Crop Recommendation System using Machine Learning

Aman Sinha[1], Pallavi Sinha[2], Ritika Rajani[3], Dr. Sumithra Devi K.A[4]

[1] B.E Student, Department of Information Science and Engineering, DSATM, Bengaluru, India
[2] B.E Student, Department of Information Science and Engineering, DSATM, Bengaluru, India
[3] B.E Student, Department of Information Science and Engineering, DSATM, Bengaluru, India
[4] Professor & HoD, Department of Information Science and Engineering, DSATM, Bengaluru, India

ABSTRACT

Farming is one of the most basic and widely practised occupations, and it is critical to the country's development. The majority of the land is used for agricultural agriculture in order to meet the needs of the local population as well as export demand. As a result, farmers face a huge problem in expanding crop production. Crop cultivation is influenced by climate (so-called seasons) and soil attributes all over the world; nevertheless, increasing crop production is influenced by a number of elements, the most important of which is temperature. Our system made use of Machine Learning processes with the purpose of recommending the best crops based on the temperature.

Keywords: *Crop suggestion, Fertilizer suggestion, Nitrogen-Phosphorus-Potassium (NPK), Machine Learning (ML)*

I. INTRODUCTION

India is one of the world's oldest countries with a thriving agricultural sector. However, due to globalisation, agricultural trends have radically changed in recent years. The state of agriculture in India has been influenced by a number of variables. Many innovative technologies have emerged to help people regain their health. Precision agriculture is one such technology. Crop suggestion is one of the most important aspects of precision agriculture. It is influenced by a number of factors. Not all precision agriculture systems are made equal. In agriculture, however, it is vital that the advice given is exact and precise, as mistakes can result in severe material and capital loss. Our country's farmers will profit as a result of this. Prior crop and yield predictions were made based on the farmers' previous experience in a given place. They will choose the preceding or neighbouring or trendier crop in the surrounding region purely for the sake of their land, and they have little awareness of soil nutrients like nitrogen, phosphate, and potassium. India's agriculture has advanced significantly in recent years. Precision agriculture relies on site-specific cultivation. Even while precision agriculture has progressed, it still has problems. Precision agriculture plays a significant part in crop recommendation. "The dataset contains various parameters like Nitrogen (N), Phosphorous (P), Potassium (K), PH value of soil, Humidity, Temperature, and Rainfall. This proposed system applied different kinds of Machine Learning algorithms like Support Vector Machine (SVM), Logistic Regression, Random Forest (RF), and XGBoost."

II. LITERATURE SURVEY

"Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique- Rakesh Kumar, M.P. Singh, Prabhat Kumar, and J.P. Singh".[1]

This study proposed the Crop Selection Method (CSM) to address the crop selection problem, maximise the crop's net production rate throughout the season, and achieve maximum economic growth for the country. The proposed strategy has the potential to increase crop net production rates.

"AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms -Zeel Doshi, Subhash Nadkarni, Rashi Agrawal, Prof. Neepa Shah".[2]

This research proposed and implemented an intelligent crop recommendation system that farmers across India can utilise. Based on a range of environmental and geographical parameters, this system would assist farmers in making an informed selection about which crop to cultivate. We've also put in place a backup mechanism called

Rainfall Predictor, which forecasts rainfall for the next 12 months. Different machine learning methods such as Decision Tree, K Nearest Neighbour (K-NN), Random Forest, and Neural Network were implemented in this proposed system, and multi-label classification was done on it. Utilizing the rainfall predictor model, this proposed system achieved 71 percent accuracy, and using a neural network technique to generate an appropriate predictor system, it reached 91.00 percent accuracy.

"Crop recommendation system to maximize crop yield using machine learning technique- Rajak, Rohit Kumar, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, and Avinash Devare". [3]

This suggested technique is based on a soil database and is used to detect certain crops. Farmers will benefit from this research in terms of enhancing agricultural productivity, minimising soil deterioration on cultivated land, lowering chemical use in crop production, and maximising water use efficiency.
Farmers will benefit from this research by enhancing agricultural productivity, minimising soil deterioration in cultivated area, reducing chemical use in crop production, and increasing water use efficiency.

"Improving Crop Productivity Through a Crop Recommendation System Using Ensembling Technique - Kulkarni, Nidhi H., G. N. Srinivasan, B. M. Sagar, and N. K. Cauvery". [4]

This proposed approach is used to accurately recommend the best crop based on the soil type and features such as average rainfall and surface temperature. Random Forest, Naive Bayes, and Linear SVM were among the machine learning algorithms used in this proposed system. The input soil dataset was categorised into two crop types by this crop recommendation system: Kharif and Rabi. Using this recommended technique, we were able to attain a 99.91 percent accuracy rate.

III. PROPOSED SOLUTION ALGORITHM USED:

K-NN Algorithm: -

There are many data mining methodologies, and the k-nearest neighbor approach is one among them which is used to analyze data. A set of points in space are created by treating each of the qualities in our training set as a distinct dimension in some space, with each observation's value for that quality acting as its coordinate. Two points in this space are from one another by a certain amount, which is used to determine how similar the two points are. The algorithm chooses which training set points are comparable enough to be taken into account by choosing the k closest data points to the new observation and choosing the most prevalent class among them for a brand-new observation, one can predict the class. Assume that p and q are feature vectors. The Euclidean metric is commonly employed to estimate the distance between p and q. if a = (a1, a2) and b = (b1,b2) then the distance is given by:

d(a, b) = math.sqrt((*b*1− *a*1)**2 + (*b*2 − *a*2)**2)

Random Forest (RF): -

The Random Forest algorithm is a machine learning technique. During the training phase, a large number of decision trees are created, and the output is separated into two categories: classification and class prediction (regression). The accuracy of the forecast is related to the number of trees used. Rainfall, perception, temperature, and productivity are among the variables included in the dataset. These variables from the dataset are used in the training phase. Only two-thirds of the total data set is taken into account. The remaining data set is used to put the method to the test. The random forest algorithm contains three parameters: n tree, which specifies the number of trees to be formed, and m try, which specifies the number of variables to be taken at each node split. The number of observations we'll need to make at terminal nodes is determined by the node size.

Support Vector Machine (SVM): -

The Support Vector Machine (SVM) is a supervised machine learning approach or model that can be used to solve classification and regression problems. However, it is mostly used to tackle categorisation issues. SVM is usually represented as a set of training data points in space divided into groups by a comprehensible gap that stretches as long as possible. Each data item is represented as a point in n-dimensional space by the SVM technique, with each feature value reflecting the value of a specific coordinate. The categorisation is then completed by determining which hyper-plane best distinguishes the two classes. The Support Vector Machine (SVM) is a supervised machine learning method for classifying binary data. Based on data from soil, temperature, humidity, moisture, subterranean water, and rainfall, SVM is used to classify whether a certain

crop can thrive in that place. This method creates a hyper plane in an N-dimensional space, where N is the number of features in a dataset that will be used to identify data points. The basic goal of the method is to locate the plane with the greatest Margin, or the maximum distance between data points of the features being plotted. The classification becomes more precise as the distance between two spots increases.

XGBoost:-

EXtreme Gradient Boosting (XGBoost) is a more flexible and improved variant of the gradient boosting method that is better for model viability, computational speed, and model performance. Other machine learning approaches have been shown to outperform XGBoost. XGBoost is a fast and accurate parallel tree boosting (also known as GBDT, GBM) technique that solves a number of data science problems.

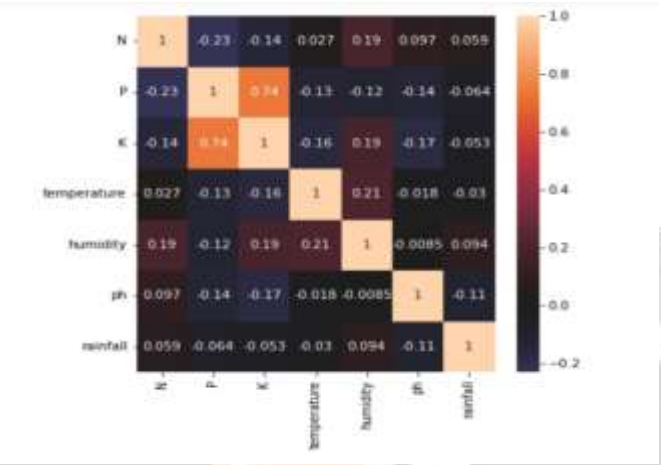

Fig- We can see that different clusters are formed based on the features in the dataset.

IV. FUTURE WORK

The following features could be added to the system to make it even better:

1) The main purpose of future work is to add new attributes to the dataset to enrich it.

2) To create an easy-to-use website and mobile app.

3) We believe that the proposed strategy will help farmers make the best crop-cultivation decisions possible.

4) Based on the system's recommendations, a farmer can plant different crops in different districts. As a result, by utilizing the method, every farmer will have the opportunity to maximize their produce and profit. Even though we are a developing country, our fundamental goal is to create more with less.

5) We're virtually out of resources trying to keep up with the rest of the world's info. Furthermore, any contribution to agriculture can be helpful to both the country and its people.

V. RESULTS

An algorithm is being developed to forecast agricultural yield rates. This application is made up of three parts. The first step is to organise your data. The second step is to test datasets, followed by the third step of data analysis. And, while working with datasets, convert them into a supported format.

	Algorithm	model	Train_score	Test_score	accuracy
0	KNN	knn	89.636364	84.272727	84.272727
1	NaiveBayes	nb	96.363636	94.727273	94.727273
2	LogisticRegression	lr	66.454545	63.909091	63.909091
3	SVM	svm	69.454545	65.181818	65.181818
4	DecisionTreeClassifier	dt	69.454545	65.181818	92.181818
5	BaggingClassifier	bg	99.454545	92.545455	92.545455
6	RandomForestClassifier	rf	100.000000	94.727273	92.545455
7	AdaBoostClassifier	ad	14.363636	12.909091	12.909091
8	GradientBoostingClassifier	gb	95.727273	90.454545	90.454545
9	XGBClassifier	xg	96.363636	91.727273	91.727273
10	lbgmClassifier	lbgm	100.000000	93.454545	93.454545

Fig- Models with a high level of precision (more than 90%).

VI. CONCLUSION

This presents a Machine Learning framework for the task of an intelligent crop recommendation system and fertilizer recommendation system, which farmers all around India can simply use. This approach would assist farmers in making an informed decision on which crop to cultivate based on a variety of factors such as N, P, K, temperature, humidity, ph., and rainfall. We can boost the country's productivity and profit from such a strategy if we use it [5]. In this way, farmers can plant the appropriate crop, enhancing their production and the country's total profitability. "Various machine learning methods such as K-Nearest Neighbour, Random Forest, and XGBoost were used in this study to convey the recommendations of various Indian crops". The analysis was carried out on various types of machine learning algorithms, and all of them will display their best accuracy results.

For the fertilizer recommendation system user need to input the N, P, K, temperature, humidity, ph., rainfall, soil type, and crop type. "Various machine learning methods such as Random Forest, Support Vector Machine (SVM), and XGBoost were used in this study to convey the recommendations for fertilizers".

VII. REFERENCES

[1]"Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique - Rakesh Kumar, M.P. Singh, Prabhat Kumar, and J.P. Singh."

[2]"AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms -Zeel Doshi, Subhash Nadkarni, Rashi Agrawal, Prof. Neepa Shah."

[3]"Crop recommendation system to maximize crop yield using machine learning technique- Rajak, Rohit Kumar, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, and Avinash Devare."

[4]"Improving Crop Productivity Through a Crop Recommendation System Using Ensembling Technique- Kulkarni, Nidhi H., G. N. Srinivasan, B. M. Sagar, and N. K. Cauvery."

[5]"Crop Recommendation Assistant Using Machine Learning – Aman Sinha, Pallavi Sinha, Ritika Rajani, and Dr. Sumithra Devi K.A."