IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13845

CROP RECOMMENDATION SYSTEM

Bhoomika.S[1], Prof. Narasimha Murthy M R[2]

Post-Graduation Student, Department of MCA, Vidya Vikas Institute of Engineering and Technology, Mysore, Karnataka[1]

Assistant Professor, Department of MCA, Vidya Vikas Institute of Engineering and Technology, Mysore, Karnataka[2]

Abstract: Agriculture plays a vital role in global food security, and farmers constantly seek ways to optimize their crop selection to maximize yields and profits. However, selecting the most suitable crop for a specific region or environment can be challenging due to various factors such as climate, soil conditions, and water availability. The Crop Recommendation System addresses this challenge by leveraging machine learning algorithms to analyze environmental data and provide personalized crop recommendations.

I. INTROUDCTION

The Crop Recommendation System developed in this project leverages cutting-edge machine learning techniques, particularly the Random Forest algorithm, to provide personalized crop recommendations based on a detailed analysis of environmental factors. The model is trained using a diverse dataset that includes essential parameters such as district, rainfall, humidity, minimum and maximum temperatures, soil moisture, and pH levels. By meticulously examining these variables, the system aims to equip farmers and agricultural stakeholders with actionable insights, allowing them to make informed decisions about crop selection. This data-driven approach not only enhances agricultural productivity but also supports sustainable farming practices by ensuring that crop choices are well-suited to the specific environmental conditions of each region. By integrating these advanced technologies, the system contributes to more efficient resource use, reduced environmental impact, and improved crop yields, ultimately fostering a more resilient agricultural sector.

Problem Statement: Farmers often face challenges in selecting the most suitable crops for cultivation due to varying environmental factors such as soil type, climate conditions, and regional characteristics. Traditional methods of crop selection are often based on historical practices or limited empirical data, leading to suboptimal yields and inefficiencies in resource use. This lack of precision in crop selection can result in reduced agricultural productivity, environmental degradation, and economic losses. There is a need for a data-driven solution that can analyse multiple environmental parameters and provide personalized crop recommendations to help farmers make informed decisions, optimize resource utilization, and improve sustainability in agriculture.

Scope: This project can be used in various agricultural settings, including farms, agricultural research institutions, and agricultural extension services. Agricultural experts and extension agents can also benefit from the system by providing tailored recommendations in their respective regions. Additionally, agricultural researchers can use the system to analyze historical data and trends to gain insights into crop suitability and adaptation to changing environmental factors. Overall, the project has the potential to support sustainable agriculture practices and improve crop yields in diverse agricultural contexts

II. LITERATURE SURVEY

Literature Review 1: Crop Recommender System Using Machine Learning Approach
By: Shilpa Mangesh Pande, Prem Kumar Ramesh, Anmol Anmol, B. R Aishwarya, Karuna Rohilla
Agriculture and its allied sectors are undoubtedly the largest providers of livelihoods in rural India. The agriculture sector is also a significant contributor factor to the country's Gross Domestic Product (GDP). Blessing to the country is the overwhelming size of the agricultural sector. However, regrettable is the yield per hectare of crops in comparison to international standards. This is one of the possible causes for a higher suicide rate among marginal farmers in India. This paper proposes a viable and user-friendly yield prediction system for the farmers. The proposed system provides connectivity to farmers via a mobile application. GPS helps to identify the user location. The user provides the area & soil type as input. Machine learning algorithms allow choosing the most profitable crop list or predicting the crop yield for a user-selected crop. To predict the crop yield, selected Machine Learning algorithms such as Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Multivariate Linear Regression (MLR), and K-Nearest Neighbour (KNN) are used. Among them, the Random Forest showed the best results with 95% accuracy. Additionally, the system also suggests the best time to use the fertilizers to boost up the yield.

DOI: 10.17148/IJARCCE.2024.13845

Literature Review 2: Crop Recommendation System using Machine Learning Algorithms.
By: Gaurav Chauhan; Alka Chaudhary
India is a predominantly agricultural country, with agriculture playing animportant part in the Indian economy and people's lives. Crops are recommended based on soil, weather, humidity, rainfall, and other variables to increase agricultural output. It benefits not just farmers, but also the country and helps to keep food costs down.This paper presents the utilisation of machine learning approaches like Random Forest and Decision Tree to predict which crop is best for which soil type based on the data sets.

Literature Review 3: Crop Prediction using Machine Learning Approaches
By: Nischitha K, Dhanush Vishwakarma, Ashwini, Mahendra N, Manjuraju M.R
As we know the fact that, India is the second largest population country in the world and majority of people in India have agriculture as their occupation. Farmers are growing same crops repeatedly without trying new verity of crops and they are applying fertilizers in random quantity without knowing the deficient content and quantity. So, this is directly affecting on crop yield and also causes the soil acidification and damages the top layer. So, we have designed the system using machine learning algorithms for betterment of farmers. Our system will suggest the best suitable crop for particular land based on content and weather parameters. And also, the system provides information about the required content and quantity of fertilizers, required seeds for cultivation. Hence by utilizing our system farmers can cultivate a new variety of crop, may increase in profit margin and can avoid soil pollution.

Literature Review 4: Crop Recommendation System using Machine Learning
By: Dhruvi Gosai, Chintal Raval, Rikin Nayak, Hardik Jayswal
A vast fraction of the population of India considers agriculture as its primary occupation. The production of crops plays an important role in our country. Bad quality crop production is often due to either excessive use of fertilizer or using not enough fertilizer. The proposed system of IoT and ML is enabled for soil testing using the sensors, is based on measuring and observing soil parameters. This system lowers the probability of soil degradation and helps maintain crop health. Different sensors such as soil temperature, soil moisture, pH, NPK, are used in this system for monitoring temperature, humidity, soil moisture, and soil pH along with NPK nutrients of the soil respectively. The data sensed by these sensors is stored on the microcontroller and analyzed using machine learning algorithms like random forest based on which suggestions for the growth of the suitable crop are made. This project also has a methodology that focuses on using a convolutional neural network as a primary way of identifying if the plant is at risk of a disease or not.

Literature Review 5: Machine Learning based Recommender Systems for Crop Selection
By: Younes Ommane, Mohamed Amine Rhanbouri, Hicham Chouikh, Ikram Chairi
Crop selection (CS) is one of the most critical elements that affects the final yield directly. As a result, selecting an appropriate crop is always a critical decision that a farmer must make, considering environmental factors. Choosing an appropriate crop for a given farm is a difficult decision including a plethora of variables that influence the final yield. Experts are frequently consulted to assist farmers with CS; but, as this alternative is time consuming and expensive, it is not available to many farms. The use of recommender systems (RSs) in agricultural management has recently brought some captivating and promising results. We propose a systematic literature review (SLR) in this article to find and provide the most relevant and high-quality publications ad- dressing the crop recommendation (CR) question. The core concept of this SLR is inspired from the guidelines of PRISMA 2020.The different CR approaches are discussed, as well as all the most important input features for recommendation, which are determined and classified. We also identified some of the biggest hurdles to using CR in agriculture. Besides, we made an inventory of the most used techniques for CR. Further, we made an inventory of evaluation criteria and evaluation approaches.

III. METHODOLOGY

Background Study
Existing methods of crop recommendation often rely on manual observations, local knowledge, and general agricultural practices passed down through generations. Farmers typically make decisions based on their own experiences, knowledge of the local climate and soil conditions, and advice from neighboring farmers or agricultural extension officers. Extension services may also provide general recommendations based on regional climate patterns and historical crop performance data. However, these methods may lack precision and may not consider specific variations in soil quality, climate conditions, and other environmental factors that can significantly impact crop suitability.

1. Rely heavily on the subjective judgment and experiences of individual farmers or agricultural experts, which can vary widely and may not always be accurate or consistent.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

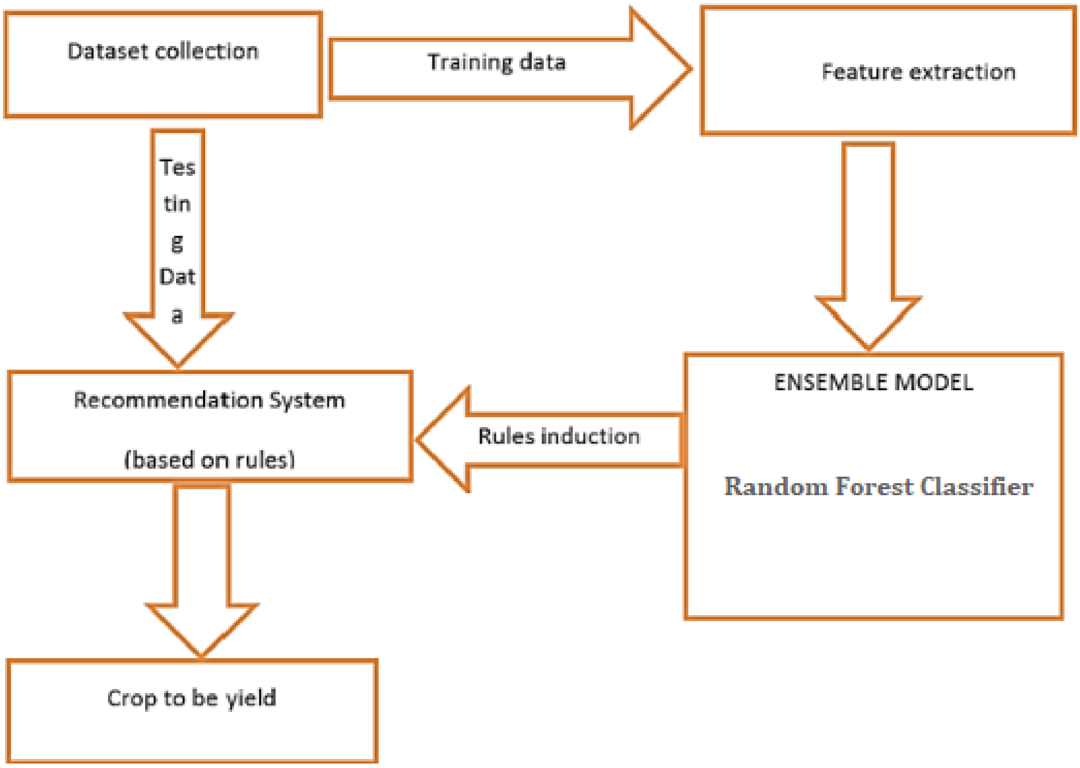
International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13845

2. May not consider a comprehensive range of factors affecting crop growth, such as soil quality, microclimates, and pest/disease prevalence. This limited scope can lead to less accurate recommendations and suboptimal crop choices.

3. Require significant time and effort from farmers or agricultural experts to collect and analyze relevant information manually. This can be inefficient and may result in delayed or suboptimal decision-making.



Proposed Methodology

Random Forest can be used for crop recommendation by leveraging its ability to handle large datasets with numerous input variables efficiently. In this context, the algorithm can analyze agricultural data comprising various factors such as district, rainfall, humidity, temperature, moisture, and soil pH to predict suitable crop choices for specific locations or conditions.

By constructing multiple decision trees and aggregating their predictions, Random Forest can generate robust recommendations based on the collective insights gleaned from the dataset. Additionally, the algorithm's flexibility allows it to capture complex relationships between input variables and crop outcomes, enabling it to adapt to diverse agricultural scenarios and provide accurate recommendations tailored to specific environments or farming practices.

1. Random Forest is known for its high accuracy in prediction tasks. It can effectively handle large datasets with numerous input variables, resulting in more precise crop recommendations based on various agricultural factors.

2. Random Forest mitigates the risk of overfitting, a common issue in machine learning models, by constructing multiple decision trees and averaging their predictions. This ensemble approach helps generalize the model's predictions to unseen data, improving its reliability in real-world scenarios.

3. The algorithm provides insights into the importance of different input variables (e.g., rainfall, humidity, temperature) in predicting crop outcomes. This information can be valuable for agricultural experts to understand which factors have the most significant impact on crop suitability.

Random Forest can handle missing values in the dataset without the need for imputation techniques. It leverages the majority voting mechanism during training, ensuring that missing values do not significantly impact the model's performance.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 13, Issue 8, August 2024
DOI: 10.17148/IJARCCE.2024.13845

ALGORITHM

Random Forest

The Random Forest technique is an ensemble learning approach that boosts the lifetime and accuracy of the project by creating many decision trees using bootstrapping sampling and feature bagging. Every tree assures variety and lowers overfitting by employing a randomly selected portion of features at every division and being trained on a randomly selected part of the data. During the training stage, these trees combine to generate the Random Forest model.

The model predicts in classification problems the class with most votes. In regression problems, the model finds the average of the predictions. This approach increases both prediction accuracy and the ability to understand the relevance of characteristics, so it is a great tool for managing complicated, high-dimensional datasets in the given project.

Ensemble Learning: Random Forest belongs to the ensemble learning methods where multiple models (in this case, decision trees) are trained and their outputs are combined to improve overall performance.

Decision Trees: A decision tree is a model in which the data is split into subsets based on the value of input attributes, therefore generating a tree structure wherein each leaf indicates a class (for classification) or a value (for regression).

Bagging (Bootstrap Aggregating): Random Forest uses a technique called bagging where multiple subsets of the training data are created by random sampling with replacement. Each decision tree is trained on a different subset.

Feature Randomness: Every decision tree's training involves a random subset of characteristics chosen at each split, which decorrelates the trees and boosts the model's resilience and variety.

Steps in the Random Forest Algorithm

Create Bootstrap Samples:
Generate and different bootstrap samples from the original dataset. Each sample is created by randomly selecting observations with replacement.

Build Decision Trees: Grow a decision tree for every bootstraps sample. When dividing a node, only examine a random subset of features (usually, the square root of the entire number of features for classification tasks).

Voting/Averaging: Every tree in the forest votes for a class for categorization; the class with the majority votes is then the final prediction.The average of all the trees' predictions serves as the last output for regression.

Advantages of Random Forest

Improved Accuracy: By combining the predictions of multiple trees, the model often achieves higher accuracy compared to individual decision trees.

Robustness: It reduces the risk of over fitting since individual trees are trained on different subsets of the data and features.

Feature Importance: Random Forest can provide insights into the importance of different features in making predictions.

Versatility: It can handle both classification and regression tasks and works well with large datasets.

Disadvantages of Random Forest

Complexity: The model is more complex and computationally intensive compared to a single decision tree.

Interpretability: The combined model is less interpretable than individual decision trees.

Training Time: Training multiple trees can be time-consuming, especially with large datasets and many trees.

DOI: 10.17148/IJARCCE.2024.13845

IV. RESULTS & DISCUSSION

Random Forest is a powerful machine learning algorithm used for both classification and regression tasks. In the context of crop recommendation, Random Forest can be applied as follows:

Data Preprocessing: Before training the Random Forest model, the dataset needs to be preprocessed. This involves steps such as handling missing values, encoding categorical variables (if any), and splitting the dataset into training and testing sets.

Feature Selection: Identify the relevant features (independent variables) from the dataset that influence crop recommendation. These features may include climatic factors (e.g., rainfall, temperature, humidity), soil characteristics (e.g., pH, moisture), geographical attributes, and crop-specific parameters.

Model Training: Once the dataset is prepared, the Random Forest algorithm is trained using the training data. During training, the algorithm builds multiple decision trees based on random subsets of the features and samples from the training dataset. Each decision tree is trained independently.

Prediction: Once the model is trained and evaluated, it can be deployed to predict the suitable crops and fertilizers for a given set of input features. Agricultural experts can input environmental and soil parameters into the model, and it will output the recommended crops based on its learned patterns and relationships.

By following these steps, Random Forest can effectively analyze agricultural data and provide accurate recommendations for crop selection based on various environmental and soil conditions.

V. CONCLUSION

In conclusion, the Crop Recommendation System developed using the Random Forest algorithm presents a promising approach to assist farmers and agricultural stakeholders in making informed decisions about crop selection. By leveraging machine learning techniques and agricultural data such as district, rainfall, humidity, temperature, and soil properties, the system offers personalized crop recommendations tailored to specific environmental conditions. Through the ensemble learning approach of Random Forest, the model achieves high accuracy and robustness while effectively handling missing data and identifying feature importance. The implementation of this system has the potential to enhance agricultural productivity, optimize resource utilization, and contribute to sustainable farming practices. Moving forward, continued refinement and validation of the model, along with integration with user-friendly interfaces for farmers, will be crucial for its successful adoption and impact in real-world agricultural settings.

REFERENCES

[1]. Narasimha Murthy MR, "Future Scope of Artificial Intelligence in Software Engineering", International Journal of Science and Research (IJSR), Volume 12 Issue 12,December2023,pp.1401-1402, https://www.ijsr.net/getabstract.php?paperid=SR231113100032

[2]. B. Choudhary, S. Dewangan, and A. Shukla, "Crop Recommendation System Using Machine Learning Algorithms," International Journal of Computer Applications, vol. 180, no. 42, 2018.

[3]. K. Patel and K. Patel, "Survey on Crop Yield Prediction Using Machine Learning Techniques," International Journal of Computer Applications, vol. 179, no. 40, 2018.

[4]. S. Tiwari, D. Sahu, S. Shrivastava, and A. Uikey, "A Crop Recommendation System Using Machine Learning Algorithms and Ensemble Technique,"International Journal of Scientific & Technology Research, vol. 9, no. 3, 2020.

[5]. R. K. Alagukannan, "Crop Yield Prediction Using Machine Learning Algorithms," International Journal of Advanced Science and Technology, vol. 29, no. 3, pp. 3574-3580, 2020.

[6]. K. Jayashree, S. M. Antony, R. Janarthanan, and R. S. Anitha, "Smart Farming: Crop Recommendation System Using Machine Learning," International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 5, no. 6, pp. 50-54, 2019.

[7]. S. Garg, A. Sharma, and A. Kumar, "Crop Recommendation System Using Data Mining and Machine Learning," International Journal of Computer Science and Engineering, vol. 6, no. 9, pp. 37-42, 2018.

[8]. P. Kamble, A. Patil, and P. S. Patil, "Farming Assistance Based on Crop Prediction Using Machine Learning," Journal of Emerging Technologies and Innovative Research, vol. 6, no. 6, pp. 545-549, 2019.

[9]. R. Prajapati, A. A. Prajapati, and M. V. Prajapati, "Crop Recommendation System Using Machine Learning for Precision Agriculture," International Journal of Engineering Development and Research, vol. 7, no. 2, pp. 238-242, 2019.

[10]. M. A. Haque, M. M. Rahman, and M. Rahman, "Development of Crop Selection Model Using Data Mining and Machine Learning Algorithms," International Journal of Scientific & Technology Research, vol. 8, no. 11, pp. 3293-3297, 2019.

