Crop Recommendation System using ML

Dr. D. Kavitha[1], R. Triveni[2], S. Harsha Vardhan[3], N. Akhil[4], V. Kumar ParasuRam[5]

Sr. Assistant Professor, Department of Information Technology[1]

B. Tech Students, Department of Information Technology[2,3,4,5]

Prasad V. Potluri Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India

Abstract: The crop recommendation system using machine learning is an intelligent decision support system that provides recommendations to farmers on the most suitable crop to cultivate based on soil and weather conditions like temperature, humidity, rainfall, nitrogen, potassium, phosphorus and pH value of the soil. This system uses machine learning algorithms like Decision Tree, Random Forest, Naïve Bayes, Support Vector Machine (SVM), and Logistic Regression to analyse data on soil properties, climate, and other relevant factors to generate personalized crop recommendations for each farmer.

Keywords: Crop Recommendation, temperature, humidity, rainfall, nitrogen, potassium, phosphorus, ph value, Decision Tree, Random Forest, Naïve Bayes, Support Vector Machine (SVM), Logistic Regression, machine Learning.

I. INTRODUCTION

The traditional methods of crop recommendation involve relying on the experience of the farmer or agricultural experts, which may not always provide accurate results. In contrast, the crop recommendation system using machine learning provides a more data-driven approach, allowing farmers to make informed decisions based on objective data analysis.

This system involves collecting data on various factors such as soil properties, climate, and crop yield, which are then analysed using machine learning algorithms. The algorithms use the collected data to develop models that can predict the most suitable crop to cultivate in a given area. The output of the system is a list of recommended crops with their estimated yield, helping farmers make informed decisions about which crop to cultivate to maximize their profits. It have the potential to revolutionize agriculture by improving crop yields, reducing input costs, and promoting sustainable farming practices. As the world population continues to grow, this technology will become increasingly important to ensure food security for future generations.
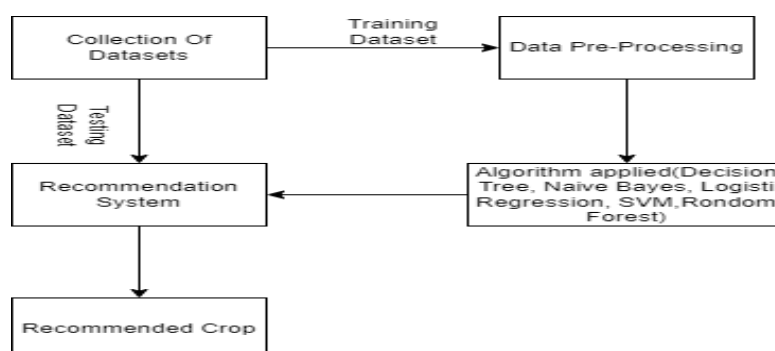
II. PROPOSED SYSTEM

Figure 1.1: Proposed system model

Proposed model recommends the crop based on rain fall, temperature, humidity, pH value of soil, nitrogen, phosphorus and potassium of the soil. It compares accuracy of the ML algorithms like Decision tree, Random Forest , SVM, Naive Bayes , logistic regression and use the one which have high accuracy.

III. METHODOLOGY

The following steps are involved while implementing the project in order to achieve the results.

3.1 Data Collection:

The first step is data collection which is most important to implement the project. Data collection is the process of gathering information.

3.2 Data Pre-Processing:

In this step the collected data will be cleaned and filtered. The real time data which was collected won't be in our desired format to use there will be some noisy data, missing values and redundant values etc. If we want accurate results the data should be accurate. In order to achieve the accurate data this step will be used to remove redundant values and to fill the missing data with appropriate values.

3.3 Model Building:

In this step we will build and train the recommendation system using the pre-processed data and the ML algorithms. To train the model the dataset need to be divided into two parts i.e., training data set and testing data set. By using the ML algorithms on training data set we will train and build the model.

3.4 Crop Recommendation:

In this project we used five ML algorithms, The accuracies of the five algorithms will be compared and the one which have highest accuracy will be used to recommend the suitable crop.

IV. DATASET

The dataset consists of parameters like Nitrogen(N), Phosphorous(P), Potassium(K), PH value of soil, Humidity, Temperature and Rainfall. The datasets have been obtained from the Kaggle website. The data set has 2200 instance or data that have taken from the past historic data. This dataset include eleven different crops such as rice, maize, chickpea, kidneybeans, pigeonpeas, moth beans, mung bean, blackgram, lentil, pomegranate, banana, mango, grapes, watermelon, muskmelon, apple, orange, papaya, coconut, cotton, jute, and coffee.

Attribute Description:

There are seven attributes in the dataset they are:

N: The ratio of nitrogen content in the soil

P: The ratio of phosphorous content in the soil

K: The ratio of potassium content in the soil

Temperature: Temperature in degree Celsius

Humidity: Relative humidity in %

pH: pH value of the soil

Rainfall: Relative rainfall in mm

V. RESULTS

We have used five algorithms to train and to build the model after data pre-processing step. After comparing the accuracies of the five ML algorithms the one which have the high accuracy will be used on testing data to recommend suitable crop.
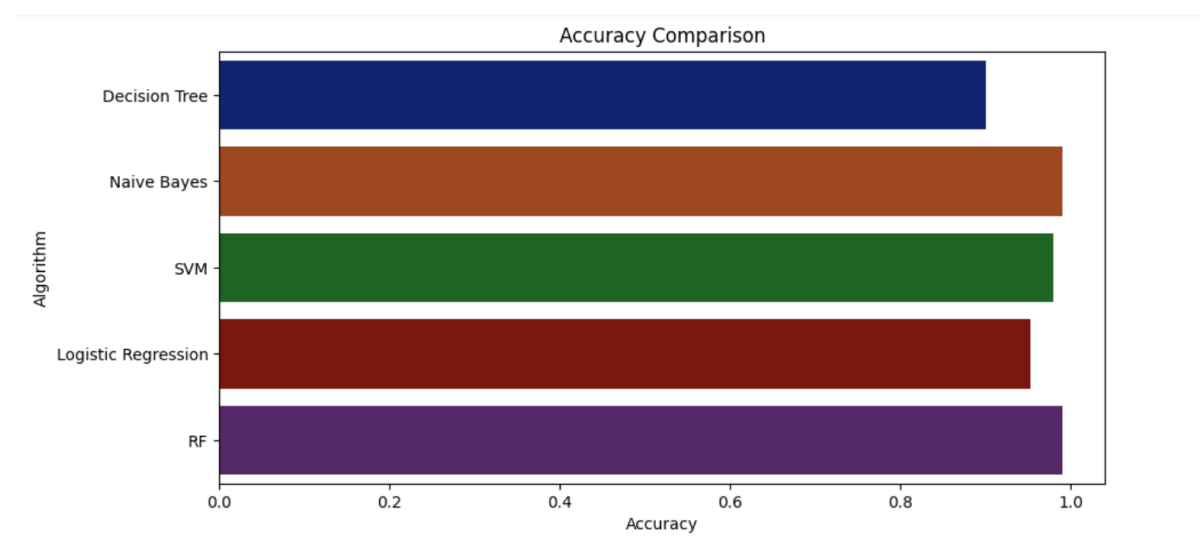

Figure 1.2: Accuracy comparison of 5 ML algorithms

```
accuracy_models = dict(zip(model, acc))
for k, v in accuracy_models.items():
    print (k, '-->', v)

Decision Tree --> 0.9
Naive Bayes --> 0.990909090909091
SVM --> 0.9795454545454545
Logistic Regression --> 0.9522727272727273
RF --> 0.990909090909091
```

Figure 1.3: Accuracy Description

From the comparison and description of accuracies we can see that both Random Forest and Naïve Bayes algorithms have the highest accuracy of 99%, followed by SVM of 97%, Logistic- Regression gave 95% of accuracy while Decision tree gave the 90% of the accuracy. So that we have used both Random Forest and Naïve Bayes algorithms to recommend the crop on test data.

Random Forest algorithm:

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. As the name suggests, Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset. Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.
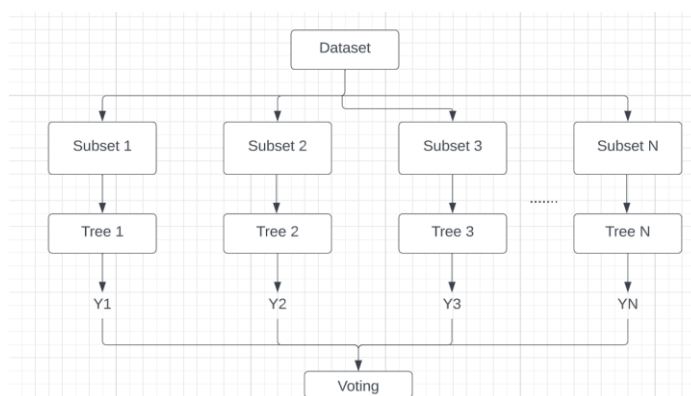

Figure 1.4: Working of Random Forest algorithm

Naïve Bayes algorithm:

Naive Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. Naive Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.

$$p\left(\frac{A}{B}\right) = \frac{P\left(\frac{B}{A}\right) * p(A)}{p(B)}$$

VI. FUTUTRE SCOPE

In this project we only did crop recommendation rather than that in future if we do crop fertilizer recommendation and crop disease prediction added to this existing project it will be very useful to the farmers. And also if we design a web application for the project in the future it would be very simple and easy to farmers to gain the related information.

VII. CONCLUSION

We have collected the dataset from the Kaggle website as it is the first step and then did the data pre-processing which is the second step followed by third step model building and training. For the model building and training we have used five algorithms i.e., Random Forest, Naïve Bayes, Logistic Regression, SVM and Decision Tree algorithms. By comparing the accuracies of five algorithms we came to know that both Random

Forest and Naïve Bayes algorithms have the high accuracy of 99% than other algorithms. So that we have used those two algorithms on test data to recommend the suitable crop.

VIII. REFERENCES

[1] https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset

[2] https://ijisrt.com/assets/upload/files/IJISRT20MAY722.pdf

[3] https://www.javatpoint.com/machine-learning-random-forest-algorithm

[4] https://www.javatpoint.com/machine-learning-naive-bayes-classifier

[5] Mrs. N. Saranya, Ms. A. Mythili" Classification of Soil and Crop Suggestion using Machine Learning Techniques" International Journal of Engineering Research & Technology (IJERT) http://www.ijert.org IJERTV9IS020315 (4.0 International License.) Vol. 9 Issue 02, February-2020