IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)
e-ISSN: 2278-1676, p-ISSN: 2320-3331, Volume 18, Issue 3 Ser. I (May. – June. 2023), PP 49-55
www.iosrjournals.org

Smart Crop Recommendation System Using Machine Learning

CH. RAJESH
Department of ECE
SRM institute of science and technology
KTR Campus, INDIA,603203
V.H.S .PRAVEEN
Department of ECE
SRM institute of science and technology
KTR Campus, INDIA,603203
G.SAI DEEPAK
Department of ECE
SRM institute of science and technology
KTR Campus, INDIA,603203
K. HARSHITH
Department of ECE
SRM institute of science and technology
KTR Campus, INDIA,603203
M. VIKRAM SAI REDDY
Department of ECE
S KTR Campus, INDIA,603203
SRM institute of science and technology

Abstract –
Countries' competitiveness and economic growth are fuelled by innovation. Software isused in sustainable agriculture to provide farmers with data and assistance on crop rotation, harvesting details, and soil management. Sensors are used to measure the soil moisture and temperature. This contributes to the development of a model that assists farmers by providing crop-related information or crop recommendations based on various attributes such as crop details, soil composition, weather conditions that crop can grow in, temperature, soil PH, and rainfall. Machine learning algorithms applied are Linear Regression, Random Forest, Decision Tree. The data gathered on experiments by the Indian Chamber of Food and Agriculture from the Kaggle repository .The effects of all three algorithms are measured on a variety of scales, including Accuracy, Precision, Recall, andMeasure.Accuracyiscalculatedbycomparinginstancesthatwerecorrectly and incorrectly predicted. The results show that the Linear Regression algorithm performs well, with the highest precision of 66%when compared to other methods, and it consumes lesser time to build the model[2].

Keywords: *Crop Recommendation, Machine learning, Multiple linear regression, Random Forest, Decision tree.*

Date of Submission: 29-05-2023 Date of Acceptance: 09-06-2023

I. INTRODUCTION

Agriculture is under a lot of stress as a result of today's challenges, such as population growth, climate change, and increased insecurity due to land, water, and energy limits. Agriculture becomes more competitive and sustainable as a result of technology developments, which can help with concerns such as environmentally sustainable practices, waste reduction, and soil optimization. Crop condition is monitored in
modern agricultural production by observing and measuring variables, and crop farmers face a substantial difficulty in controlling all these variables. The rapid expansion of precise monitoring of agricultural productivity and its health evaluation is crucial for the efficient use of farming resources and crop yield management.

Precision agriculture is a sort of farm management that employs information technology to ensure that crops receive exactly what they need to thrive. It is also a modern agricultural approach that uses research data

on soil qualities, soil types, and crop yield data to advise farmers on which crop to plant based on their specific site conditions. This reduces crop selection errors and increases output.

It was also known as satellite agriculture and site-specific crop management. Precision agriculture is strongly reliant on sophisticated machinery, software, and information technology services. The strategy includes access to real-time data on crop, soil, and atmospheric conditions, as well as other related factors such as highly accurate weather forecasts, labour expenses, and equipment availability. Farmers use satellite data to obtain real-time photos of their fields, which are subsequently processed for immediate and long-term decision-making.

It also supports farmers in identifying fields that need to be treated and deciding the quantity of water, fertilizer, and pesticides to apply. This

enables the farmer in eliminating resource waste and preventing run-off, ensuring that the soil has the correct amount of flavoring for maximum health, saving expenses, and regulating the farm's impact on the environment.

A variety of factors are rice, wheat, maize, etc.. influence crop recommendations. Precision agriculture attempts to establish these criteria site by site in order to address crop selection of crops problems. Although "site-specific" procedures enhanced results, the outputs of the systems must still be evaluated. Not all agricultural systems produce reliable outcomes. In agriculture, however, it is vital that the advice provided is correct and precise, as mistakes might result in capital loss. Researchers are doing several investigations in order to produce an accurate and efficient crop forecast model. Crop recommendation systems have piqued the interest of the machine learning community. Different techniques to rule learning are studied in this research.

II. REALTED WORK

Smart agriculture process recommends the use of machine learning approach to know the crop that best suits for the soil to be grown [1]. Accordingly, by considering all the degrees we selected Linear regression algorithm for implementation. It can be easily learnt and understood. Next, A proper dataset with 2200 values with both chemical and environmental conditions containing maximum number of parameters which is collected to predict the result accurately [2]. With some references a linear regression code with suitable import libraries is written, which divides the dataset into two variables and gives the output of the crop that suits the soil and environmental conditions perfectly. Some values from the dataset are given to the machine to learn the data before predicting. Then the system plots the graphs by taking some parameters in x, and others on y axis. By giving the input the system predicts the output. The accuracy value the output is tested and finally the results are taken.

III. PROPOSED METHOD

Linear Regression:

The more precise method. This is a statistical procedure to find the best fit for a set of data points by minimizing the sum of the offsets or residuals of points from the plotted curve. This is the primary technique used in regression analysis

RANDOM FOREST:

These are disassembled learning techniques that construct multiple decision trees and inverted multiple decision trees utilizing a bootstrap database of original data and randomly picking a subset of variables at each decision tree phase. The model then chooses the mode of all projections from each decision tree, and by depending on the majority winds model, the danger of mistake from individual trees is reduced

Dataset Collection:

The Kaggle platform provides data sets to its users, as well as the capability to build models in an environment that is generally web-based and data-oriented. It's essentially a community for machine learning and data science enthusiasts to share data and a platform to showcase their work. It also offers events in which people can compete and refine their machine learning skills while also gaining research ideas**.**[4]

Training Dataset:

From the above diagram training dataset, we have collected data set [3]. In dataset we have input and labelled output. The input data consist of N, P, K, Temperature, Humidity, PH range, Rainfall. Output data is Label (Crop Name).

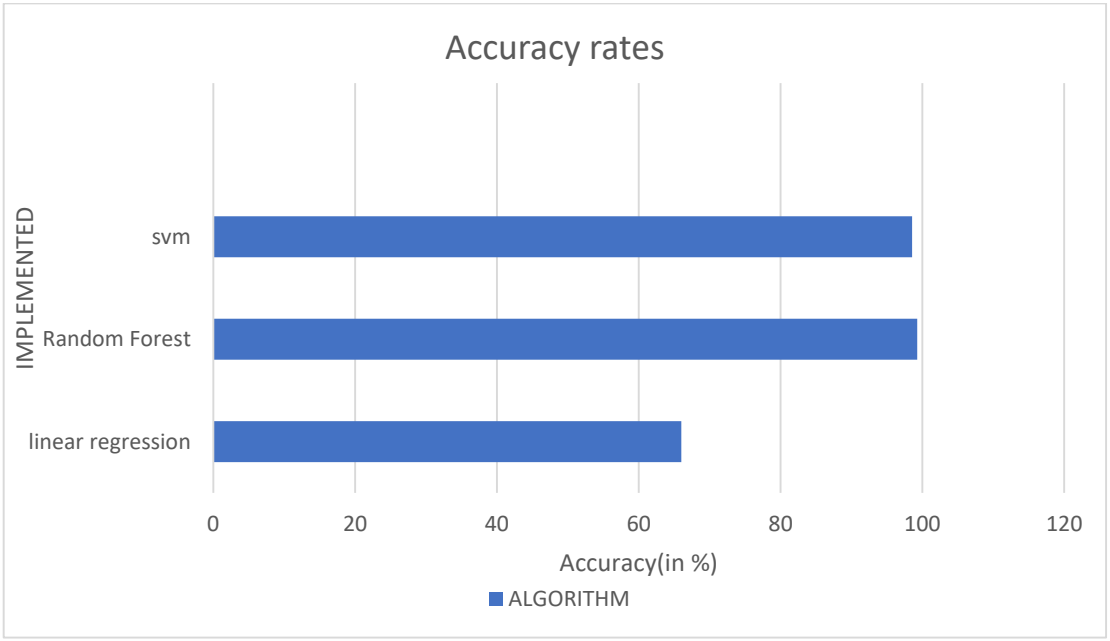
INPUT DATA:

To give inputs to a machine learning model, You have to create a NumPy array, where you have to input the values of the features you used to train your machine learning model. Then we can use that array in the model predict () method, and at the end , it will give the predicted values as an output based on the inputs.

CLASSIFIER:
A classifier in machine learning is an algorithm that automatically orders or categorizes data into one or more of a set of "classes" one of the most common examples is emails to filter them by class label.

PREDICTED OUTPUT:
Prediction refers to the output of an algorithm after it has been trained on a historical dataset and applied to new data. when forecasting the likelihood of a particular outcome**.**

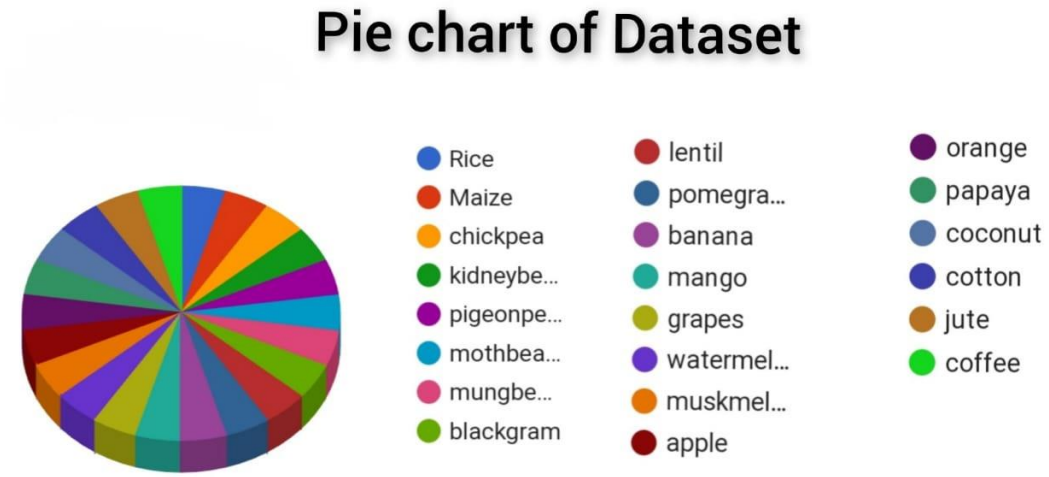
Machine Learning Algorithm:
In pre-harvesting machine learning is used to capture the parameters of soil. Seeds quality, fertilizer application, pruning, genetic and environmental conditions and irrigation, focusing on each component it is important to minimize the overall losses in production**.**



RANDOM FOREST:
Similarly, the greater the number of trees in a random forest algorithm, the higher its accuracy and problem-solving ability, random forest is a classifier that contains serval decision trees on various subsets of given dataset and takes the averages to improve the predictive accuracy of that dataset[5].so that random forest are the best algorithm among them. So, we get 99.27% in random forest[6].

LINEAR REGRESSION:
Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variables values is called the independent variable. In our project we had done the string to float conversion so that we get the 66% accuracy value.

SUPPORT VECTOR MACHINE (SVM):
SVM algorithm is used for classification to classify the different parameters of the soil and predict the most suitable crop.The proposed algorithm is simulated in anaconda navigator to
analyze the soil parameters and recommend a suitable crop. The SVM algorithm is considered for classification.[7]
It gives very good result in terms of accuracy when the data are linearly separable when the data are linearly separable. The SVMs result is a separating hyper plane which maximizes the margin of separation between classes, measured along a line perpendicular to the hyper plane[8]



IV. RESULT AND DISCUSSION:

Finally we got the result accordingly to dataset when we have taken the soil sample for testing purpose . After testing the soil sample which is suitable for that soil**.**

Table 1: DATA SET DESCRIPTION

S.NO	ATTRIBUTE	DESCRIPTION
1	N	Nitrogen
2	P	Phosphorous
3	K	Potassium
4	T	Temperature in Celsius
5	H	Humidity (%)
6	PH	Soil PH
7	RF	Rainfall in mm
8	LABEL	Crop Grown

Data Set Description

The Kaggle Repository was used to get the data for this investigation. This dataset was generated by complementing existing India rainfall, climate, and fertilizer records. These records were compiled by the Indian Chamber of Food and Agriculture (ICFA). This dataset has 2200 occurrences. The dataset has eight characteristics, and the value of one class reflects the crop cultivated under the climatic circumstances. The characteristics of the dataset are listed in Table 1. Table 2 presents the example dataset, which illustrates the relevant values for several variables for a certain crop. The total data set is split into two parts: 60% for training and 40% for testing[3-5].

Table 2 : sample instances in the dataset

CROP	N	P	K	T	H	PH	RF
Rice	**90**	**42**	**43**	**20.8**	**82.0**	**6.5**	**202.9**
Maize	**79**	**51**	**16**	**25.3**	**68.4**	**6.5**	**96.4**
Chickpea	**43**	**79**	**79**	**19.4**	**18.9**	**7.8**	**80.2**
Kindney Beans	**17**	**77**	**23**	**24.5**	**20.8**	**5.6**	**64.1**
Piegeon Peas	**28**	**59**	**22**	**30.9**	**52.7**	**7.0**	**170.9**

The number of instances of each crop type in the data set is shown in table.2 . The dataset contains compared to other classifiers,and SVM, which correctly classified 98% of incidents which is highest and best of other three ..algorithms the details about various parameters of twenty-two crops.

Experimental Results

The performance of the three algorithms is compared interms of model building time is shown in table 3 and thesame is pictorially represented in figure 3. It is inferred thatthe PART algorithm consumes lesser time compared to other algorithms.

TABLE3.MODEL BUILDING TIME

ALGORITHM	Model Building Time(sec)
Linear Regression	0.25
Decision	0.42
Random Forest	0.06

Table.3 Time taken for model training

The confusion matrix and the prediction accuracy of all the three algorithms were tabled in Table4[10]. The confusion matrix shows the number of incidences correctly classified and wrongly classified out of the total number of incidences used .From the table, Random Forest classifier correctly classified 99.27% of incidents, followed by Decision Table ,which correctly classified 98.54% of incidents compared to other classifiers, and SVM , which correctly classified 98% of the incidents which is highest and best of other three algorithms. [11]

Table 4 PERFORMANCE OF THE ALGORITHMS AND CONFUSION MATRIX

ALGORITHM	Evaluated instance		Predication on Accuracy (%)
	Correct Instance	Incorrect Instance	
Linear Regression	633	27	66
Decision	577	83	98.54
Random Forest	649	11	99.27

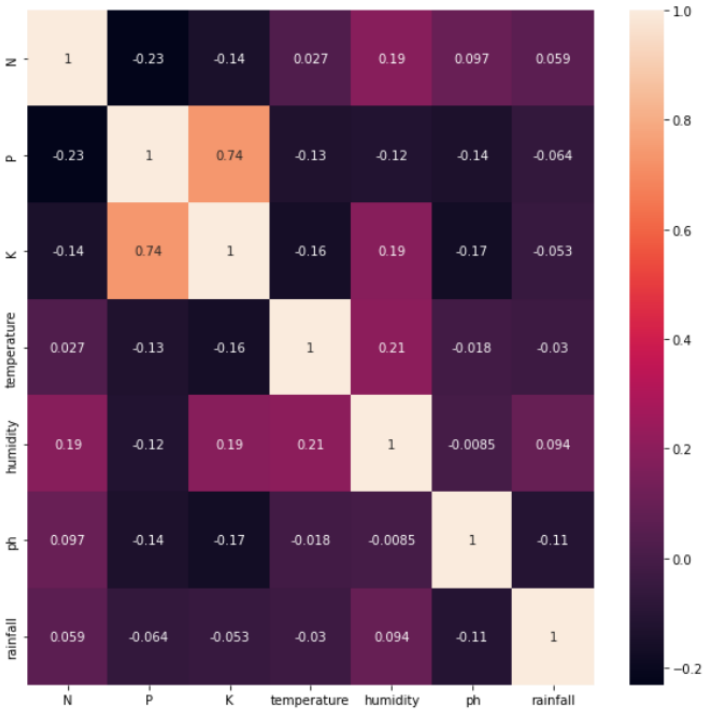


table.4 depicts the prediction accuracy, which is defined as the percentage properly predicted instances over the full testing set and the result shows the outcome of different algorithms for the crop recommendation.

Table5: Performance of the classifiers

Furthermore, the algorithms' accuracy was compared usingthe TP rate, FP rate, Precision, Recall, and F-measure, as shown in Table 5.

Algorithm	**TP Rate**	**FP Rate**	**Precision**	**Recall**	**F-Measure**
Linear Regression	0.962	0.002	0.963	0.962	0.961
Decision	0.886	0.005	0.938	0.886	0.901
Random Forest	0.958	0.000	1.000	0.958	0.991

From the results it is evident that the Random Forest showing its superiority by giving the best prediction accuracy of 99.13% for crop recommendation

V. CONCLUSION

This paper examined the application and performance of various rule-based algorithms for crop recommendation. Thefocus is to direct farmers on the best crops to grow depending on a variety of factors and to assist them in making an informed decision prior to planting. The various performance matrices are calculated such as Precision, Accuracy, F-Measure, and Recall. Our analysis shows that Random Forest and Decision Table classified the correct incidents, with 99 percent and 98 percent, respectively, while Linear Regression classified the most correct incidents, with 66 percent. The result of this study illustrates that the PART algorithm out performs compared to other methods. It is also the most precise classifier. This research will assist farmers in developing a clear picture of how to cultivate a specific crop, by considering the weather conditions, soil, and water requirements.

The Linear Regression algorithm out perform so the, according on the findings of this study. It is also the most accurate classifier out there.

VI. FUTURE ENHANCEMENTS

Acceptance and use of climate statistics primarily grounded completely ranch advisories is in all liability to arise gradationally. growers want time to strive out new statistics, enjoy the benefits, and take delivery of the results. Technology is converting hastily while the station of the growers variations sluggishly. perfecting pack of practices for foremost foliage conserving in view the climate touchy crop categories and climate touchy ranch operations for dwindling cost of civilization and enhancing yield and growing internet returns. The effect exploration must be replicated in different foliage of the region.

References:

[1]. "Machine Learning in Agriculture: Applications and Techniques" in sciforce blog for 2019.
[2]. P.Priya, U.Muthaiah. Balamurugan.Predicting the yield of the crop using a machine learning algorithm. International Journal of Engineering Science Research Technology.
[3]. https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset
[4]. R.Kavitha, M.Kavitha, R.Srinivasan Crop Recommendation in Precision Agriculture using Supervised Learning Algorithm.
[5]. SHILPA MANGESH PANDE, DR. PREM KUMAR RAMESH, ANMOL, B.R AISHWARYA, KARUNA ROHILLA, KUMAR SHAURYA, "Crop Recommender System Using Machine Learning Approach", IEEE Conference 2021.
[6]. S Umamaheswari, S Sreeram, N Kritika and DJ Prasanth, "BIoT: Blockchain-based IoT for Agriculture", *11th International Conference on Advanced Computing (ICoAC)*, pp. 324-327, 2019 Dec 18.
[7]. Crop Recommendation Using Machine Learning Algorithm Daneshwari Modi; Ashok V. Sutagundar; Vijayalaxmi Yalavigi; Anupama Aravatagimath.
[8]. Crop Recommendation System using Machine Learning Aman Sinha, Pallavi Sinha, Ritika Rajani, Dr. Sumithra Devi K.A
[9]. Machine Learning Approaches for Crop Recommendation P.Parameswari, N.Rajathi, K.J.Harshanaa.
[10]. Crop Recommendation System Using Machine Learning Algorithms Atharva Jadhav, Nihar Riswadkar, Pranay Jadha ,Yash Gogawale4
[11]. Keerthan Kumar T G, Shubha C, Sushma S A" Random Forest Algorithm for Soil Fertility Prediction and Grading using Machine Learning "International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, November 2019.