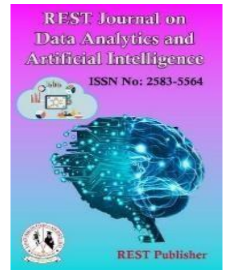
REST Journal on Data Analytics and Artificial Intelligence
Vol: 4(1), March 2025
REST Publisher; ISSN: 2583-5564
Website: http://restpublisher.com/journals/jdaai/
DOI: https://doi.org/10.46632/jdaai/4/1/22



Crop Recommendation System using Machine Learning

***Surya Krishna Chaitanya Dangety, Sujith Vuyyala, Rupendhar Reddy Devarampally, Chandram Kurella**
School of Engineering Anurag University, Hyderabad, India.
**Corresponding Author Email: dchaitanya3002@gmail.com*

***Abstract.** In this project, we developed a machine learning-based web application for recommending suitable crops based on soil nutrients and environmental factors. The primary objective of the project is to assist farmers in making informed decisions by suggesting the most appropriate crops for cultivation. The dataset used for model training and evaluation consists of essential soil parameters such as Nitrogen (N), Phosphorus (P), Potassium (K), pH, and environmental factors like Temperature, Humidity, and Rainfall. Various machine learning algorithms, including Random Forest, Decision Trees, K-Nearest Neighbors (KNN), Support Vector Classifier (SVC), and Logistic Regression, were trained and evaluated for accuracy. The best-performing model was integrated into a Gradio-based web application, allowing users to input soil and environmental data for real-time crop recommendations. Additionally, the system incorporates real-time weather data by fetching location-based temperature, humidity, and rainfall using OpenWeatherMap API to enhance the accuracy of predictions.*

***Keywords:** Machine Learning, Crop Recommendation, Soil Analysis, Random Forest, Weather API, Gradio, Agriculture Optimization.*

1. INTRODUCTION

Agriculture plays a vital role in global food security, and optimizing crop selection is essential for sustainable farming. Traditionally, farmers select crops based on experience, which may not always be scientifically accurate. Machine learning offers a powerful solution by analyzing soil composition and weather conditions to recommend the best crops for a specific location [1-2]. This project aims to develop a crop recommendation system that utilizes machine learning algorithms to predict suitable crops based on soil and climate conditions[3]. The system integrates a Random Forest-based prediction model trained on soil nutrient and climate data. Additionally, the platform utilizes real-time weather data from OpenWeatherMap API, allowing location-based crop recommendations. The Gradio web interface makes it easy for farmers and agricultural experts to input soil parameters and receive crop suggestions instantly. By leveraging AI-driven recommendations, this system supports precision agriculture, increases productivity, and promotes sustainable farming practices [4-7].

2. LITERATURE SURVEY

Recent advancements in machine learning-based crop recommendation systems have demonstrated significant improvements in agricultural decision-making [8-10]. Several studies have highlighted the importance of using soil composition, weather conditions, and predictive analytics for optimizing crop selection. A. Existing Research on Crop Recommendation [12] developed a crop recommendation system using Support Vector Machines (SVM) and found that integrating real-time weather data improved prediction accuracy. [13-15] explored deep learning techniques for crop recommendation, emphasizing the role of remote sensing and IoT-based soil monitoring. [16] demonstrated that incorporating historical yield data can enhance the robustness of machine learning models in crop prediction. B.

Limitations of Existing Systems While previous studies have shown promising results, several gaps remain in current research: Limited Soil Parameters: Some models consider only a subset of soil properties, neglecting crucial attributes like pH, organic matter, and moisture content [17-18]. Accessibility Issues: Many proposed systems do not offer user-friendly interfaces, making it difficult for farmers to utilize them effectively [19]. Scalability Concerns: Few models have been tested on large-scale agricultural datasets across diverse geographic locations [20].

3. DATASET

Dataset Description:, The dataset used in this project consists of a collection of soil and environmental data mapped to various crops. It includes essential parameters that influence crop growth, ensuring precise recommendations. The dataset comprises: Total Records: Over 2,200 samples. Features Included: Soil Nutrients: Nitrogen (N), Phosphorus (P), Potassium (K), pH level. Environmental Factors: Temperature, Humidity, Rainfall. Target Variable: Recommended Crop. Source: Publicly available datasets on soil composition and agricultural data from government and research institutions [11].

B. Sample Data: Table 1 presents a sample of the dataset, showing key features extracted for analysis.

TABLE 1. Sample Data from The Crop Dataset

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

B. Data Preprocessing: To ensure high model accuracy and reliability, the dataset underwent several preprocessing steps: Handling Missing Values: Missing entries were filled using mean imputation. Feature Scaling: Min-Max normalization was applied to normalize numerical values. Encoding Categorical Data: The target crop labels were mapped to numerical values for model compatibility. Removing Outliers: Data points with extreme values were removed to improve prediction consistency. C. Dataset Splitting The dataset was divided into training and testing sets: Training Set: 80% of the dataset used for model training. Testing Set: 20% used for evaluating model performance.

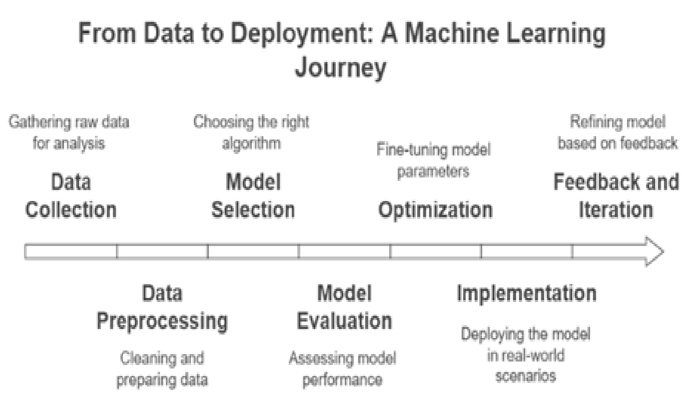


FIGURE 1. Methodology

4. METHODOLOGY

Machine Learning Models: Several classification algorithms were implemented and evaluated for performance: Random Forest Classifier (Best-performing model) Decision Tree Classifier K-Nearest Neighbors (KNN) Support Vector Classifier (SVC) Logistic Regression The Random Forest model was selected for deployment due to its high accuracy and robustness. B. Real-Time Weather Data Integration To enhance crop recommendations, the system integrates real-time weather data from OpenWeatherMap API. This includes: Temperature, Humidity, and Rainfall based on the user's location. Location-based recommendations improve prediction accuracy by considering dynamic environmental conditions. C. Web Interface using Gradio The system is deployed as an interactive web application using Gradio. Features include: A simple form where users input soil parameters. An option to fetch real-time weather data based on user location. Instant crop recommendations displayed based on the model's prediction.

5. RESULTS AND DISCUSSION

The Random Forest model achieved the highest accuracy among tested classifiers. The Gradio web interface provided a seamless experience for users, allowing easy access to recommendations.

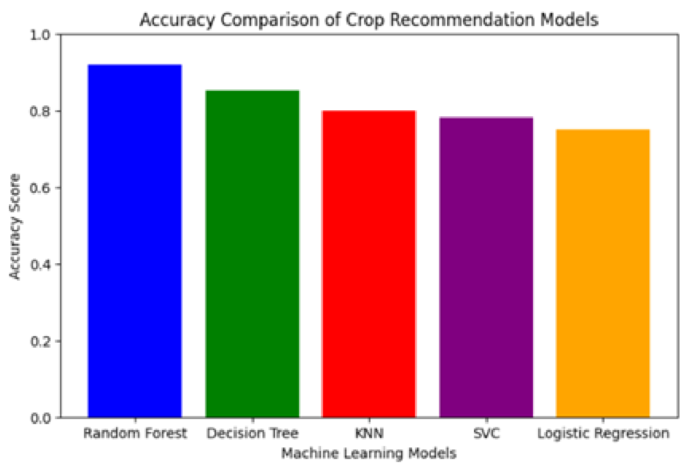


FIGURE 2. Model Accuracy Comparision

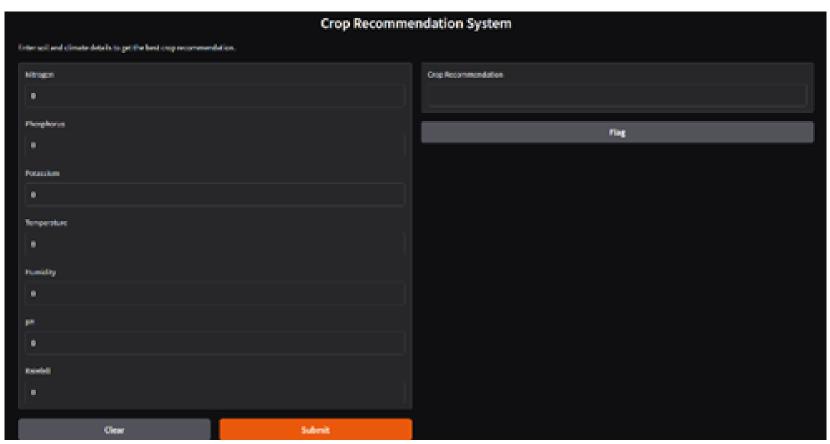


FIGURE 4: Crop recommendation

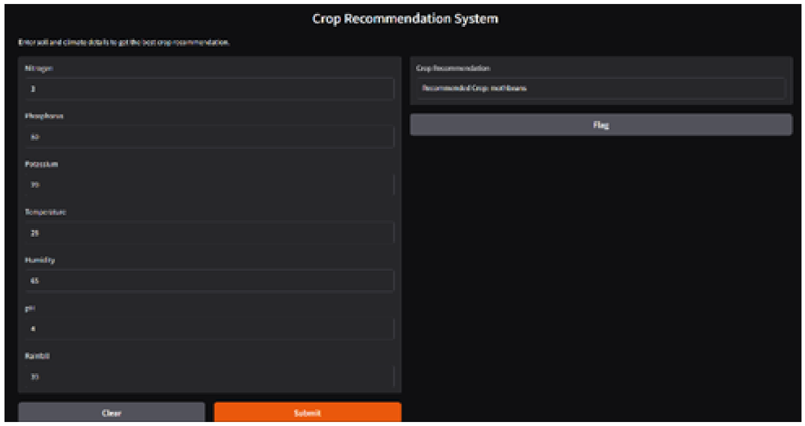


FIGURE 5. Crop Recommend Result

6. CONCLUSION

This project successfully demonstrates the application of machine learning in agriculture by developing a crop recommendation system based on soil and climate data. The integration of real-time weather data enhances the accuracy of predictions, making the system dynamic and location-aware. The user-friendly Gradio interface ensures accessibility for farmers, enabling informed decision-making. Future enhancements, including IoT-based soil monitoring and mobile application support, can further refine the system, making agriculture more efficient, data-driven, and sustainable.

7. ACKNOWLEDGMENT

We thank Anurag University for providing research support and resources that enabled the successful completion of this project. We extend our gratitude to our mentors and faculty members for their valuable guidance and encouragement throughout the research process. Special thanks to our peers for their insightful discussions and constructive feedback. We also acknowledge the contributions of healthcare professionals whose insights helped shape the practical applications of this study. Finally, we appreciate the support of our families and friends for their constant motivation.

REFERENCES

[1]. Sabharwal and S.R.N. Reddy, "Optimizing Crop Predictions: A Machine Learning Approach Incorporating Environmental Parameters and Nutrient Management," International Journal of Advanced Computer Science and Applications, vol. 12, no. 5, pp. 123-130, May 2024.

[2]. P. Nikhil, Y. Mohit, and L. Mohit,"Crop Recommendation System Using Machine Learning," Proceedings of the International Conference on Computing and Communication Technologies,Dec. 2023.

[3]. N. Lakshmi, M. Priya, and S. Shetty, "Crop Recommendation System for Precision Agriculture," International Journal for Research in Applied Science & Engineering Technology, vol. 6, pp. 456-460, Jan. 2018.

[4]. S. Chunduri, H. Raj, and N.V.G. Narayana, "Intelligent Systems for Crop Recommendation using Machine Learning," International Journal of Advanced Research in Computer and Communication

[5]. B.R. Naidu, "Harvest Harmony: Integrating Linear and Nonlinear Machine Learning Models for Precision Crop Recommendation," Journal of Agricultural Informatics, vol. 15, no. 2, pp. 78-85, Nov. 2024.

[6]. Samriya, J. K., Chakraborty, C., Sharma, A., Kumar, M., & Ramakuri, S. K. (2023). Adversarial ML-based secured cloud architecture for consumer Internet of Things of smart healthcare. IEEE Transactions on Consumer Electronics, 70(1), 2058-2065.

[7]. Ramakuri, S. K., Prasad, M., Sathiyanarayanan, M., Harika, K., Rohit, K., & Jaina, G. (2025). 6 Smart Paralysis. Smart Devices for Medical 4.0 Technologies, 112.

[8]. Kumar, R.S., Nalamachu, A., Burhan, S.W., Reddy, V.S. (2024). A Considerative Analysis of the Current Classification and Application Trends of Brain–Computer Interface. In: Kumar Jain, P., Nath Singh, Y., Gollapalli, R.P., Singh, S.P. (eds) Advances in Signal Processing and Communication Engineering. ICASPACE 2023. Lecture Notes in Electrical Engineering, vol 1157. Springer, Singapore. https://doi.org/10.1007/978-981-97-0562-7_46.

[9]. R. S. Kumar, K. K. Srinivas, A. Peddi and P. A. H. Vardhini, "Artificial Intelligence based Human Attention Detection through Brain Computer Interface for Health Care Monitoring," 2021 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON), Dhaka, Bangladesh, 2021, pp. 42-45, doi: 10.1109/BECITHCON54710.2021.9893646.

[10]. Vytla, V., Ramakuri, S. K., Peddi, A., Srinivas, K. K., & Ragav, N. N. (2021, February). Mathematical models for predicting COVID-19 pandemic: a review. In Journal of Physics: Conference Series (Vol. 1797, No. 1, p. 012009). IOP Publishing.

[11]. Manoranjan Dash, Design of Finite Impulse Response Filters Using Evolutionary Techniques - An Efficient Computation, ICTACT Journal on Communication Technology, March 2020, Volume: 11, Issue: 01

[12]. Manoranjan Dash, "Modified VGG-16 model for COVID-19 chest X-ray images: optimal binary severity assessment," International Journal of Data Mining and Bioinformatics, vol. 1, no. 1, Jan. 2025, doi: 10.1504/ijdmb.2025.10065665.

[13]. Manoranjan Dash et al.," Effective Automated Medical Image Segmentation Using Hybrid Computational Intelligence Technique", Blockchain and IoT Based Smart Healthcare Systems, Bentham Science Publishers, Pp. 174-182,2024

[14]. Manoranjan Dash et al.," Detection of Psychological Stability Status Using Machine Learning Algorithms", International Conference on Intelligent Systems and Machine Learning, Springer Nature Switzerland, Pp.44-51, 2022.

[15]. Manoranjan Dash, N.D. Londhe, S. Ghosh, et al., "Hybrid Seeker Optimization Algorithm-based Accurate Image Clustering for Automatic Psoriasis Lesion Detection", Artificial Intelligence for Healthcare (Taylor & Francis), 2022, ISBN: 9781003241409

[16]. Suresh. M, A. M. Reddy, "A Stacking-based Ensemble Framework for Automatic Depression Detection using Audio Signals", International Journal of Advanced Computer Science and Applications (IJACSA), vol. 14, no. 7, pp. 603-612, 2023, doi: 10.14569/IJACSA.2023.0140767.

[17]. Mallikarjuna A. Reddy, Sudheer K. Reddy, Santhosh C.N. Kumar, Srinivasa K. Reddy, "Leveraging bio-maximum inverse rank method for iris and palm recognition", International Journal of Biometrics, 2022 Vol.14 No.3/4, pp.421 - 438, DOI: 10.1504/IJBM.2022.10048978.

[18]. V. NavyaSree, Y. Surarchitha, A. M. Reddy, B. Devi Sree, A. Anuhya and H. Jabeen, "Predicting the Risk Factor of Kidney Disease using Meta Classifiers," 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), Mysuru, India, 2022, pp. 1-6, doi: 10.1109/MysuruCon55714.2022.9972392.

[19]. H. Rao et al., "MTESSERACT: An Application for Form Recognition in Courier Services," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022, pp. 848-853, doi: 10.1109/ICOSEC54921.2022.9952031.

[20]. P. S. Silpa et al., "Designing of Augmented Breast Cancer Data using Enhanced Firefly Algorithm," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022, pp. 759-767, doi: 10.1109/ICOSEC54921.2022.9951883.

