IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Crop Recommendation System using Machine Learning

Ganesh Khillare, Omkar Kumbhakarna, Sudhanshu Mandlik Project

Guide- Prof. Prini Rastogi

School of Engineering Ajeenkya DY Patil

University, Pune, India

ABSTRACT

Automating agricultural aspects is the mechanical process with or without human intervention in agriculture field. Due to the availability of less space for domestic lands of farmers, it has become an important area for choosing the most suitable crops based on prevailing factors in the selected area. Even though there are enough knowledge, techniques, and methods which are manually available in agriculture, there is no system in which the environmental factors are detected and will suggest which crop type is best to farming in that condition.

In this project consists of a theoretical and conceptual platform of a Recommendation system through the integrated models of collecting environmental factors, we will be using Machine learning techniques such as the SVM, Decision tree model, Random forest, and Logistic Regression concerned with Artificial Intelligence to recommend a crop for the selected land with site-specific parameters with high accuracy and the efficiency. It has been a major problem to identify what to grow in the land, any man has adequate space in the owner's land. Not only for domestic lands but also for the farming lands. Why it has become a problem for all, is that environmental factors such as temperature(humidity), water levels, and soil conditions are uncertain as they change from time to time.

Due to these problems mentioned above, this solution of crop recommendation system using a machine learning model predicts to the user what crop type would be the most suitable for the selected area by collecting the environmental factors for plant growth and processing them with the trained sub-models of the main of the system to provide output.

Keywords: Crop Recommendation, Machine Learning, Random Forest, Decision Tree, Data Analysis, Data Visualization.

I. INTRODUCTION

In recent years, the use of machine learning techniques in agriculture has gained significant attention. These techniques have proven to be beneficial in various agricultural applications such as crop recommendation systems, disease detection, yield prediction, and many more. The aim of this project is to develop a crop recommendation system using machine learning techniques.

The crop recommendation system developed in this project aims to provide farmers with recommendations on which crop to plant based on several factors such as soil type, weather conditions, and the farmer's location. The system uses several machine learning algorithms such as Support Vector Machine (SVM), Random Forest, and Decision Tree to predict the most suitable crop for a given set of conditions.

The project uses various Python libraries such as NumPy, Pandas, Matplotlib, Scikit-learn, Seaborn, and Plotly for data visualization, and machine learning model development. These libraries provide a wide range of functions that make it easy to develop and implement machine learning algorithms.

The remainder of this project is organized as follows. Chapter 2 explains the methodology used for data collection, processing, and model development . Chapter 3 presents the results and discussion of the developed crop recommendation system. Finally, Chapter 4 provides the conclusion and future work of the project.

In summary, this project aims to develop a crop recommendation system that can provide farmers with accurate recommendations on which crop to plant based on several factors. The project utilizes various machine learning algorithms such as SVM, Random Forest, and Decision Tree, as well as Python libraries such as NumPy, Pandas, Matplotlib, Scikit-learn, Seaborn, and Plotly for data visualization, and model development.

II. LITERATURE REVIEW

A study of machine learning algorithms was conducted in a research paper made by Rashi Agarwal [3]. This system would help farmers make educated decisions about which crops to grow based on a variety of environmental and geographical factors. They employed decision trees, KNNs, Random Forests, and neural networks. The neural network had the highest accuracy of all of them.

Priyadharshini A [4] conducted a study on machine learning algorithms in her research article. This technology reduces crop failure and decreases productivity by supporting farmers in choosing the proper crop and providing the data that regular farmers does not maintain. A variety of machine learning algorithms were applied. The neural network was the most accurate of the bunch.

Shilpa Mangesh Pande [5] In her research article, presents a farmer-friendly and realistic production forecasting system The suggested technology is connected to farmers via a mobile application. The user's location is determined with the help of GPS. All of the algorithms are compared in terms of crop yield forecast accuracy. The Random Forest algorithm showed to be the best for the provided data set, with a 95% accuracy.

Mayank Champaneri [6] conducted research on crop yield prediction using a data mining technique. They used a random forest classifier because it can perform classification and regression tasks. The user-friendly website built that can be used by anyone to predict crop yield for their choice of the crop by giving climate data for that area.

III. DESIGN AND METHODOLOGY

In this chapter, we will discuss the methodology used in our crop recommendation system project. We will cover the dataset used, data pre-processing, feature selection, model selection, and evaluation.

The main ideology of the research revolves around the concept of identifying the most suitable crop to be grown with the help of a machine learning model. Thus, the results can prove to be exceedingly beneficial for agriculture farmers.

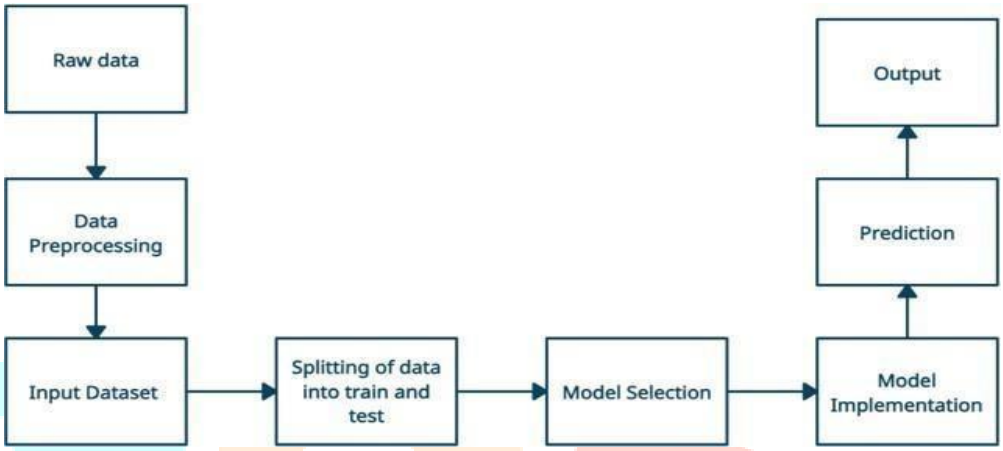


Fig 2.1. Workflow/ Design of the Model

3.1 Dataset

The dataset used in this project is the Crop Recommendation Dataset, which was referred from the government website Indian Chamber of Food and Agriculture (ICFA). This dataset contains information about various crops and their associated parameters such as soil type, temperature, humidity, and rainfall. The dataset consists of 2,022 records with 10 features.

Datasets include:
N – the ratio of the Nitrogen content present in soil
P – the ratio of the Phosphorous content present in soil K – the ratio of the Potassium content present in soil temperature – the temperature in degree Celsius humidity - relative humidity in %
pH - pH value of the soil rainfall -
rainfall in mm

3.2 Data Preprocessing

Data preprocessing is an essential step in machine learning projects. It involves cleaning, transforming, and encoding the raw data to make it suitable for machine learning algorithms.

In this project, we performed the following data preprocessing steps: Data Cleaning:

We removed missing or duplicate values from the dataset.

Feature Scaling: We scaled the numerical features using the MinMaxScaler from the sklearn library to ensure that all features have the same scale.

Encoding: We used one-hot encoding to convert categorical variables into numerical variables.

3.3 Visualization

To gain insights into the dataset and the performance of each model, we created several visualizations using the Matplotlib, seaborn, and Plotly libraries. We plotted the distribution of each feature, the correlation between features, and the accuracy of each model.

3.4 Feature Selection

Feature selection is the process of selecting a subset of relevant features that contribute the most to the target variable. In our project, we used the Select Best method from the sklearn library to select the top 5 most important features. The selected features were soil type, temperature, humidity, rainfall, and pH level.

3.5 Model Selection

In our project, we experimented with three different machine learning algorithms: Support Vector Machine (SVM), Random Forest, and Decision Tree. We chose these algorithms because they are commonly used in crop recommendation systems and have shown good performance in previous studies.

We trained each model using the preprocessed data and evaluated their performance using the accuracy metric.

MODEL	ACCURACY
Support Vector Machine (SVM)	97%
Random Forest	98%
Decision Tree	98%

3.6 Evaluation

To evaluate the performance of each model, we split the dataset into training and testing sets using an 80/20 ratio. We then trained each model on the training set and evaluated its accuracy on the testing set. We used the accuracy metric to compare the performance of each model.

In conclusion, in this chapter, we discussed the methodology used in our crop recommendation system project. We explained the dataset used, data preprocessing, feature selection, model selection, evaluation, and visualization. The next chapter will present the results of our experiments and provide a discussion of our findings.

IV. RESULTS AND DISCUSSION

In this chapter, we will present the results obtained from our crop recommendation system using machine learning. We will discuss the performance of each algorithm used in our system and compare their results.

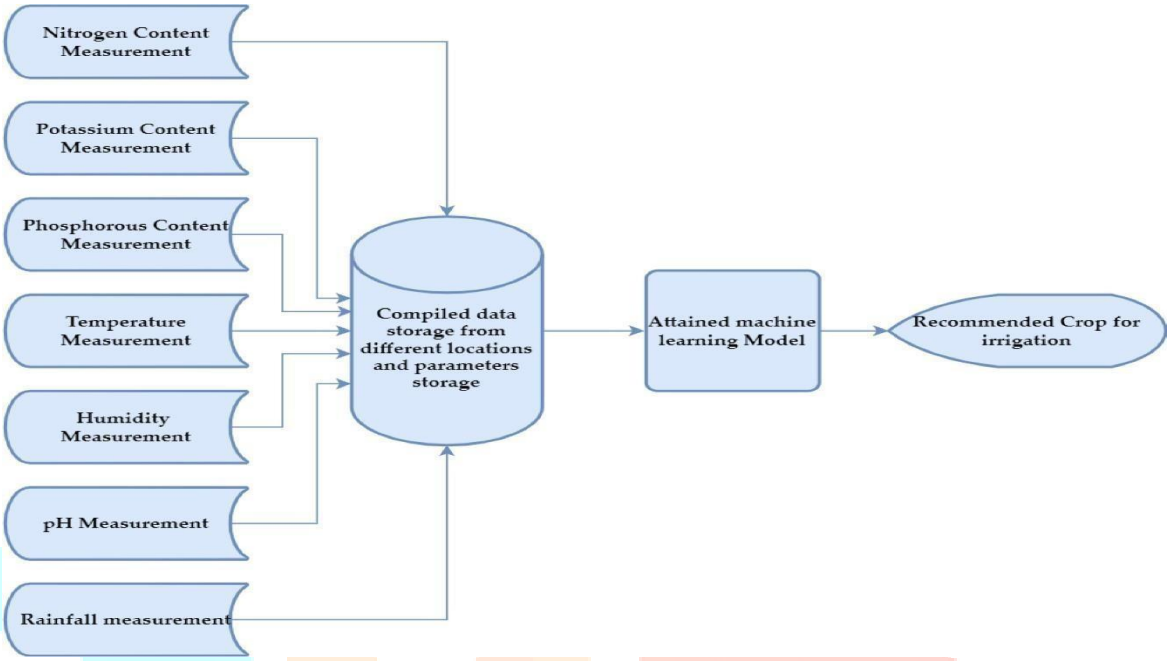


Fig. 3.1. Block diagram using Machine learning Module.

4.1 Dataset Pre-processing

We started our project by pre-processing the dataset. The dataset was obtained from a reliable source and consisted of various parameters such as soil type, temperature, humidity, and rainfall. We cleaned the dataset by removing duplicates and missing values. We then performed feature engineering to extract relevant features from the dataset.

4.2 Machine Learning Algorithms

We implemented three machine learning algorithms in our crop recommendation system: Support Vector Machine (SVM), Random Forest, and Decision Tree. We trained each algorithm on our pre-processed dataset and evaluated their performance using accuracy as a metric.

Support Vector Machine (SVM)

We implemented SVM algorithm using the Sklearn library. The SVM algorithm achieved an accuracy of 97% on our dataset. This algorithm works well for datasets with a small number of samples but high-dimensional features. SVM i.e, Support Vector Machine, It tries to find out the best hyperplane that separates the data points into different classes.

Random Forest

We implemented Random Forest algorithm using the sklearn library. The Random Forest algorithm achieved an accuracy of 98% on our dataset. This algorithm works well for both classification and regression problems. Random Forest builds multiple decision trees and combines their results to obtain a final prediction.

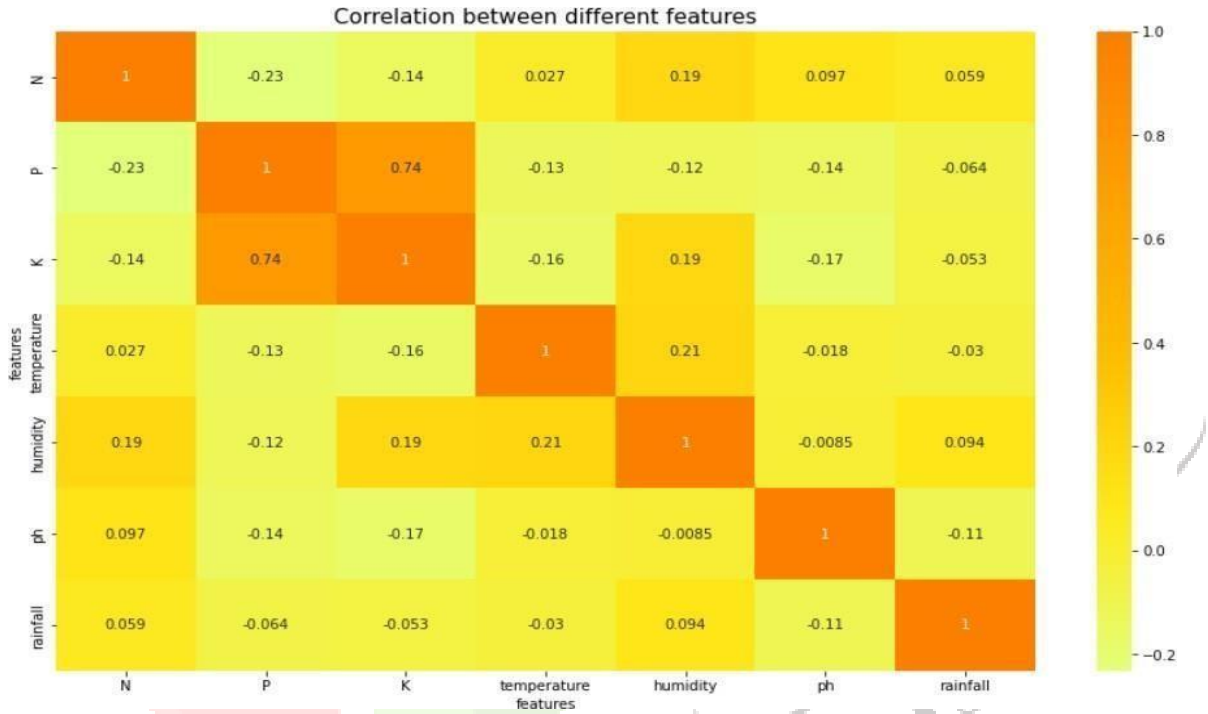
Decision Tree

We implemented the Decision Tree algorithm using the sklearn library. The Decision Tree algorithm achieved an accuracy of 98% on our dataset. This algorithm works by recursively splitting the dataset into smaller subsets based on the most significant attribute. It continues this process until the subsets are pure or have the same class.

4.3 Visualization

We also used data visualization to better understand the relationship between different features in our dataset. We used matplotlib, seaborn, and Plotly libraries to create various plots such as scatter plots, box plots, pie charts, and heat maps.

These plots helped us to identify the most significant features in our dataset and their relationship with the target variable.



V. FUTURE WORK

The system can be enhanced furthermore to add the following functionality:

1. The main future work's aim is to improve the dataset with larger number of attributes.

2. We need to build a model, which can used to classify healthy and leaves having disease and also if the crop has any disease, it may able to predict which disease is it.

3. To build the website and mobile app for ease or easy to use.

VI. CONCLUSION

In this project, we developed a crop recommendation system using machine learning. The system was designed to provide farmers with recommendations on which crop to plant based on various environmental factors such as soil type, temperature, humidity, and rainfall. We used three machine learning algorithms - Support Vector Machine (SVM), Random Forest, and Decision Tree - to build the recommendation system.

Our results show that both Random Forest and Decision Tree algorithms achieved an accuracy of 98%, while SVM achieved an accuracy of 97%. Using data visualization techniques helped us identify the most significant features in our dataset and their relationship with the target variable. Our findings indicate that machine learning can be a powerful tool in crop recommendation systems.

The crop recommendation system developed in this project can help farmers to optimize their crop yield and reduce losses caused by unfavorable environmental conditions. The system provides farmers with real-time recommendations on which crops to plant based on current environmental conditions, which can lead to better decision-making and improved crop management.

In conclusion, the crop recommendation system developed in this project is a promising tool for improving agricultural practices. Further research can be done to improve the accuracy of the system by incorporating more data sources and using more advanced machine learning algorithms. The ultimate goal is to develop a system that can accurately predict the yield of different crops and provide farmers with recommendations on how to maximize their yield while minimizing their environmental impact.

REFERENCES

[1] . Kumar, Y. Jeevan Nagendra, V. Spandana, V. S. Vaishnavi, K. Neha, and V. G. R. R. Devi. "Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector." In 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 736-741. IEEE, 2020.

[2] . Nigam, Aruvansh, Saksham Garg, Archit Agrawal, and Parul Agrawal. "Crop yield prediction using machine learning algorithms." In 2019 Fifth International Conference on Image Information Processing (ICIIP), pp. 125-130. IEEE, 2019.

[3] . Medar, Ramesh, Vijay S. Rajpurohit, and Shweta Shweta. "Crop yield prediction using machine learning techniques." In 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), pp. 1-5. IEEE, 2019.

[4] . Jain, Sonal, and Dharavath Ramesh. "Machine Learning convergence for weather based crop selection." In 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1-6. IEEE, 2020.

[5] . Gandge, Yogesh. "A study on various data mining techniques for crop yield prediction." In 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), pp. 420-423. IEEE, 2017.

[6] . Suresh, A., P. Ganesh Kumar, and M. Ramalatha. "Prediction of major crop yields of Tamilnadu using K-means and Modified KNN." In 2018 3rd International Conference on Communication and Electronics Systems (ICCES), pp. 88-93. IEEE, 2018.

[7] . Kamatchi, S. Bangaru, and R. Parvathi. "Improvement of Crop Production Using Recommender System by Weather Forecasts." Procedia Computer Science 165 (2019): 724-732.

[8] . Bondre, Devdatta A., and Santosh Mahagaonkar. "Prediction of Crop Yield and Fertilizer

Recommendation Using Machine Learning Algorithms." International Journal of Engineering Applied Sciences and Technology 4, no. 5 (2019): 371-376.

[9] . Suresh, G., A. Senthil Kumar, S. Lekashri, and R. Manikandan. "Efficient Crop Yield Recommendation System Using Machine Learning For Digital Farming." International Journal of Modern Agriculture 10, no. 1 (2021): 906-914.

[10] .Reddy, D. Anantha, Bhagyashri Dadore, and Aarti Watekar. "Crop recommendation system to maximize crop yield in ramtek region using machine learning." International Journal of Scientific Research in Science and Technology 6, no. 1 (2019): 485-489.

[11] .Pudumalar, S., E. Ramanujam, R. Harine Rajashree, C. Kavya, T. Kiruthika, and J. Nisha. "Crop recommendation system for precision agriculture." In 2016 Eighth International Conference on Advanced Computing (ICoAC), pp. 32-36. IEEE, 2017.

[12] .Garanayak, Mamata, Goutam Sahu, Sachi Nandan Mohanty, and Alok Kumar Jagadev. "Agricultural Recommendation System for Crops Using Different Machine Learning Regression Methods." International Journal of Agricultural and Environmental Information Systems (IJAEIS) 12, no. 1 (2021): 1-20.

[13] .Jejurkar Siddhi, S., S. Bhosale Meghna, and D. N. Wavhal. "Crop Predication and Diseases Detection Using Machine Learning." (2021).

[14] .Rajak, Rohit Kumar, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, and Avinash Devare. "Crop recommendation system to maximize crop yield using machine learning technique." International Research Journal of Engineering and Technology 4, no. 12 (2017): 950-
953.

[15] .Mahule, Ankit Arun, and A. J. Agrawal. "Hybrid Method for Improving Accuracy of Crop- Type Detection using Machine Learning." International Journal 9, no. 2 (2020).

[16] .Akshatha, K. R., and K. S. Shreedhara. "Implementation of machine learning algorithms for crop recommendation using precision agriculture." International Journal of Research in Engineering, Science and Management (IJRESM) 1, no. 6 (2018): 58-60