

**CROP RECOMMENDATION SYSTEM USING MACHINE LEARNING ALGORITHMS**

**E. Sateesh N**, Assistant professor, Department of Computer Science and Engineering, Amrita Sai institute of science and technology, Paritala, Kanchikacherla mdl, Andhra Pradesh India.

**Bhanusri, Sk. Khaja Mohiddin Pasha, Sk. Sameer, P. Gopala krishna, A. Sunitha**  
Graduate students, Department of Computer Science and Engineering, Amrita Sai institute of science and technology, Paritala, Kanchikacherla mdl, Andhra Pradesh India.

**Abstract**

Agriculture is a major source of income and employment in India. The most prevalent problem faced by Indian farmers is that they do not select the appropriate crop for their land and do not use the appropriate fertilizer. They will experience a significant drop in production as a result of this. Crop recommendation has been used to solve the farmers' difficulty. Crop recommendation is a modern farming strategy that employs research data on soil ph, soil nutrient contents, weather conditions to recommend the best crop to farmers as well as fertilizer recommendations based on site-specific features. This decreases the number of times a crop is chosen incorrectly and increases productivity. In this paper, this problem is solved by proposing a recommendation system through ML models with majority voting technique using Random Forest, Naive Bayes, Support Vector Machine (SVM), Logistic Regression, Decision tree, XGBoost, as learners to recommend a crop for the site-specific parameters with high accuracy and efficiency. In Addition to that we are performing fertilizer recommendation. The fertilizer recommendation system is purely python logic based. In this we compare the data (optimum nutrients for growing the crop) with the user's entered data, and according to that suggestions will be fetched.

**Keywords:** Decision tree, Random Forest, Support Vector Machine(SVM), LogisticRegression, Naïve Bayes, XGBoost

**1.Introduction**

A farmer's decision about which crop to grow is generally clouded by his intuition and other irrelevant factors like making instant profits, lack of awareness about market demand, overestimating a soil's potential to support a particular crop, and so on [1]. A very misguided decision on the part of the farmer could place a significant strain on his family's financial condition. Perhaps this could be one of the many reasons contributing to the countless suicide cases of farmers that we hear from media on a daily basis [2]. In a country like India, where agriculture and related sector contributes to approximately 20.4 per cent of its Gross Value Added (GVA) , such an erroneous judgment would have negative implications on not just the farmer's family, but the entire economy of a region [3]. For this reason, we have identified a farmer's dilemma about which crop to grow during a particular season, as a very grave one [4]. The need of the hour is to design a system that could provide predictive insights to the Indian farmers, thereby helping them make an informed decision about which crop to grow [5]. With this in mind, we propose a system, an intelligent system that would consider environmental parameters (temperature, rainfall, humidity) and soil characteristics (N, P, K, pH value) before recommending the most suitable crop to the user [6]. In addition to that a fertilizer suggestion is also made which is based on the optimum nutrients of the crops grown [7].

**2.Literature Review**

The following research papers were referred by us before doing our project. While referring each of these papers we have come across various different findings discussed below.

**Title:** Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector

**Author:** Kumar, Y. Jeevan Nagendra, V. Spandana, V. S. Vaishnavi, K. Neha, and V. G. R. R. Devi.

**Method Used:** Random Forest and Decision tree.

**Title:** Crop recommendation system to maximize crop yield using machine learning technique  
**Author:** Rajak, Rohit Kumar, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, and Avinash Devare

**Method Used:** support vector machine (SVM) classifier, ANN classifier, Random Forest and Naïve Bayes

**Title:** Crop recommendation system to maximize crop yield in ramtek region using machine learning

**Author:** Reddy, D. Anantha, Bhagyashri Dadore, and Aarti Watekar

**Method Used:** random forest, CHAID, K- Nearest Neighbour and Naïve Bayes

**Title:** Improving Crop Productivity Through A Crop Recommendation System Using Ensembling Technique

**Author:** Kulkarni, Nidhi H., G. N. Srinivasan, B. M. Sagar, and N. K. Cauvery **Method Used:** Random Forest, Naive Bayes, and Linear SVM

**Title:** Survey of Crop Recommendation Systems

**Author:** Dighe , Deepti, Harshada Joshi, Aishwarya Katkar , Sneha Patil, and Shrikant Kokate

**Method Used:** CHAID, KNN, K-means, Decision Tree, Neural Network, Naïve Bayes, C4.5, LAD, IBK and SVM

### **3 Problem Definition**

In India, agriculture is one of the most important professions. Many of the people do agriculture but are unable to determine which types of crops are more suitable to their soil. Means there are variety of crops which are only suitable for wet soil, some requires medium humidity in the soil to grow but this knowledge is less known to farmers as well as newbies who develop some interest in farming. As of now there are very less resources as well as software's which will help them to improve quality. Such type of software is Crop recommendation system using machine learning.

### **4. Methodology**

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

#### **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. The outliers should be removed.

#### **Feature Selection**

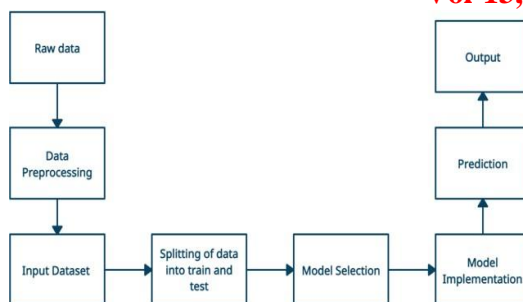
It is important that we select only those features that will be necessary to determine the type of crop to grow. For this, we have created a correlation matrix that shows the linear relationship of a feature with every other features. If features are highly correlated then that feature should be dropped.

#### **Model Building**

The next step is building the machine learning model. While building the machine learning model, first we need to split our dataset into 2 parts i.e.: training data and test data. We have split the data in the ratio of 80-20.

Taking the training data, we apply our machine learning algorithms on the features of the dataset. We have used machine learning algorithms on our training dataset and the algorithms that gives us the highest accuracy will be selected for recommendation.

Architecture diagram of our machine learning model



### Building a UI

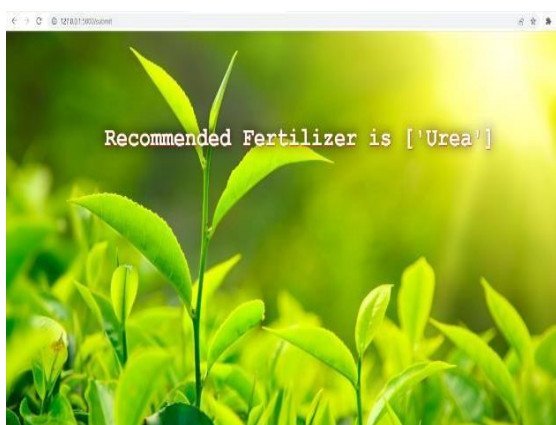
In the next step, we have built a UI for a user to input his data. so that once he enters the information such as N, P, K, ph values of soil, temperature, humidity, rainfall etc., the model will process the data and will recommend the appropriate type of crop to be grown in such a condition and N, P, K, values of soil, temperature, humidity, moisture, soil type, crop type will recommend the appropriate fertilizer to be used. Once the user enters the following values and submits the machine learning model will predict the result. Below is a screenshot of the UI which we have made for crop recommendation.



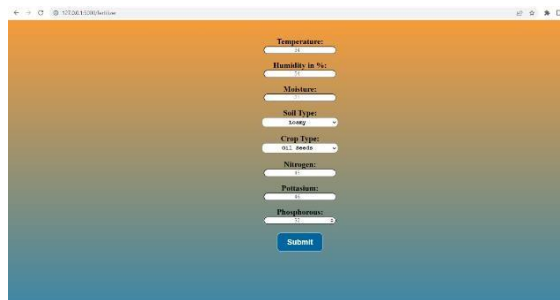
After putting the values by the user, we get the results of the model in the way mentioned below



Below is a screenshot of the UI which we have made for fertilizer recommendation



After putting the values by the user, we get the results of the model in the way mentioned below

A screenshot of a web application interface for crop recommendation. The form is centered on a light blue background with a white border. It contains several input fields: Temperature (text), Humidity in % (text), Moisture (text), Soil Type (dropdown menu), Crop Type (dropdown menu), Nitrogen (text), Potassium (text), and Phosphorus (text). Below these fields is a blue Submit button.

## 5. Dataset

The dataset for this topic was taken from Kaggle.

## 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

Decision Tree is one of the most popular machine learning algorithms used mostly in classification problems but can also be used for regression type of problems. 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.

### K-Nearest-Neighbour:

K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.

### Support Vector Machine:

The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.

### Naive Bayes algorithm:

It is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.

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.

### Logistic Regression

It is one of the simplest algorithms in machine learning. It is used for solving classification problems. 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

## 7. CONCLUSION

Using machine learning algorithms, we are predicting the right crop to grow and fertilizer to be

used if the farmer adopts to this technology which would not only make his life easier but also help him in making decisions that do not exploit the environment.

## **8. Results**

The algorithms we have used are Decision Tree, Random Forest, Logistic Regression, KNN, Support vector machine, Naïve bayes and XGBoost. After implementing the algorithms on our dataset, we can see that XGBoost gives us the highest accuracy out of all the algorithms.

## **REFERENCES**

- [1] 2019, 10th International Conference on Computing, Communication and Networking Technologies, “Low-cost IOT+ML design for smart farming with multiple applications”, Fahad Kamraan. Radhika, Narendiran, “Kind of Crops and Small Plants Prediction using IoT with Machine Learning,” International Journal of Computer & Mathematical Sciences, 2018.
- [2] “Crop Recommendation on Analyzing Soil Using Machine Learning” Anguraj.Ka, Thiyaneswaran.Bb, Megashree.Gc, Preetha Shri.J.Gd, Navya.Se, Jayanthi. Jf, 2020.
- [3] “Classification of Soil and Crop Suggestion using Machine Learning Techniques”, A. Mythili , IEEE 2019.
- [4] Mehta, P., Shah, H., Kori, V., Vikani, V., Shukla, S., & Shenoy, M.,2018. “Survey of unsupervised machine learning algorithms on precision agricultural data”, IEEE
- [5] “IOT based Crop Recommendation, Crop Disease Prediction and Its Solution” Rani Holambe, Pooja Patil, Padmaja Pawar, Saurabh Salunkhe , Mr. Hrushikesh Joshi, 2019 IRJET