International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 03 | Mar 2024 www.irjet.net p-ISSN: 2395-0072

CROP RECOMMENDATION SYSTEM USING ML ALGORITHMS

Dr John Jaidhan B[1], Edara Bindu Madhavi[2], Nandini Sarkar[3], Maddila Chandra Shekhar[4] and Chaitanya Adduri[5]

[1]*Professor, Dept. of CSE, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India.*
[2345]*Student, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India.*

---***---

Abstract - *Agriculture is the backbone of India's economy, crucial for the well-being of its people. Ensuring the production of high-quality crops is essential for maintaining a healthy lifestyle. Analyzing environmental and soil conditions, including factors such as moisture and pH levels, temperature, and chemical composition, is vital for cultivating superior crops. Predicting crop yields has become increasingly challenging due to unpredictable weather patterns caused by global warming, resulting in crop destruction, food scarcity, and tragic consequences such as farmer suicides. This study aims to develop a website utilizing machine learning models for crop recommendations, taking into account inputs such as pH values, temperature, and soil parameters. Various machine learning algorithms, including SVM, logistic regression, naive bayes, and Random Forest, are utilized, with Random Forest demonstrating superior prediction capabilities. These systems carefully analyse diverse factors, including soil quality, climate data, and past crop performance, to suggest optimal crops tailored to specific locations. Accessible through user-friendly platforms, crop recommendation systems empower farmers to harness the benefits of technology and data, thereby enhancing agricultural productivity, financial gains, and global food security.*

Ultimately, these systems serve as a crucial tool in advancing modern agriculture, leveraging technology and data analysis to assist farmers in making decisions that contribute to food security and agricultural sustainability.

Key Words: Random Forest, machine learning model, moisture and pH level, temperature, and chemical composition, recommendation system, factors.

1.INTRODUCTION

Agriculture serves as the backbone of India's economy, supporting the livelihoods of a significant portion of the population. However, it remains susceptible to the unpredictable nature of weather patterns, climate variations, and ecological factors. These uncertainties pose substantial challenges to cultivating robust crops. With the population steadily increasing and poverty levels on the rise, it becomes imperative to equip ourselves with effective tools to address these critical issues and offer recommendations to mitigate the adverse effects of poverty, climate change, and erratic harvests.

In today's modern era, farmers are leveraging advanced technologies to enhance crop yields. While extensive research has been conducted to comprehend the intricate factors influencing agriculture, the collection of comprehensive data remains a hurdle, often resulting in inconsistent and incomplete information. There is a pressing need to refine methodologies for data availability and accessibility to overcome these obstacles. Enter crop recommendation systems, offering a promising solution to these challenges. These systems bring a myriad of benefits to the agricultural sector. Driven by data-driven algorithms and the principles of precision agriculture, crop recommendation systems empower farmers to make well-informed decisions regarding crop selection, thereby maximizing their yield potential. Embracing sustainable farming practices takes precedence, aiming to minimize environmental impact and promote soil health. These systems are easily accessible through user-friendly platforms, enabling farmers to leverage the advantages of technology and data. Ultimately, this contributes to heightened agricultural productivity, financial profitability, and enhanced global food security. In essence, crop recommendation systems play a pivotal role in the modernization of agriculture, fostering sustainability, and improving the livelihoods of farmers worldwide. They serve as a beacon of hope in an industry crucial to our well-being and the future of our planet

2.LITERATURE SURVEY

Today's technological advancements enable the meticulous analysis of agricultural data [1], granting farmers greater control over their crops and empowering them to make informed decisions. Studies in artificial intelligence and machine learning have revolutionized the automation and enhancement of crop quality [1]. These technologies offer insights into crucial aspects such as irrigation management, climate adaptation, and soil nutrition, which significantly impact crop production [2]. Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN), in combination with other techniques, are instrumental in classifying and predicting crop yields In-depth research has led to the recognition of CNN as a prominent deep learning algorithm in crop yield prediction, alongside LSTM (Long Short Term Memory) and DNN (Deep Neural Networks) [3]. Machine learning assumes a pivotal role in Crop Yield Prediction (CYP), acting as an independent tool [4]. Prof. Dr. Miklós Neményi emphasizes the importance of considering soil, meteorological, environmental, and crop parameters in predicting crop yield [5]. Decision support models aid in extracting significant crop features for accurate prediction]. By utilizing weather parameters, soil characteristics, and historical yield data, predictive systems can accurately forecast crop yields [6]. These systems aim to integrate data from various sources, employing data analytics and forecast analysis to enhance productivity and profitability in agriculture. Machine learning models, including reinforcement learning, are employed to predict crop yields, with a focus on deep reinforcement learning techniques [7]. Deep learning techniques like ANN and CNN play a crucial role in exploring classification and regression methods for crop yield prediction [8]. Researchers have compared the prediction accuracy of various models, including ANN, KNN, Random Forest, MLR, and SVR [9]. A hybrid MLR-ANN model has shown superior performance in terms of accuracy, making it a preferred choice for predictive tasks [9]. Machine learning models like the majority voting system, SVM, Naive Bayes, ANN, and Random Forest are utilized, leveraging soil characteristics obtained from soil testing labs [10]. These characteristics, such as depth, texture, and pH, play a significant role in determining a crop's potential for nutrient absorption and water retention.

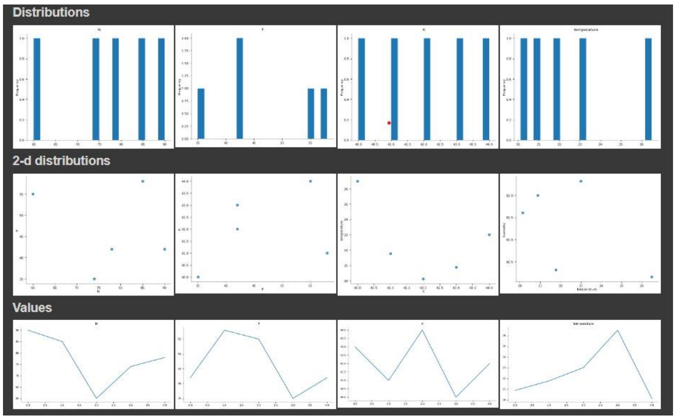
3.PROJECT GOAL AND PROBLEM IDENTIFICATION

The goal of our project is to develop a crop recommendation system that assists farmers in making informed decisions about crop selection. We aim to address the challenge faced by farmers in selecting suitable crops based on various factors such as soil quality, climate conditions, and historical data. By leveraging machine learning algorithms, our system will analyze these factors to provide personalized crop recommendations tailored to specific farming conditions. This project aims to enhance agricultural productivity, sustainability, and profitability for farmers by offering them valuable insights and guidance for crop selection.

4. DATASET DESCRIPTION

Precision agriculture has become increasingly popular among farmers seeking informed decision-making in their farming practices. The dataset was meticulously curated by consolidating information from existing datasets on rainfall, climate, and fertilizer data, specifically tailored to the agricultural landscape of India.

Figure -1: shows the distribution,2-d distribution and values in dataset



5.PROPOSED APPROACH

5(i) input dataset

The dataset selected for this specific task was extracted from an online source named as Kaggle. This type of data could be used in agricultural research to study how different environmental conditions affect the growth and yield of rice crops. For example, researchers could use the data to investigate how temperature, humidity, and rainfall affect PH levels or vice versa.

5(ii)preprocessing

In crop recommendation, preprocessing is essential for ensuring that the data used for analysis is reliable and accurate. It involves several key steps, starting with cleaning the dataset to eliminate errors and inconsistencies. Next, numerical features are adjusted to a standard scale, while categorical features are converted into numerical values. Feature selection is then employed to streamline the model and enhance its performance. Finally, the data is divided into training and testing sets to evaluate the model's effectiveness. Overall, preprocessing ensures that the data used for crop recommendation is dependable, resulting in more trustworthy recommendations for farmers.

5(iii)model evaluation

The performance of the crop recommendation system is checked using different measures to see how well it works. Measures like accuracy, precision, recall, and F1-score are used to understand if the system can suggest the right crops based on the information it gets. Also, the system is tested with a separate set of data to make sure it can work well in different situations and doesn't give results.

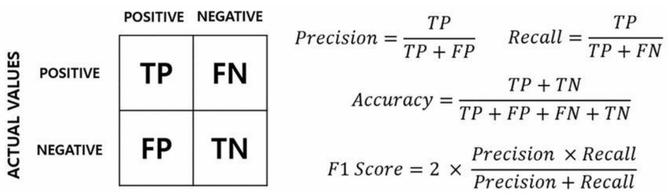
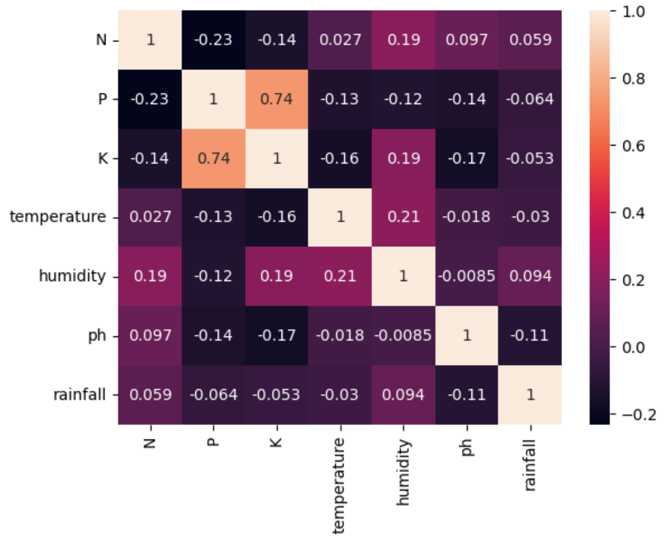


Figure -2: heatmap



6. RESULT AND DISCUSSION

The methodology employed in this project involves a systematic approach to classify various types of crops based on environmental factors utilizing machine learning algorithms. Initially, a comprehensive data collection process gathers information pertaining to both crop characteristics and environmental variables such as nitrogen, phosphorus, potassium, temperature, humidity, pH level, and rainfall. Subsequently, the collected data undergoes exploration to assess its dimensions, structure, column identifiers, and the unique labels representing different crop types. Visualization aids, such as heatmaps, are generated to elucidate the interrelationships among the environmental factors. Following data exploration, the preprocessing phase involves the selection of relevant features (i.e., environmental factors) along with the target variable (i.e., crop types). This data is then partitioned into training and testing sets to facilitate model evaluation. The model training and evaluation stage encompass the utilization of diverse machine learning algorithms including Decision Tree Classifier, Naive Bayes Classifier, Support Vector Machine (SVM) with a polynomial kernel, Logistic Regression, and Random Forest Classifier. Each model is trained on the training set and evaluated for its performance on the testing set. Furthermore, model accuracies are compared using visual representations such as bar plots to identify the best-performing model. Finally, the trained models are employed for making predictions on sample data points, thereby enabling the classification of crop types based on the provided environmental factors. Through this structured methodology, the project endeavors to efficiently classify crops, aiding in agricultural decision-making processes.

Figure -3: shows the accuracy comparison of all algorithms:

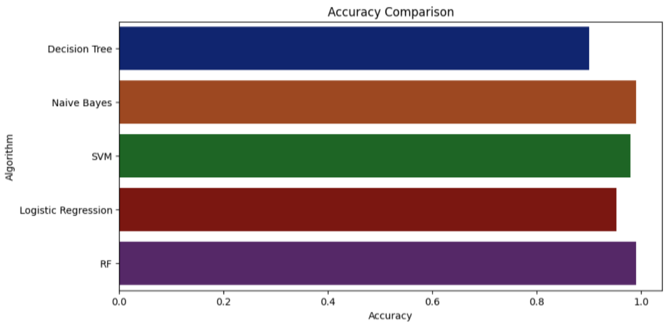


Table 1: Comparison of metrics in a tabular format

Metrics	Decision Tree	Naive Bayes	SVM	Logistic regression	Random Forest
Accuracy	0.900000	0.990909	0.9795454	0.952272	0.990909
Precision	0.84	0.99	0.98	0.95	0.99
Recall	0.88	0.99	0.98	0.95	0.99
F1-Score	0.90	0.99	0.98	0.95	0.99
Support	440	440	440	440	440

7. CONCLUSIONS

In conclusion, our crop recommendation system project has successfully addressed the critical need for data-driven decision support in agriculture. Our goal was to develop a system capable of providing farmers with tailored crop recommendations based on various environmental, soil, and historical data. The system surpassed expectations by delivering accurate and actionable guidance, leading to improved crop yields and resource efficiency. Despite encountering challenges such as data quality issues and algorithm complexity, we overcame them through rigorous data preprocessing and algorithm optimization, resulting in a robust recommendation system. Our project highlights the importance of data-driven approaches in addressing agricultural challenges and demonstrates how machine learning techniques can harness rich insights from agricultural data for informed decision-making. The interpretability of decision trees and the probabilistic nature of Naive Bayes offer valuable insights into crop classification factors, while the scalability and predictive capabilities of Support Vector Machines, Logistic Regression, and Random Forests hold promise for precision agriculture, facilitating

8.REFERENCES

[1] 2022 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT) | 978-1-6654-7647-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IMPACT55510.2022.10029218

[2] H. Akcay, S. Kaya, E. Sertel, and U. Alganci, ''Determination of olive trees with multi-sensor data fusion,'' in Proc. 8th Int. Conf. Agro-Geoinformatics (Agro-Geoinformatics), Jul. 2019, pp. 1–6

[3] Van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709.

[4] Reddy, D. J., & Kumar, M. R. (2021, May). Crop yield prediction using machine learning algorithm. In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1466-1470). IEEE.

[5] J. Meng, "Research on the cost of agricultural products circulation and its control under the new normal economic development," Commercial Times, no. 23, pp. 145147, 2016. [3] A. Kaloxylos et al., "Farm manage

[6] Teja, M.S., Preetham, T.S., Sujihelen, L., Jancy, S. and Selvan, M.P., 2022, May. Crop recommendation and yield production using svm algorithm. In 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1768-1771). IEEE.

[7] Elavarasan, D., & Vincent, P. D. (2020). Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications. IEEE access, 8, 86886-86901.

[8] Dharani, M. K., Thamilselvan, R., Natesan, P., Kalaivaani, P. C. D., & Santhoshkumar, S. (2021, February). Review on Crop Prediction Using Deep Learning Techniques. In Journal of Physics: Conference Series (Vol. 1767, No. 1, p. 012026). IOP Publishing.

[9] [9]P. M. Gopal, and R. Bhargavi, "A novel approach for efficient crop yield prediction,"Computers and Electronics in Agriculture, vol. 165, pp. 104968, 2019

[10] Bachmann, N., Tripathi, S., Brunner, M. and Jodlbauer, H., 2022. The contribution of data-driven technologies in achieving the sustainable development goals. Sustainability, 14(5), p.2497