International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 12, Issue 4 [April. 2023] PP: 19-26

AGRO ASSIST

Y.Sowjanya[1], B.Bhavya[2], K.Ruchitha[3], N.Abhijith[4]

[1]*Assistant Professor in Department of IT Anurag Group of Institutions,*
[2,3,4]*UG scholar In Department of IT Anurag Group Of Institutions*

Abstract— *One of the key industries that effect the nation's economic development is agriculture. Farmers must be highly aware of the types of nutrients present in the soil in order to boost quality and productivity. In order to plan or safeguard their crop, farmers must have access to the weather forecast for at least a few days in a row. Many farmers can lack a thorough understanding of the many types of soil and the products that can be cultivated there. Numerous plant diseases and their management strategies remain unknown to farmers. It needs an agricultural officer or expert advisor. We propose crop recommendations for cultivation using machine learning techniques. The Random forest classifier and support vector machine algorithms are used in the crop prediction module to forecast the appropriate crop that the user should grow depending on the criteria provided. Using the Random Forest technique, the suggested model achieves 99% accuracy, and the Support Vector Machine algorithm, 95% accuracy. The crop prediction module employs an algorithm to foretell the suitable crop that the user should cultivate in accordance with the parameters given.*

Keywords— *Agriculture, Machine Learning (ML), Crop Recommendation, Climatic Conditions, Random Forest (RF),Support Vector Machine (SVM).*

Date of Submission: 18-03-2023 Date of acceptance: 03-04-2023

I. INTRODUCTION

For both human existence and the Indian economy, agriculture is a crucial sector. It is among the major careers that are necessary for human existence. Additionally, it makes a significant contribution to our daily lives. Most of the time, farmers commit suicide as a result of production loss because they are unable to repay the bank debts as they have taken out for their businesses. We observed that the environment has been steadily changing right now, which is bad for the crops and pushes farmers into more debt and suicide.

When various mathematical or statistical techniques are used to data, these risks can be reduced [1]. Using these techniques, we can advise the farmer on the best crop to plant on his agricultural area in order to assist him make the more profit possible. In India nowadays, agriculture has greatly advanced.

The machine learning algorithms such as supervised, unsupervised, and reinforcement learning, has both advantages and downsides. In supervised learning, a set of data that includes the intended inputs and outputs is used to build a mathematical model. A mathematical model is built using an unsupervised learning technique from a set of data that just comprises inputs and no specified output labels. In situations when some of the sample input lacks labels, semi- supervised learning algorithms are used to build mathematical models.

[2] Input characteristics such as nitrogen (N), phosphorus (P), potassium (K), soil PH, humidity, temperature, and rainfall are used in this paper to suggest the best crop. The most appropriate crop is suggested based on the accuracy of the future production of eleven different crops, including rice, maize, chickpeas, kidney beans, pigeon peas, moth beans, mungbeans, black gram, lentil, pomegranate, banana, mango, grapes, watermelon, muskmelon, apples, oranges, papayas, coconuts, cotton, jute, and coffee. The dataset includes a number of factors like humidity, temperature, rainfall, nitrogen (N), phosphorus (P), potassium (K), and soil PH. To serve as training data for the machine learning system, the data is gathered and analyzed. The software displays better results with more data.

The machine learning techniques used in this suggested system are Random Forest (RF), Support Vector Machine (SVM).

The following is a list of the paper's main contributions:

I. The programme offers farmers a comprehensive platform where they may access several machine learning techniques through a single online application.

II. When compared to the current systems, the machine learning model accurately predicts the outcomes,

and it continuously adjusts itself to increase accuracy.

III. The major goal of the suggested approach is to enable rapid and effective information retrieval.

IV. The system is very simple to use and also have a user-friendly interface.

II. LITERATURE REVIEW

[3] Globally, economic losses from natural disasters are increasing, and the agriculture sector is particularly susceptible to them. The United Nations Office for Disaster Risk Reduction (UNISDR) (2018) estimates that from 1998 to 2017, disaster-affected nations suffered direct economic losses totaling US$ 2908 billion.

Disasters related to climate made up 77 percent of all losses. In the recent past, the agriculture industry has been more severely affected by climate change. According to the 2018 Economic Survey conducted by the Government of India,the negative consequences of climate change result in an annual loss of $9–10 billion. Different gases that impact the climate and lead to natural disasters are described in figure 1.

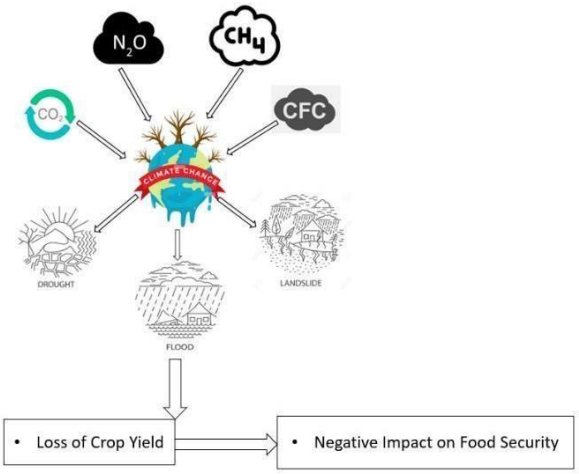

Figure 1. Impact of Climate Change on Agriculture

Since farming is mostly dependent on weather and soil characteristics, a change in climate could affect productivity, leading to a shortage of food. Figure 2 below illustrates how crops in India are impacted by extreme natural occurrences.

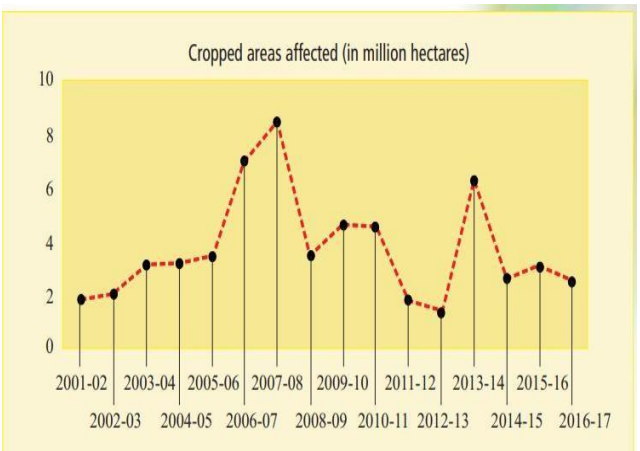

Figure 2. Impacted crops

[4] Examining and obtaining relevant information from data is a procedure known as "data mining." Data mining is used in many industries, including finance, retail, medicine, and agriculture. Data mining is applied for the analysis of the many biotic and abiotic components in agriculture.

[5] The machine learning techniques used by the proposed system included random forest, CHAID, K-

Nearest
Neighbor, and Naive Bayes. By using the suggested system, we can forecast a certain crop under a specific set of weatherconditions, as well as state and district values. In order to boost national productivity, our suggested work would assist farmers in planting the appropriate seed depending on soil requirements.

A crop recommendation system using machine learning is to maximize agricultural productivity [6]. Utilizing soil database, the proposed method is able to identify a certain crop. This system's design was tested with a wide range of crops, including groundnut, legumes, cotton, vegetables, bananas, rice, sorghum, sugarcane, and coriander, as well as with a number of soil properties, including depth, texture, pH, soil color, permeability, drainage, water holding, and erosion[7]. The suggested approach utilized a number of machine learning classifiers, including Nave Bayes, Random Forest, Support Vector Machine (SVM), and ANN, to accurately and efficiently select a crop for a site- specific parameter. This research will aid farmers in improving agricultural productivity, preventing soil erosion on cultivated land, reducing the use of chemicals in crop production, and for proper usage of water resources.

Ensemble technique helps to improve crop productivity through crop recommendation system. This suggested approach is used to accurately select the best crop dependingon the unique type and characteristics of the soil, such as the average amount of rainfall and the surface temperature. The machine learning methods used by this suggested system included Random Forest, Naive Bayes, and Linear SVM. Moreover, the input soil dataset was categorized by this crop recommendation system into two types of crops: Kharif and Rabi. Applying the suggested technique led to a 99.91%accuracy rate.

III. METHODOLOGY

The proposed system's main objective is crop recommendation based on several factors like soil condition, weather condition and moisturizer levels in the air. Thissystem increases the productivity of the crop which in turn helps the farmer to get more profits. To achieve this goal, authors used machine learning algorithms such as Random Forest and Support Vector Machine.

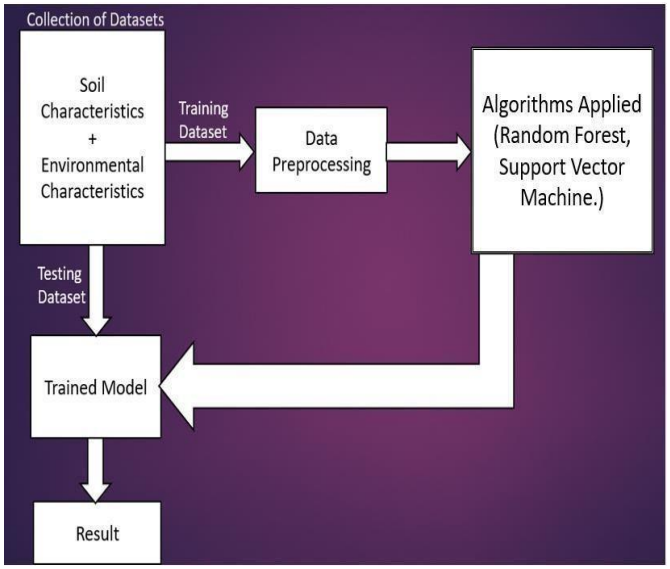

Figure 3. Crop Recommendation System Architecture

Figure 3 describes the architecture of the proposed system which is crop recommendation system and its process flow. Each phase is described in the below sections.

Data Collection: The most widely used technique for gathering and interpreting data from various sources is data collection [8]. Several parameters taken into consideration for advising the crops such as soil pH, humidity, nitrogen, phosphorus, potassium, temperature, and rainfall [9].

Nitrogen	Phosphorous	Potassium	Temperature	PH	Rainfall	Label
68	58	38	23.223	6.336	221.209	Rice
87	35	25	21.445	6.178	65.889	Maize
34	76	80	20.656	7.985	65.238	Chickpea
101	70	48	25.360	6.012	116.553	Banana
32	14	37	22.730	6.825	104.684	Pomegranate

Figure 4: Crop recommendation dataset

Figure 4 represents the dataset[15] which contains one dependent variable called "Label" contains eleven different crops such as rice, maize, kidney beans, pigeon peas, moth beans, chickpea, pomegranate, banana, musk melon, lentil and coffee.

Data Preprocessing: Following the data collection from several sources, dataset preprocessing takes place prior to model training. The data preprocessing begins with reading the obtained dataset and ends with cleaning the data. In order to create crop recommendation datasets, data cleaning involves removing some redundant attributes from the datasets. To increase accuracy, it is necessary to remove pointless characteristics and fill in any missing values in datasets with unfavourable Nan values. The aim of the model follows data cleaning, and the dataset split using the Sklearn software into a train data and a test data.

Algorithms used:

Machine learning approaches determine the likelihood of future outcomes based on historical records. The objective is to provide the most accurate prediction of what will transpire in the future in addition to comprehending what has already occurred. We employed supervised machine learning with the subcategories of classification and regression in our system. For our system, Random forest classifier works well for crop forecasting [10].

A. Random Forest (RF):

Instead of relying solely on a single decision tree, the random forest tends to take the estimate from each tree and bases its prediction of the final output on the majority votes of predictions [11]. Random Forest is a classification technique that contains a classification tree on various subsets of the particular dataset and tends to take the overall score to enhance the forecasting ability of that dataset. Higher accuracy and overfitting are prevented by the larger number of trees in the forest.

First, N decision trees are combined to generate the random forest, and then predictions are made for each tree that was produced during the initial phase.

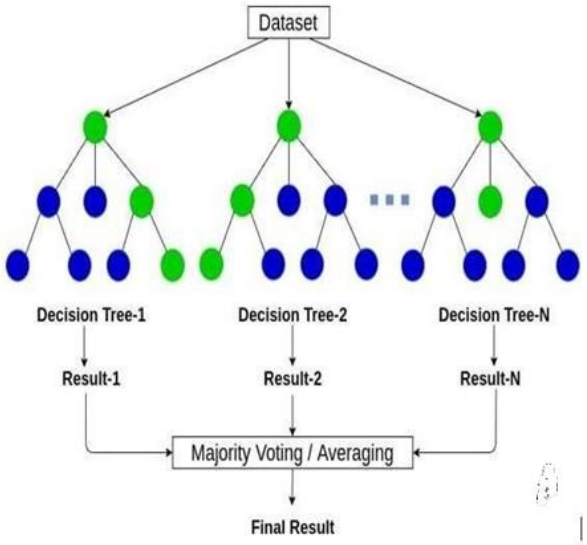

Figure 5: Random Forest Classifier

[12] The following stages can be used to show how theprocess flow works:
Step 1: Randomly select K data points from the training dataset.
Step 2: Create the decision trees connected to the selecteddata points (Subsets).
Step 3: Choose N for the decision tree size you want tocreate.
repeat steps 1 and 2.
Step 4: By searching up each decision tree's predictions for the new data points, add new data points to the category that obtains the most votes as shown in figure 5.
The following code represents how to build the model and to fit the model with training data using random forest.

```
# Random Forest
from sklearn.ensemble import RandomForestClassifier
ran = RandomForestClassifier(n_estimators=20, random_state=0)
ran.fit(Xtrain,Ytrain)
```

B. Support Vector Machine (SVM):

The Support Vector Machine (SVM), one of the most popular supervised learning algorithms, is used to resolve classification and regression problems. The Machine Learning is mostly used to solve classification problems. To swiftly categorize new data points in the future, the SVM approach seeks to create the ideal decision boundary or line that can divide dimensional space into classes. This ideal decision boundary is referred to as a "hyperplane". Figure 6 describes how SVM chooses the extreme points and vectors to build the hyperplane. The Support Vector Machine approach gets its name from the support vectors that are utilized to represent these extreme occurrences [13].

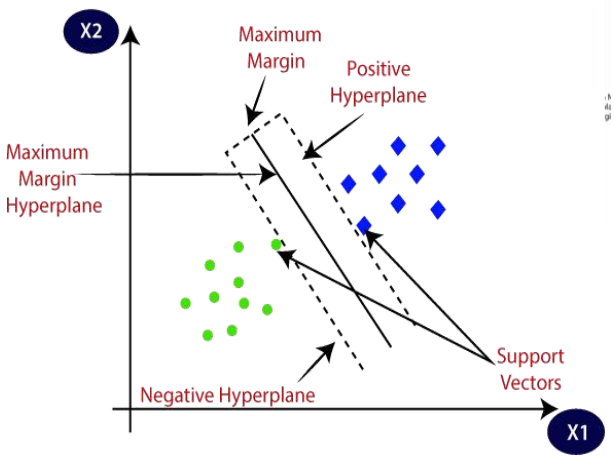

Figure 6. Support Vector Machine

SVM is available in two forms such as linear and non-linear SVM.
Data that can be linearly separable into two classes by a single straight line is subjected to linear SVM classifier. A dataset is deemed non-linear if it cannot be categorized alonga straight line, and the classifier used is referred to as a Non- linear SVM classifier. In SVM, there may be multiple linesor decision boundaries can be used to split classes in n- dimensional space, but we must choose the one that is the ideal boundary known as SVM hyperplane will help us to classify the data points most effectively.
In this, the data points or vectors that are closest to the hyperplane and have the greatest bearing on its location are called support vectors.
The following code represents the SVM approach and fit the model with training data.

```
# Support Vector Machine(SVM)
from sklearn.svm import SVC
SVM = SVC(gamma='auto')
SVM.fit(Xtrain, Ytrain)
```

IV. EXPERIMENTAL RESULT:

All the aspects are taken into account from the standpoint of the farmer and the plant in a modern setting with less space and less agricultural knowledge, the farmer is appropriately guided till the harvest. Before choosing a plant to cultivate, it's critical to be aware of the variables that affect its growth and how to maintain or manage them. In view of all these aspects, proposed system suggests the best suitable crop tobe grown. Moreover, it requires less maintenance and does not require expert advice.

To meet the goal of proposed system, the authors suggested Random Forest Classifier and Support Vector Machine algorithms.

The amount of accurate predictions made by a model in relation to the total number of predictions made is measured by the performance metric known as the accuracy score in machine learning. It can be calculated by using the followingequation.

$$\text{Accuracy} = \frac{number\ of\ correct\ predictions}{total\ number\ of\ predictions}$$

RF algorithm fits a number of decision trees and averages the results, or uses majority voting to increase the predicted accuracy. Random Forest classifier gives an accuracy of nearly 99%.

SVM algorithm is mostly employed in classification issues where the data is sparse. Finding the hyper-plane that effectively distinguishes between the two classes is how we accomplish classification. With the Support Vector Machine,we attained an accuracy of nearly 95%. Figure 7 shows the performance comparison between RF and SVM. Here Random forest gives more accuracy compared to SVM.

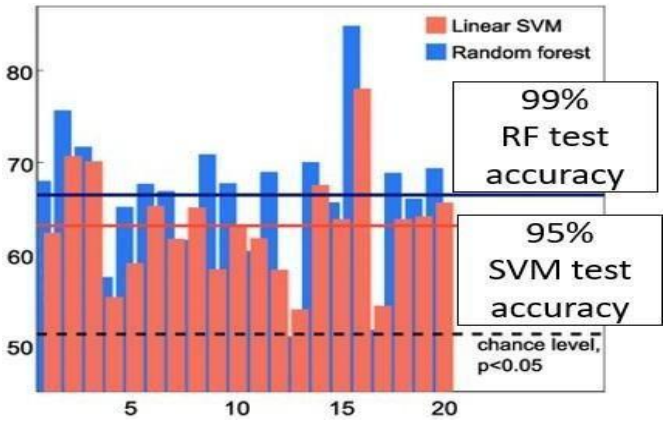

Figure 7. Comparison of Random Forest & SVM

For user convenience, the suggested system is designed as a web application. In order for the online programme to estimate the appropriate crop based on the available parameters, users can submit the necessary parameters and then click the predict button. This programme will be simple to use and controllable from the convenient mobile device aswell [13,14].

We tested the system on a dataset – Crop Recommendation [7,15]. The dataset contains various characteristics, such as soil pH, temperature, humidity, rainfall, and NPK levels. The forecasts based on the aforementioned dataset are tabulated below in figure 8.

S No	Nitrogen	Phosphorous	Potassium	Ph Level	Rainfall (mm)	Predicted Crop
1	8	72	17	5.7	87.8	Kidney Beans
2	38	77	22	6.3	35.3	Lentil
3	9	8	40	6.5	111.6	Pomegranate
4	25	45	78	4	12	Muskmelon
5	32	123	113	101	122	Chick Pea

Figure 8. Results of Various Recommended Crops

V. CONCLUSION

The crop recommendation system based on several factors was proposed and implemented by employing various machine learning algorithms. Random forest and support vector machine approaches were used to suggest the best suitable crop based on the performance metrics. Crop recommendation dataset was used in this approach to analyze different crops to harvest. After preprocessing the data, build the model using random forest algorithm and tested the performance. Similarly, by applying support vector machine technique, it results in 95% accuracy. As random forest gives 99% accuracy, it is suggestible to maximize that crop's yield.

(The suggested machine learning model suggests which crop should be cultivated and what the best ratio for combining various chemicals should be in order to maximize that crop's yield. Choosing the proper distribution for the dataset and concluding with precise classifications. This system suggests Random forest algorithm which gives 99% accuracy compared to SVM classifier.

By utilizing numerous additional Feature Engineering approaches, numerous optimization strategies, and various Machine Learning Algorithms, we may expand upon this methodology. The user can determine which crop is best suited for growing or cultivating based on the current climatic circumstances by employing this notion. They can also determine what adjustments should be made to the chemical makeup of the fertilizer to increase crop yield.)

FUTURE WORK

This GUI crop recommendation could eventually be improved to include fertilizer forecasts, disease forecasts, and crop comparisons. As a further improvement, we may also do market research for crop forecasting in order to plant the most lucrative crop. Relevant numbers of fertilizers will be predicted when the sickness is predicted. Additionally, the crop can be compared to learn more about its chemical make- up and pricing. Future improvements will includecrop comparison, disease detection, and fertilizer prediction.

REFERENCES

[1]. Priyadharshini, A., Swapneel Chakraborty, Aayush Kumar, and Omen Rajendra Pooniwala. "Intelligent crop recommendation system using machine learning." In 2021 5th international conference on computing methodologies and communication (ICCMC), pp. 843- 848. IEEE, 2021.

[2]. Parikh, Dhruv Piyush, Jugal Jain, Tanishq Gupta, and Rishit Hemant Dabhade. "Machine learning based crop recommendation system." *Int J Adv Res Sci Commun Technol* 6, no. 1 (2021): 891-897.

[3]. Ch, Srinivasarao, R. S. Prasad, and T. Mohapatra. "Climate change and Indian agriculture: impacts, copingstrategies, programmes andpolicy." *TechnicalBulletin/PolicyDocument* (2019).

[4]. Pudumalar, S., E. Ramanujam, R. Harine Rajashree,

[5]. 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.

[6]. 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.

[7]. 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.

[8]. Deshmukh, Mangesh, Amitkumar Jaiswar, Omkar Joshi, and Rajashree Shedge. "Farming assistance for soil fertility improvement and crop prediction using XGBoost." In ITM Web of Conferences, vol. 44, p.03022. EDP Sciences, 2022.

[9]. Priyadharshini, A., Swapneel Chakraborty, Aayush Kumar, and Omen Rajendra Pooniwala. "Intelligent crop recommendation system using machine learning." In 2021 5th international conference on computing methodologies and communication (ICCMC), pp. 843- 848. IEEE, 2021.

[10]. Manjula, Aakunuri, and G. Narsimha. "Crop Yield prediction with aid of optimal neural network in spatial data mining: New approaches." *International Journal of Information and Computation Technology,(6)* 1 (2016): 25-33.

[11]. Yadav, Jagdeep, Shalu Chopra, and M. Vijayalakshmi. "Soil analysis and crop fertility prediction using machine learning." *Machine Learning* 8, no. 03 (2021).

[12]. Doshi, Zeel, Subhash Nadkarni, Rashi Agrawal, and Neepa Shah. "AgroConsultant: intelligent crop recommendation system using machine learning algorithms." In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), pp. 1-6. IEEE, 2018.

[13]. 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.

[14]. Onyelowe, Kennedy C., Chilakala B. Mahesh, Bandela Srikanth, Chidobere Nwa-David, Jesuborn Obimba-Wogu, and Jamshid Shakeri. "Support vector machine (SVM) prediction of coefficients of curvature and uniformity of hybrid cement modified unsaturated soil withNQF inclusion." Cleaner Engineering and Technology 5 (2021): 100290.

[15]. Adebiyi, Marion Olubunmi, Roseline Oluwaseun Ogundokun, and Aneoghena Amarachi Abokhai. "Machine learning–based predictive farmland optimization and crop monitoring system." Scientifica 2020 (2020).

[16]. https://www.kaggle.com