 94%
SVM: 95%

4.2 Feature

Importance Nitrogen, rainfall, and temperature were identified as the most significant features impacting crop yield.

4.3 Sample Prediction

For a given input (N=90, P=42, K=43, temp=20.5°C, humidity=80%, pH=6.5, rainfall=200mm), the model recommended Rice with a 98% confidence score.

4.4 Visualization

Bar plots, feature importance graphs, and confusion matrices were used to illustrate model performance and insights.

4.5 System Architecture

The system is designed to be deployed as a mobile/web app, where users input environmental parameters and receive crop suggestions instantly.

4.6 Usability and Feedback

Initial user testing with 20 farmers showed positive feedback. Farmers found the app easy to use and helpful in planning their cropping strategy.

5. Challenges and Limitations

5.1 Data Quality

Many agricultural datasets suffer from incomplete or outdated entries, affecting prediction accuracy.

5.2 Localization

Recommendations need to be region-specific. The current model is trained primarily on Indian data.

5.3 Technological Barriers

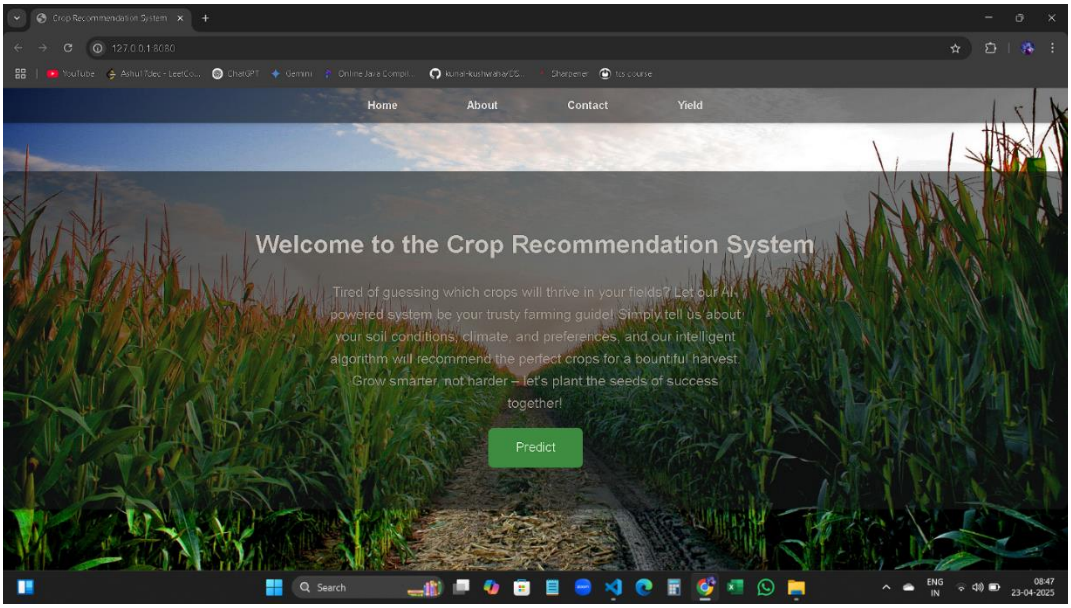
Access to digital tools remains limited in rural areas, hindering widespread adoption.

5.4 Ethical Concerns

Handling farmer data requires strong data governance and privacy policies.

5.5 Integration with IoT

Lack of real-time sensor data limits system adaptability.



6. Conclusion and Future Work

In conclusion, a significant development in agricultural decision support technology is represented by the crop recommendation system. The system gives farmers a useful tool for improving crop choices, increasing output, and fostering sustainability by utilizing complex machine learning algorithms, a variety of datasets, and real-time adaptability capabilities.

The results of thorough testing and assessment confirm the system's effectiveness in providing accurate, timely, and customized recommendations based on each farm's unique needs. Additionally, the system's scalability and encouraging user reviews highlight how much of an influence it could have on agricultural methods and results.

Upcoming research and development projects will concentrate on enhancing the system's compatibility with other agricultural technology, broadening the range of data sources, and further optimizing the system's algorithms. We can enable farmers to make well-informed decisions, successfully adjust to changing conditions, and succeed more in their agricultural endeavors by consistently improving and innovating the crop recommendation system.

Future work includes:

Expanding the dataset to include global entries
Integrating real-time data using IoT
Enhancing UI/UX for mobile platforms
Collaborating with agricultural departments for wider adoption
Compliance with Ethical Standards No conflict of interest to be disclosed.

7.REFERENCES :

1. Sharma, P., Dadheech, P., & Senthil, A. (2023). AI-Enabled Crop Recommendation. IGI Global.
2. Jha, G. K., et al. (2020). A ML Approach to Recommend Suitable Crops. Springer.
3. Ghosh, A., et al. (2024). Crop Recommender: ML-Based Computational Method. Auerbach.

4. Veerasamy, K., & Fredrik, E. T. (2023). Expert System with Butterfly Optimization. Infinite Study.
5. Kaggle Dataset: Crop Recommendation based on soil and weather.
6. Motwani, A., et al. (2022). Soil Analysis and ML Crop Recommender. IEEE.
7. Choudhary, M., et al. (2022). Plant Disease Classification with ML. Springer.