



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 5, May 2025

Data Driven Recommendation for Tailored Crop Selection

R.Ponmayil, Tamilmozhi S

Assistant Professor , Department of Computer Application, Vivekanandha Institute of Information and Management
Studies, Tiruchengode, Tamil Nadu, India

Master of Computer Applications, Vivekanandha Institute of Information and Management Studies, Tiruchengode,
Tamil Nadu, India

ABSTRACT: A data-driven crop recommendation system using machine learning to help farmers select optimal crops based on soil and environmental conditions. The process begins with exploratory data analysis, including data cleaning, descriptive statistics, and correlation analysis. Several machine learning models are applied—Random Forest (as the primary model), Decision Tree, SVM, and Logistic Regression—to predict suitable crops. These models are evaluated using metrics like accuracy, precision, recall, and F1 score. Hyper parameter tuning through Grid Search is used to optimize performance. The system ultimately identifies the most effective model to ensure accurate and reliable crop recommendations.

KEYWORDS: Data-driven crop recommendation systems leverage machine learning, predictive modeling, and data analytics to provide tailored crop selection based on soil properties, climate data, and agro-environmental factors. By integrating remote sensing, GIS mapping, and sensor data, these systems offer site-specific recommendations that optimize crop yield, enhance resource management, and support sustainable agriculture.

I. INTRODUCTION

Data driven Recommendation for tailored for crop selection In modern agriculture selecting the right crop for specific environmental conditions is crucial for maximizing yield, ensuring sustainability, and improving profitability. Traditional crop selection methods often rely on expert knowledge and general regional data, which may not capture the unique characteristics of individual farms. Data-driven recommendation systems leverage advances in machine learning, big data, and remote sensing to analyze diverse factors such as soil quality, climate variables, and market trends. By integrating these datasets, these systems provide tailored, site-specific crop suggestions that help farmers optimize resource use, adapt to changing environmental conditions, and make informed decisions. This approach represents a significant step toward precision agriculture, enabling more efficient, sustainable, and profitable farming practice

II. OBJECTIVES

The objective of a data-driven crop recommendation system is to assist farmers in selecting the most suitable crops based on specific environmental and soil conditions. It aims to enhance agricultural productivity by identifying crops that are best adapted to local climate, temperature, humidity, and soil nutrient levels. By leveraging machine learning algorithms, the system can analyze large datasets to uncover patterns that inform smarter crop choices. One of the key goals is to maximize yield while minimizing input costs, helping farmers achieve higher profitability. The system also promotes efficient use of resources such as water, fertilizers, and pesticides, contributing to environmentally sustainable farming. It provides location-specific and real-time recommendations, which improve decision-making under variable conditions. Market trends are also considered, ensuring that selected crops are both agronomically and economically viable. Reducing crop failure risks due to poor planning or unsuitable conditions is another important goal. The system supports climate-smart agriculture by suggesting resilient crops that can withstand weather extremes. It fosters precision agriculture by customizing recommendations at the farm level. Hyperparameter tuning and model evaluation are used to enhance the accuracy of the predictions. The system encourages data-driven decision-making, replacing guesswork with evidence-based strategies. It empowers small and large-scale farmers alike to make informed, timely

decisions. By continuously learning from new data, the system remains adaptive and relevant. Overall, the aim is to build a reliable, smart solution that supports sustainable and productive farming practices.

III. LITERATURE SURVEY

Recent studies have shown significant progress in the development of data-driven crop recommendation systems using machine learning and AI. Patil and Biradar (2020) developed a crop prediction model using soil nutrients, pH, and climate data, applying classification algorithms like Random Forest for accurate results. Kagalkar et al. (2021) combined IoT sensor data and Support Vector Machines to provide real-time, site-specific crop suggestions. Dhanabal and Chandramathi (2017) utilized data mining techniques such as Naïve Bayes and clustering to analyze historical yield and environmental data. Sujithra and Karthikeyan (2022) introduced deep learning models with LSTM networks to offer dynamic recommendations based on weather forecasts. Ensemble approaches, such as the one proposed by Kishore et al. (2019), combined Random Forest and Gradient Boosting to consider both agronomic and market factors. FAO and CGIAR reports emphasize integrating satellite imagery and GIS for scalable advisory systems. Shrivastava et al. (2021) used remote sensing data to generate crop suitability maps, helping farmers adapt to climate variability. Mishra et al. (2018) focused on a user-friendly web-based recommendation tool for small-scale farmers. Pal et al. (2020) employed multi-objective optimization to balance yield, water use, and economic return. Zhang et al. (2019) leveraged deep neural networks and high-resolution remote sensing data for large-scale crop performance prediction. Collectively, these works highlight the growing role of intelligent systems in promoting sustainable, profitable, and localized crop selection strategies.

IV. EXISTING SYSTEM

Several existing systems have been developed to support data-driven crop recommendation using machine learning and big data technologies. The Mobile and web-based tools like **Krishi Network** and **RML AgTech** in India provide localized crop suggestions and agricultural advice using regional data. Research-based systems such as those developed by **ICAR** and **IARI** use soil health cards, weather data, and historical yields to suggest optimal crops to farmers. Platforms like **CropIn** and **AgNext** combine satellite imagery, IoT sensors, and AI to offer precision agriculture solutions, including crop recommendations. Additionally, academic prototypes such as those built using **Random Forest**, **Decision Tree**, and **SVM** models have shown success in predicting suitable crops based on datasets containing soil nutrients, pH, temperature, rainfall, and humidity.

DISADVANTAGES

- **Data Dependency** - Requires high-quality, accurate, and up-to-date data.
- **High Setup Cost** - Implementing sensors, data collection tools.
- **Maintenance Challenges** - Regular updates, calibration, and technical support are needed.
- **User Interface Issues** - not user-friendly for farmers with limited digital literacy.
- **Technical Complexity** - it can be difficult for non-experts to understand or models

V. PROPOSED SYSTEM

The proposed system aims to provide tailored crop recommendations using machine learning based on specific soil and environmental conditions. It collects input data such as soil nutrients (N, P, K), pH, temperature, humidity, and rainfall. The data undergoes preprocessing and exploratory analysis to ensure accuracy and consistency. Machine learning models like Random Forest, Decision Tree, and SVM are trained to predict the most suitable crops. Model performance is evaluated using accuracy, precision, recall, and F1-score. The final system delivers user-friendly, location-specific crop suggestions to support smarter and more sustainable farming decisions.

VI. METHODOLOGIES

This methodology outlines a comprehensive approach for developing a data-driven crop recommendation system that leverages machine learning techniques to assist farmers in making informed decisions. By integrating data collection, preprocessing, model selection, evaluation, and deployment, this system aims to optimize crop selection processes and

promote sustainable agricultural practices by leveraging historical agricultural data and advanced machine learning algorithms, the project aims to optimize crop selection, thereby enhancing yield, resource efficiency, and sustainability for farmers. The system seeks to analyze various factors such as soil properties, weather patterns, and regional agricultural practices to provide tailored recommendations that can help farmers make informed decisions, reduce risks, and maximize profitability. Ultimately, the goal is to create a tool that supports sustainable agriculture by promoting the best crop choices for given conditions, leading to increased food security and economic resilience in farming communities.

MODULES

- Data Collection
- Data Preprocessing
- Exploratory Data Analysis (EDA)
- Model Building
- Model Evaluation & Optimization
- System Deployment

MODULE DESCRIPTION

Data collection

Data collection is the foundational step in developing a crop recommendation system. It involves gathering information from multiple sources such as soil health reports, weather stations, agricultural research databases, and satellite imagery. Key parameters collected include soil nutrients (N, P, K), pH level, temperature, humidity, rainfall, and past crop performance.

Data preprocessing

Data preprocessing is a crucial step that prepares raw data for effective analysis and modeling. It involves cleaning the dataset by handling missing values, removing duplicates, and correcting inconsistencies. Numerical features such as soil nutrients and temperature are normalized to ensure uniformity, while categorical variables are encoded for compatibility with machine learning algorithms.

Exploratory data Analysis (EDA)

Exploratory Data Analysis (EDA) helps understand the structure and patterns within the dataset. It involves using statistical summaries and visualizations to explore relationships between features like soil nutrients, temperature, and crop types.

Model building

Model building involves training machine learning algorithms to predict suitable crops based on processed data. Common models used include Random Forest, Decision Tree, Support Vector Machine (SVM), and Logistic Regression. The models learn from input features like soil nutrients, pH, temperature, and rainfall to classify the best crop.

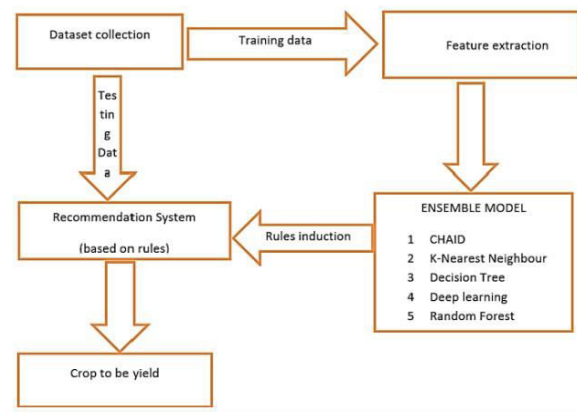
Model Evaluation & Optimization

Model evaluation and optimization involve assessing the performance of trained models using metrics like accuracy, precision, recall, and F1-score. This step helps identify the most effective algorithm for crop prediction. Optimization techniques such as hyperparameter tuning (e.g., Grid Search) are applied to improve model accuracy and reliability..

System deployment

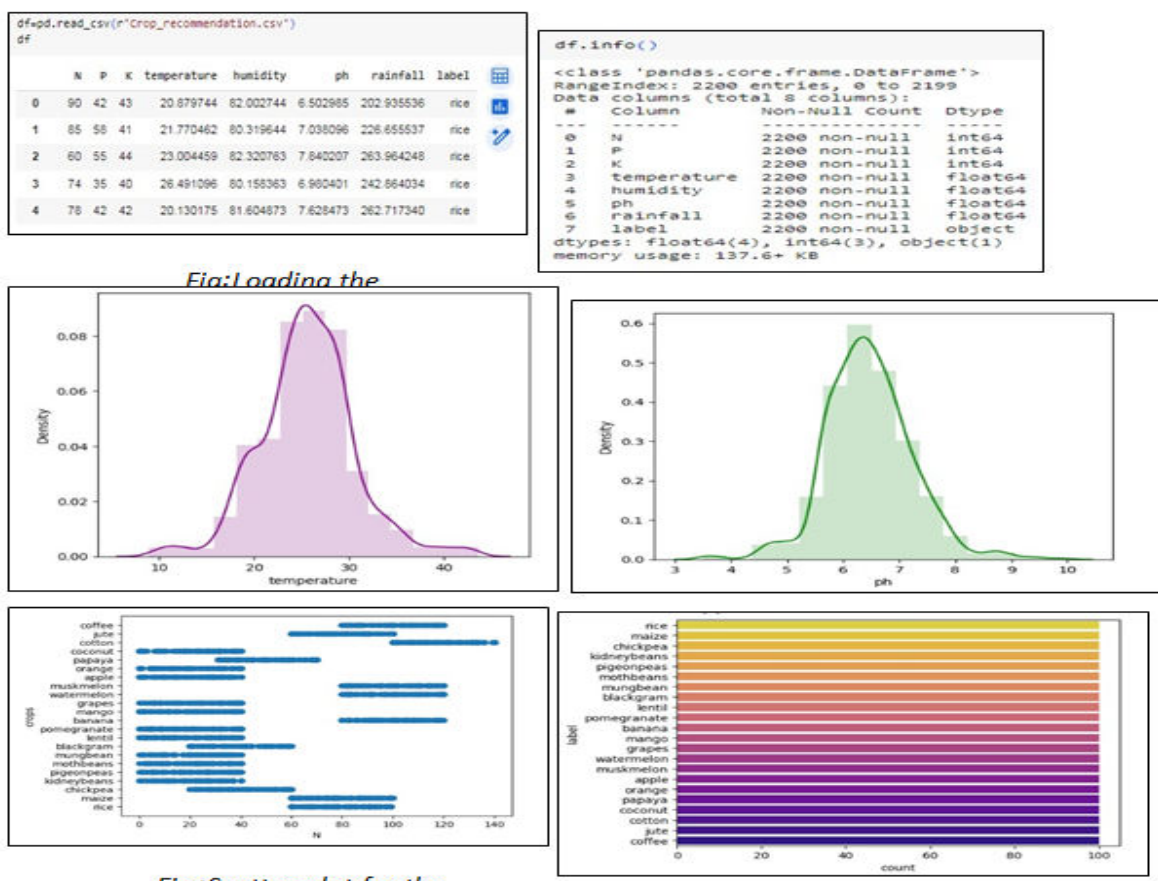
System development involves designing and building the complete crop recommendation platform. It includes integrating machine learning models with a user interface that is accessible and easy to use. Backend systems manage data input, processing, and prediction delivery efficiently. The goal is to create a reliable, scalable, and farmer-friendly application for real-world use.

SYSTEM ARCHITECTURE



VII. EXPERIMENTAL RESULT

Figure shows the result (a)loading the data (b) Checking the information of the dataset(c) Checking the kernel density estimator for ph and density(d) Checking the kernel density estimator for temperature and density(e) Scatter plot for the distribution(f) Count plot for the distribution of the dataset.



VIII. CONCLUSION

The crop recommendation project using machine learning offers a significant advancement in agricultural decision-making by leveraging data-driven insights to guide crop selection. By analyzing a wide range of factors such as soil properties, climate conditions, and market trends, the system provides farmers with tailored recommendations that can enhance productivity, sustainability, and profitability. The successful implementation of machine learning models, like the Random Forest Classifier, demonstrates the potential of AI in transforming traditional farming practices.

IX. FUTURE ENHANCEMENT

In the crop recommendation project using machine learning, there are several avenues for future enhancement to improve the system's accuracy, usability, and scalability:

Integration of Real-Time Data: Incorporating real-time weather forecasts and climate change projections can make recommendations more dynamic and responsive to changing conditions. Using IoT devices in fields to collect real-time data on soil moisture, pH, and temperature can provide more accurate inputs for the recommendation system.

Advanced Algorithms: Exploring deep learning techniques such as Convolutional Neural Networks (CNNs) for image-based soil analysis or Recurrent Neural Networks (RNNs) for time-series climate data could enhance the model's predictive power. Combining multiple machine learning algorithms (e.g., Random Forest, Gradient Boosting, and SVM) could improve the robustness of the recommendations by leveraging the strengths of different models.

REFERENCES

1. Amudha, T. (2021). Artificial intelligence: A complete insight. In Artificial intelligence: Theory, models, and applications (pp. 1–24). Auerbach Publications.
2. Bryant, B. P., Kelsey, T. R., Vogl, A. L., Wolny, S. A., MacEwan, D., Selmants, P. C., ... Butterfield, H. S. (2020). Shaping land use change and ecosystem restoration in a water-stressed agricultural landscape to achieve multiple benefits. *Frontiers in Sustainable Food Systems*, 4, 138. <https://doi.org/10.3389/fsufs.2020.00138>
3. Falkeis, A., Shesterikova, A., James, B., & Tingen, M. (2022). Nonlinear Urbanism: Towards multiple urban futures. Birkhäuser.
4. Gardezi, M., Abuayyash, H., Adler, P. R., Alvez, J. P., Anjum, R., Badireddy, A. R., ... Dadkhah, A. (2024). The role of living labs in cultivating inclusive and responsible innovation in precision agriculture. *Agricultural Systems*, 216, 103908. <https://doi.org/10.1016/j.agsy.2022.103908>
5. Jayapandian, N. (2023). Machine learning-based spam e-mail detection using logistic regression algorithm. In *Proceedings of the 2023 IEEE International Conference on ICT in Business Industry & Government (ICTBIG)*.
6. Jhariya, M. K., Meena, R. S., & Banerjee, A. (2021). Ecological intensification of natural resources towards sustainable productive system. In *Ecological intensification of natural resources for sustainable agriculture* (pp. 1–28).

WEBSITE REFERENCES

1. <https://www.mdpi.com/2077-0472/13/11/2141>
2. <https://arxiv.org/abs/2412.16196>
3. <https://arxiv.org/abs/2502.04054>
4. <https://www.nature.com/articles/s41598-025-93417-3>
5. <https://medium.com/@1032211306/developing-an-intelligent-crop-classification-and-recommendation-system-with-machine-learning-a-cdce3ddbba09>
6. <https://github.com/611noorsaeed/Crop-Recommendation-System-Using-Machine-Learning>
7. <https://github.com/KRUTHIKTR/Crop-Recommendation-System-Using-Machine-Learning>
8. <https://www.kaggle.com/code/nirmalgaud/crop-recommendation-system-using-machine-learning>
9. <https://www.geeksforgeeks.org/crop-recommendation-system-using-tensorflow>
10. https://www.youtube.com/watch?v=T7_GR8I-N5A
11. <https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset>

12. https://sist.sathyabama.ac.in/sist_naac/documents/1.3.4/1822-b.e-cse-batchno-336.pdf
13. https://en.wikipedia.org/wiki/Precision_agriculture
14. https://en.wikipedia.org/wiki/Digital_agriculture
15. https://en.wikipedia.org/wiki/Data_mining_in_agriculture



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details