

## Synopsis for Minor Project-II

Roll No	Name of the Student
20CSE036	Anurag Thakur
20CSE038	Piyush Kumar
20CSE052	Priya Khetan

**Title Of Project: Crop disease recommendation using machine learning algorithm**

### **Introduction-**

Crop recommendation in agriculture science is the primary concern for every country, as the food demand is increasing at a fast rate due to an increase in population. Moreover, the increased use of technology today has increased the efficiency and accuracy of detecting diseases in crops. Most of the times it is observed that farmers tend to sow the crop according to its market value and possible financial profits rather than taking factors like soil conditions, sustainability etc. in to the account. The detection process marks the beginning of a series of activities to fight the diseases and reduce their spread. Some diseases are also transmitted between plants, making it hard to fight them. Machine learning algorithms allow choosing the most profitable crop list or predicting the crop yield for a user-selected crop. Our focus is to clarify the details about the diseases and how to detect diseases in crops automatically use of machine learning algorithm.

### **Methodology:**

While implementing the project, the following steps were implemented in order to achieve the results:

#### **1. Data Cleaning and Preprocessing-**

One of the first steps is to make sure that the dataset we are using is accurate. The dataset should not have any missing values and if the dataset does have missing values, they should be replaced by the appropriate values. The data should also be checked to see if there is a normal distribution for its features.

#### **2. Data Analysis and Visualization-**

After performing the cleaning and preprocessing of the data, we perform data analysis and visualizations on our dataset. We try to analyze our data more clearly to find any trends or patterns in the dataset. We have created several visualizations of our dataset in order to understand the data properly.

#### **3. Feature Selection-**

It is important that we select only those features that will be necessary to determine the type of crop to grow. For this, we have created a correlation matrix that shows the linear relationship of a feature with every other features. If features are highly correlated then that feature should be dropped.

#### **4. Building a UI-**

In the next step, we have built a UI for a user to input his data so that once he enters the information such as N, P, K values of soil, temperature, humidity, rainfall etc., the model will process the data and will recommend the appropriate type of crop to be grown in such a condition. Once the user enters the following values and submits the machine learning model will predict the crop that the person must grow. We will connect our machine learning model with our front end by using Flask.

## **5. Dataset-**

The dataset for this topic was taken from Kaggle. It is a crop recommendation dataset giving us information about various types of crops and the features that decide which crop is suitable for growing.

## **6. Machine Learning Algorithms Used-**

- **Random Forest**

Random Forest is a supervised ensemble machine learning algorithm used in both classification as well as regression problems. It contains various decision trees and an average of it is taken so as to give the output.

- **Decision Tree**

The working of it is based on a simple technique, wherein a yes/no question is asked and according to the answer the tree is split in smaller nodes. The split of the nodes can either happen by calculating Gini impurity (calculates the measure of impurity) or information gain (calculates the change in the entropy). Decision Trees are prone to overfit and hence this may lead to getting a lower accuracy. This problem can be solved by using random forest algorithm.

- **Logistic Regression**

It uses a sigmoid function to mathematically calculate the probability of an observation and accordingly, the observation is then put into its respective class.

- **XGBoost**

XGBoost is one of the most popular algorithms used today. It is a tree-based algorithm using gradient boosting framework. This algorithm is based on a feedback approach, where feedback from the decision tree is taken so as to improve the decision tree further (also known as boost) as it “boosts” the efficiency of the tree and helps in bringing out a better accuracy.

## **Future Scope**

At present we have limited our scope only to a crop recommendation system, but this in the future can be extended to other areas such as fertilizer recommendations, wherein a person can get the information of the fertilizer suitable for his crop. Another extended application of this could be plant disease classification using CNN and also giving a remedy to that disease. As agriculture is quite an unexplored area, the scope for such a type of project is tremendous. While we have made a working website for our model, in the future this can also be made into an app so that it is easier for the farmers to access. Both the website and

the app can be made in regional languages so that people would be more comfortable using them.

### **REFERENCES-**

- Keerthan Kumar T G, Shubha C, Sushma S A” Random Forest Algorithm for Soil Fertility Prediction and Grading using Machine Learning
- Mrs. N. Saranya, Ms. A. Mythili” Classification of Soil and Crop Suggestion using Machine Learning Techniques.
- Zeel Doshi, Subhash Nadkarni, Rashi Agrawal, Prof. Neepa Shah “AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms.
- Archana Chougule, Vijay Kumar Jha and Debajyoti Mukhopadhyay “Crop Suitability and Fertilizers Recommendation using Data Mining Techniques

**Group No:**

**Supervisor Name :**

**Supervisor  
Signature:**

**Class Teacher**

**Project Coordinator  
3<sup>rd</sup> year**

**HOD, CSE  
3<sup>rd</sup> year**