**CROP RECOMMENDATION SYSTEM**  
  
 Authors

1. Nagendra
2. Pappala Archana
3. Penuguluru Bhanodhar
4. Pentakota Yashoda

Authors Affiliations

*Department of Information Technology,*

*Jawaharlal Nehru Technological University - Gurajada, Vizianagaram, Andhra Pradesh, India***Abstract** The complete Crop Recommendation System presented in this study uses machine learning algorithms to give farmers data-driven crop recommendations. By evaluating a range of soil and environmental factors, this methodology is intended to support sustainable farming methods by determining which crops are most suited for a particular area. This Crop Recommendation System demonstrates the transformative potential of artificial intelligence in agriculture, helping farmers to adapt to changing environmental conditions, optimize land use.  
 Future development plans aim to incorporate real-time data from IoT sensors and weather stations to further enhance predictive accuracy and responsiveness, aligning with the goals of precision agriculture. To guarantee high relevance and accuracy, the predictive model incorporates important elements including temperature, humidity, pH, rainfall, and soil composition (including amounts of nitrogen, phosphorus, and potassium).   
 This algorithm finds patterns in the data and forecasts the best crops that can flourish in particular environmental circumstances by combining a number of machine learning approaches. To increase predicted accuracy, models like Random Forest, Decision Trees are enhanced.

**I.Introduction** Precision agriculture relies heavily on crop recommendation, which helps farmers choose the best crops based on a range of soil and environmental factors. Resource efficiency, environmental sustainability, and agricultural output can all be greatly impacted by choosing the right crops. Conventional crop selection techniques frequently rely significantly on farmers' past practices, experience, and intuition, all of which can be arbitrary and occasionally ineffectual, particularly as soil conditions and climatic patterns evolve. The analytical depth and data-driven insights required to adjust to the changing agricultural concerns of today, such as soil degradation, water scarcity, and unpredictable climatic variations, are frequently absent.   
 It can be advanced with the use of the crop suggestion model and the Streamlit interface. This tool not only encourages better agricultural yields and resource management but also allows farmers to make data-informed choices, potentially decreasing hazards associated with poor crop selection. By boosting agricultural productivity through scientific insights and increasing agricultural efficiency and flexibility, this approach supports farmers in addressing environmental concerns and advancing sustainable agriculture's larger objectives.  
  
**II. Objectives**

The main objectives of this study are:

* To develop a machine learning-based system that recommends suitable crops based on environmental data.
* To preprocess the dataset to ensure data quality by removing duplicates and outliers.
* To compare different machine learning models for accuracy in crop prediction.
* To implement a user-friendly Streamlit interface for interaction with the recommendation

system.

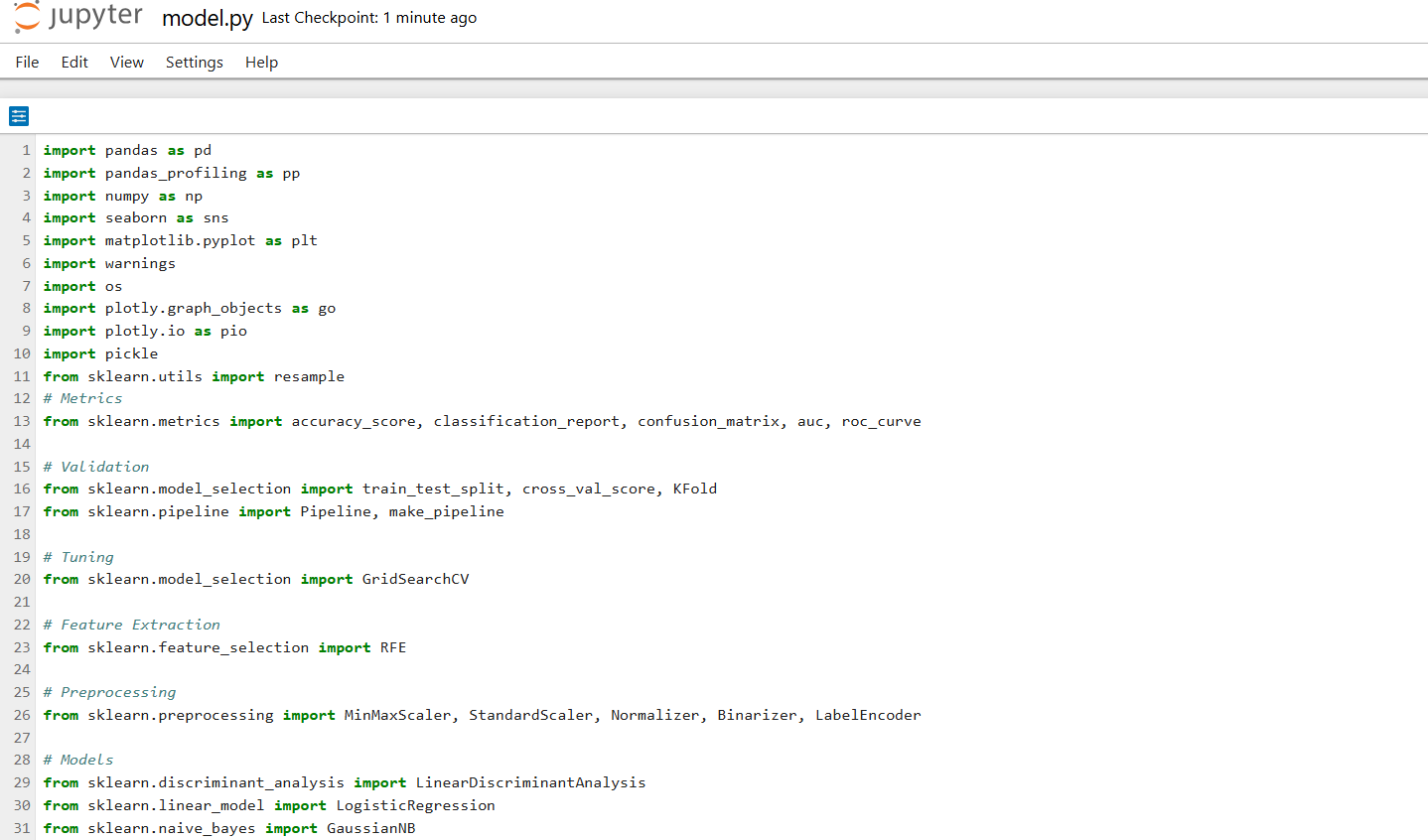
# **III.LiteratureReview** More and more advanced machine learning algorithms, including Decision Trees, Random Forest, Support Vector Machines (SVM), and ensemble approaches, are being used in crop recommendation systems in recent years. These algorithms examine the intricate connections between crop adaptability and environmental conditions to provide incredibly precise, data-driven predictions. Traditional approaches have been revolutionized by the use of machine learning to agriculture, allowing for predictions that are more accurate, faster, and more suited to changing soil and also based on the various kinds of the different types of the climate conditions. Research in this area emphasizes how crucial model selection, evaluation, and data pretreatment are. Preprocessing techniques like outlier detection, normalization, and feature scaling are essential for crop recommendation because they guarantee that the data entered into the algorithms is reliable and pertinent. Furthermore, model selection is important since the algorithm used affects the system's accuracy and efficiency. Finding the optimal model to address the unique difficulties of crop recommendation, such as managing different data sizes, seasonal variations, and geographical variations in soil composition, requires comparing algorithms and adjusting hyperparameters.

# **IV. Data Description And Preprocessing** The Crop Recommendation dataset, which includes 2201 rows and 6 columns and offers crucial environmental characteristics that affect crop development and selection, was used for this investigation. The following characteristics make up the dataset: • Nitrogen (N): The amount of nitrogen in the soil that is necessary for plant growth. • Phosphorus (P): The amount of phosphorus in the soil that affects root growth. • Potassium (K): The concentration of potassium in the soil helps prevent disease. • Temperature: The local ambient temperature (°C). • Humidity: The percentage of humidity that influences transpiration rates. • pH: The pH of the soil affects the availability of nutrients. • Rainfall: One of the main determinants of water availability is annual rainfall (mm). To guarantee data quality, the data pretreatment technique included deleting any duplicate items and applying the Interquartile Range (IQR) approach to remove outliers.

# **V. Methodology**

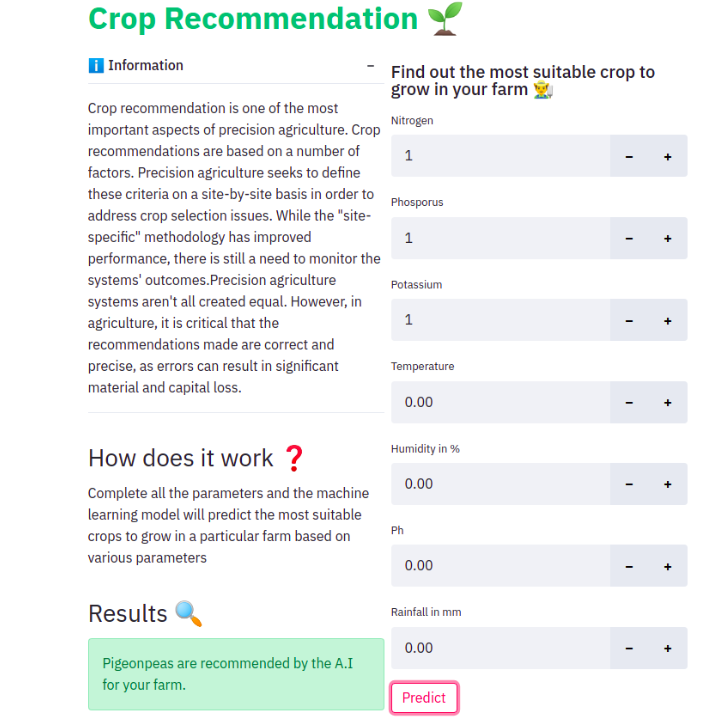
# **A**. **Model Selection**

When characteristics and the target variable have a roughly linear relationship, logistic regression—a linear model that calculates the likelihood of a binary outcome—is the best option. Although it is straightforward to interpret and computationally efficient, it might not perform as well on more complicated, non-linear datasets. Decision Trees are A non-linear model that creates a structure resembling a tree by dividing data according.  
 Although decision trees are interpretable and capable of handling both numerical and categorical data, they have a tendency to overfit when improperly adjusted. When characteristics and the target variable have a roughly linear relationship, logistic regression—a linear model that calculates the likelihood of a binary outcome—is the best option. Although it is straightforward to interpret and computationally efficient.   
 Decision Trees are A non-linear model that creates a structure resembling a tree by dividing data according to feature values. Each model and method was evaluated on key metrics, including accuracy, precision, recall, and F1-score, and compared to assess which performed best under different conditions. Ultimately, the choice of model depends on the specific requirements of the task, such as the trade-off between accuracy and interpretability, and the available computational resources.



**VI. Streamlit Application**

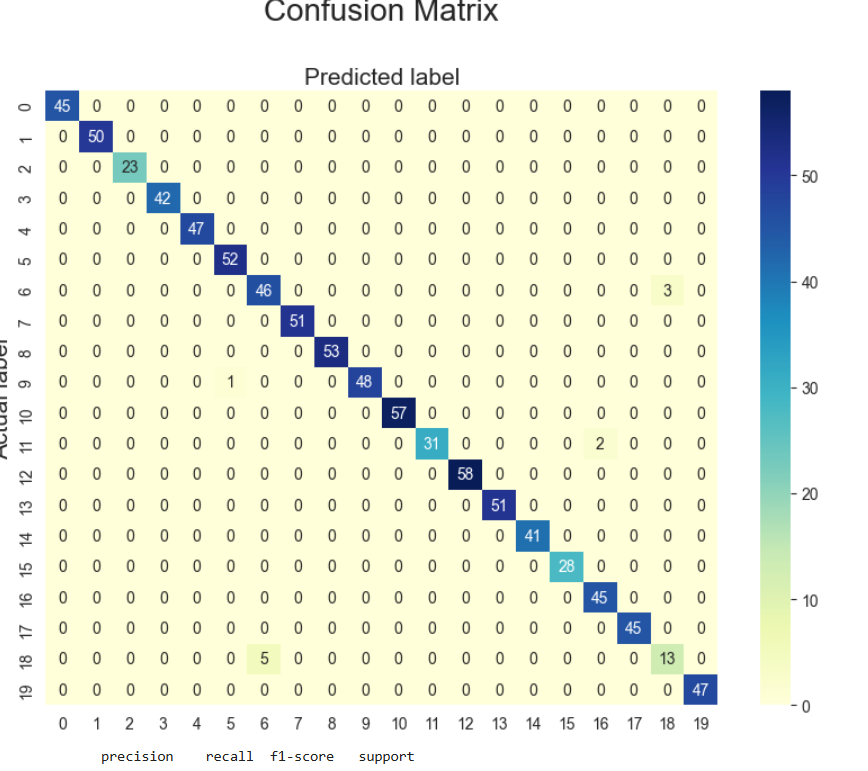
A user-friendly Streamlit application was developed to allow users to input environmental parameters and receive crop recommendations.



**VII.ModelEvaluation** A variety of measures were used to evaluate the machine learning models' performance in order to give a thorough assessment of each model's efficacy in various areas. These parameters, which collectively provide information about the model's overall performance, its capacity to accurately identify positive cases, and its ability to balance erroneous positives and false negatives, which are generally included accuracy, precision, recall, and F1-score.  
 The performance outcomes highlight how effective ensemble approaches—Random Forest in particular—are at producing solid and dependable forecasts. Because of this, it is a great option for practical uses like agriculture, where precise forecasts are crucial for making decisions. Random Forest could be used, for example, in agricultural contexts to evaluate soil conditions, forecast crop illnesses, or optimize irrigation schedules.  
 The Random Forest model stood out with the highest accuracy at 89%, indicating its robust performance. Its strong performance can be attributed to its ability to reduce overfitting through averaging multiple decision trees, thus providing a more generalized model that performs well on unseen data. The ensemble nature of Random Forest allows it to effectively capture complex relationships and patterns within the data, making it less sensitive to individual noise or outliers compared to single models like Decision Trees.

# **VIII.DiscussionAndFutureWork** The crop recommendation system developed in this project demonstrates the growing potential of machine learning in revolutionizing the agricultural industry. By leveraging data-driven insights, the system can assist farmers and agricultural professionals in making more informed decisions about which crops to grow under specific environmental conditions. The results obtained from the Random Forest model highlight its robustness and suitability for agricultural applications, with an accuracy of 89% indicating that it can deliver reliable recommendations. The interactive Streamlit app enhances the accessibility and usability of the system, offering users the ability to input real-time data and receive immediate crop recommendations. This system is a significant step toward optimizing agricultural productivity, especially in regions with variable environmental factors. It showcases how modern technologies like machine learning, data analytics, and cloud-based applications can contribute to more sustainable farming practices, improved resource management, and increased food security. However, while the current implementation is promising, there are several areas for improvement and future development that can enhance its applicability and accuracy.

# **IX.ResultsAndAnalysis** The Streamlit app provides an intuitive, user-friendly interface that enables real-time interaction with the machine learning model, allowing users to input data dynamically and receive immediate feedback on crop recommendations. The interactive nature of the application makes it ideal for scenarios where users need to adjust environmental parameters and view updated predictions quickly, ensuring that the decision-making process is efficient and based on the latest data. y enabling users to easily modify environmental conditions and receive instant, data-driven crop recommendations, the Streamlit app provides a powerful tool for agricultural decision-making. Whether for farmers looking to optimize their crop yields or researchers studying the impact of climate on agriculture, this app delivers actionable insights in a seamless, interactive manner.



# **X. Conclusion**

In conclusion, this project represents a significant step toward utilizing machine learning to improve agricultural practices by providing accurate, data-driven crop recommendations. The Streamlit app offers a practical, accessible platform for farmers to interact with the model and make informed decisions based on environmental conditions. With future enhancements such as real-time weather integration, regional optimization, and broader scaling, this system has the potential to become a vital tool in precision agriculture, helping farmers around the world to increase yields, reduce resource waste, and adapt to changing climatic conditions. The project demonstrates that machine learning has the capability to drive sustainable agriculture, contributing to the global effort to meet the food production needs of a growing population while protecting the environment.  
 This project illustrates the transformative potential of machine learning in the agricultural sector, offering a practical application that can significantly enhance farming practices. The crop recommendation system developed here leverages environmental data to provide data-driven insights, guiding farmers in selecting the most suitable crops based on their specific environmental conditions.

**REFERENCES**[1] [Sharma, P., et al. AI-Enabled Crop Recommendation System Based on Soil and Weather Patterns.In Artifi](https://doi.org/10.1007/springer1)

[2] [Jha, G. K., et al. A machine learning approach to recommend suitable crops and fertilizers for agriculture](https://doi.org/10.1016/2fsj.srl.2020.110.22) [(](https://doi.org/10.1016/2fsj.srl.2020.110.22)[Bukhari, S. N. H., et al. Crop Recommender: Machine Learning-Based Method for Best Crop Selection U](https://doi.org/10.1016/2fsj.srl.2024.210.21)

[3] Gosai, D., Raval, C., Nayak, R., Jayswal, H., & Patel, A. (2021). Crop recommendation system using machine learning. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 7(3), 558-569.

[4] Attri, I., Awasthi, L. K., Sharma, T. P., & Rathee, P. (2023). A review of deep learning techniques used in agriculture. Ecological Informatics, 102217.

[5] Dutta, M., Gupta, D., Juneja, S., Shah, A., Shaikh, A., Shukla, V., & Kumar, M. (2023). Boosting of fruit choices using machine learning-based pomological recommendation system. SN Applied Sciences, 5(9), 241.