IRJMETS
e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:08/August-2022 Impact Factor- 6.752 www.irjmets.com

PROJECTION OF CAPITULATE VALUES BY MACHINE LEARNING FOR COFFEE

Mohith GM*1, KR Sumana*2

*1PG Student, The National Institute Of Engineering, Mysuru, Karnataka, India.

*2Assistant Professor, The National Institute Of Engineering, Mysuru, Karnataka, India.

ABSTRACT

Over the past 20 years, climate change has had a negative effect on the performance of the bulk of India's agricultural products. In order to take the proper actions for marketing and storage, policymakers and farmers would benefit from knowing the yield of the coffee crop before it is harvested. Before cultivating on the agricultural field, this project will assist the planters in learning the yield of their crop, enabling them to make the best choices. In our study, we took into account the soil data for estimating the coffee yield and the area of production as elements that contribute to coffee production. The Project uses RNN (Recurrent Neural Network), Linear Regression, and Support Vector Machine (SVM) to predict a crop's appropriateness for a specific soil type and improve the overall quality of agricultural production.

Keywords: Machine Learning, Support Vector Machine (SVM), Linear Regression, Recurrent Neural Network (RNN).

I. INTRODUCTION

Coffee is a popular beverage around the world. Researchers have looked into coffee's many health benefits, including its ability to increase energy levels, support healthy weight management, enhance athletic performance, and prevent chronic disease. Around the world, 20 to 25 million families make their living by growing coffee. However, it is still difficult for agricultural system modellers to make an accurate prediction of coffee yield based on environmental, meteorological, and soil fertility variables. The management of resources to provide for basic needs has always been a challenge since the inception and evolution of humans on the planet. According to the Upanishads, Vedas, and classical Hindu literature, soil, water, and air are the three basic resources, and it is our primary duty to manage them wisely and maintain their cleanliness. So we are using Machine Learning Techniques to predict the coffee yield using soil information.

Arabica and Robusta are two varieties of coffee plants that we used in our investigation, we have collected the soil information of both types of plant and the soil information contains the following, age of plant, Phosphorus, potassium ,organic carbon, pH and yield of particular year respectively.

II. LITERATURE SURVEY

This paper discusses [1] and highlights the advantages of using deep learning to anticipate crop yields, the best remote sensing method for obtaining the necessary data, and the various aspects that influence crop yield prediction. Long Short-Term Memory (LSTM) and Convolutional Neural Networks are the most often used deep learning approaches for forecasting crop productivity, according to this study (CNN).

This paper [2].explains how they used machine learning Techniques predict crop yield of different crops like cotton and sugar cane using Linear regression, KNN and decision tree algorithm.

This paper[3] describes the machine learning methods that were utilised to analyse the acquired dataincorporating Naive Bayes, Multinomial Logistic Regression, Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), and Random Forest. To evaluate the efficacy of each technique, five parameters—accuracy, recall, precision, specificity, and f-score—were assessed. In this article [4] They have employed a number of indicators in this dataset, including temperature, rainfall, pH level, relative humidity, and area. As a machine learning method, SVM is used, while as deep learning algorithms, LSTM and RNN are used.

III. DATASET COLLECTION AND PRE-PROCESSING

A. Dataset acquisition

Collecting of datasets for predicting the yield of Coffee yield is very necessary, collecting datasets in very large amount leads to better prediction of Coffee yield, so we have collected datasets from agriculture officer. The datasets that we use in our project are given by the agriculture officer on request; he helped us in getting the datasets from the nearest agriculture office in Kodagu. We have collected datasets about of the soil information where the plant is planted, and yield that was obtained in the following years from 2008 to 2020. Collecting the soil information was necessary than collecting the images of plant and leafs. Soil plays a major role for the better yield of coffee plant, so we have collected the soil information. The dataset that are collected has total of 20074 rows and each row contains total of 9 columns namely, Coffee type-coffee plant is Arabica or Robusta plant, Age- age of the coffee plant, Organic carbon- the percentage of organic carbon in the soil, phosphorus- the amount of phosphorus content in the given same soil, potassium- the amount of potassium in the soil, pH – the amount pH in the soil, Year- the year in which the soil information is collected, place- the places where the soil is collected that was from different taluk wise data, Yield- the yield of the coffee plant in the particular soil.

B. Pre-processing Datasets

Collection of datasets from various source and pre-processing the dataset is very important in machine learning, we here in this section explain how we collected datasets and how we pre-process the datasets collected.

The Pre-processing of is datasets in needed, there is need of converting the raw data in terms that can be suitable for machine learning model. The collected datasets of soil information of coffee plant may contain noises, missing value and formats that can unused by the machine learning model. By doing the pre-processing the datasets filters only required format so we are making the model more accuracy and which leads to better prediction.

1) Getting Datasets: The first step in pre-processing is getting the datasets, for our project we have already collected the datasets. We have soil information of coffee plant.

2) Importing Libraries: For pre-processing of datasets we have to import some of the pre-defined library needed for pre-processing. These are very much helps in many tasks, the libraries that we use in our project is Numpy, Matplotlib and pandas.

3) Handle missing data in the datasets: Machine learning model may have a lot of trouble if our dataset has some missing data. As a result, the dataset needs to manage missing values.

4) Splitting the data set into training set and test set: We divide our data set into a training set and a test set for machine learning data pre-processing. One of the most important data pre-treatment processes, this one enables us to boost the functionality of our machine learning model.

5) Feature scaling: The final stage of machine learning's data pre-processing is feature scaling. It is a method for standardizing the data set's independent variables inside a specified range. With feature scaling, we equalize the range and scale of all of our variables to prevent any one variable from outweighing the others.

IV. METHODOLOGY

The following processes are included in the proposed system for predicting agricultural yield utilising the linear regression approach, Support Vector Machine (SVM), and Recurrent Neural Network (RNN).

A. Linear Regression

Step1: Import Pre-Defined Library

Step 2: Pre-processing the Datasets

Step3: Use Linear Regression Model

B. Support Vector Machine (SVM)

Step1: Import Pre-Defined Library

Step 2: Pre-processing the Datasets

Step 3: Use Support Vector Regression Model

C. Recurrent Neural Network (RNN)

Step 1: Import Pre-Defined Library

Step 2: Pre-processing the Datasets

Step 3: Define a neural network in Keras in the form of sequence of layers.

Step 4: Compile the network that requires various specified parameter.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:08/August-2022 Impact Factor- 6.752 www.irjmets.com

Step 5: Fit the network that requires the specified training data, both an input patterns matrix X and matching output patterns array y.

Step 6: Evaluate the network on the training data. Evaluation of the model on a test or validation set is done rarely.

Step 7: Make required predictions which can be achieved in a format given by the network's output layer.

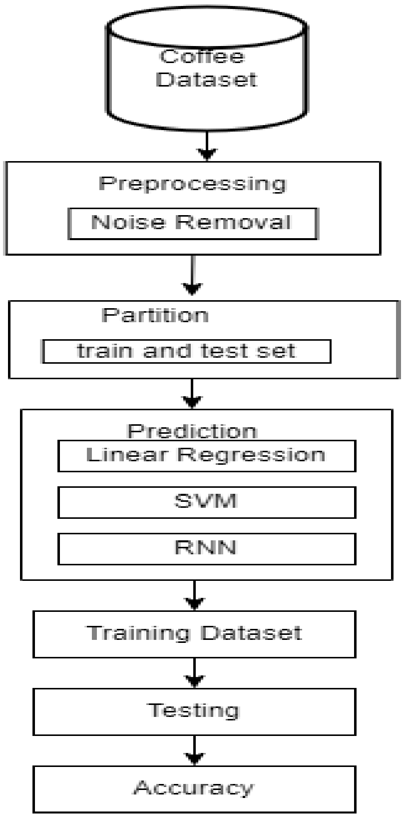

Fig. 1: Coffee Yield prediction Model

V. PERFORMANCE AND METRICS

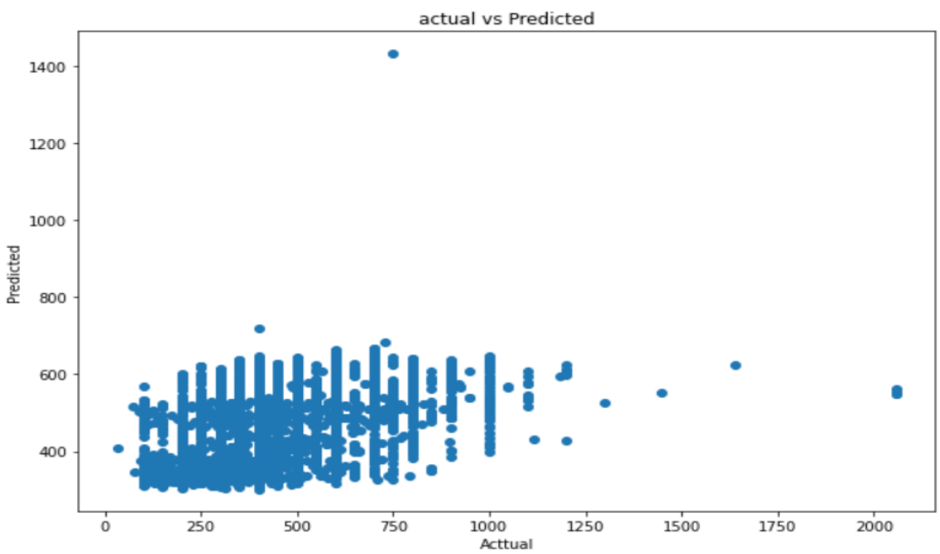

Fig. 2: Actual vs Predicted in Linear Regression

In the above figure 2 shows how the actual vs predicted values and their ranges in Linear Regression, as we can see the actual values in x axis ranges from to 0 to 2000, and y axis ranges from 0 to 1400, as we can see most of the data points fall in same place in linear regression.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:08/August-2022 Impact Factor- 6.752 www.irjmets.com

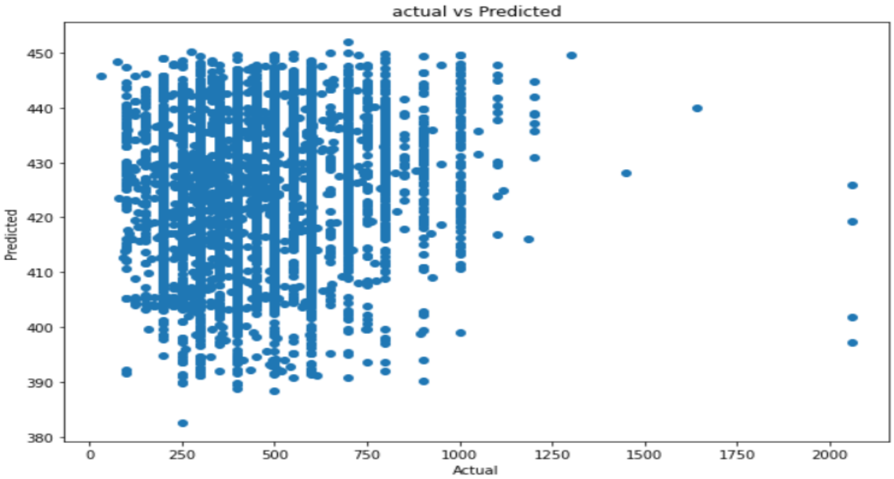

Fig. 3: Actual vs Predicted in Support Vector Machine

The above figure 3 shows how the predicted values and actual values in support vector regression. In x axis there is actual values, in y axis there is predicted values, x axis ranges from 0 to 2000 and y ranges 380 to 450, this are data points that is the predicted yield.

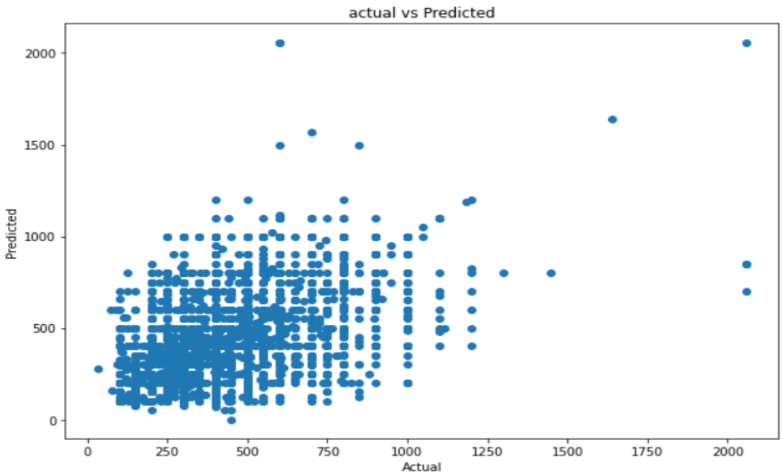

Fig. 4: Actual vs Predicted in RNN

The above figure 4 shows the predicted and actual value in recurrent neural network how the predicted values and actual values in support vector regression. In x axis there is actual values, in y axis there is predicted values, x axis ranges from 100 to 900 and y ranges 200 to 550, these are data points plotted in x and y axis.

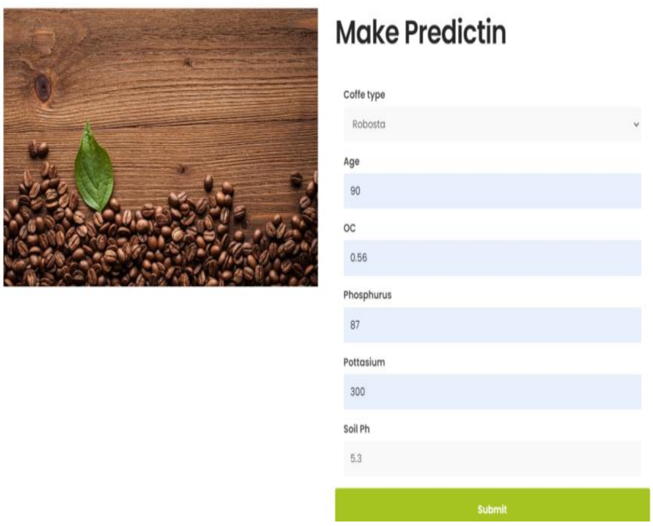

Fig. 5: Home Page

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:08/August-2022 Impact Factor- 6.752 www.irjmets.com

Fig. 6: Prediction page

The figure 5 shows 6 boxes that we have to select the coffee type Robusta or Arabica, second is a text boxes are the age of the plant, organic carbon, phosphorus, potassium, soil ph, after entering all this text box, then we select submit button. The figure 6 shows the prediction of the yield of each coffee plant and each soil.

VI. CONCLUSION

This project is a software tool for the agriculture sector that aids farmers in forecasting coffee yields using data from prior years. By entering their plot's historical data, farmers can check their expected coffee yield. Automation is effective and economically faster for predicting coffee yield. Applying regression algorithms for predicting coffee yield by using linear regression, support vector regression (SVR), and RNN regression algorithms for providing algorithm comparison results enabled the task to be completed effectively. Data mining technology encompasses the classification method. These algorithms use the prior datasets as an input to forecast the yield of coffee using those prior datasets.

VII. REFERENCES

[1] A Systematic Literature Review on Crop Yield Prediction with Deep Learning and Remote Sensing Priyanga Muruganantham , Santoso Wibowo , Srimannarayana Grandhi, Nahidul Hoque Samrat and Nahina Islam. Remote Sens. 2022, 14, 1990. https://doi.org/ 10.3390/rs14091990

[2] Crop Yield Prediction Using Machine Learning Techniques Ashwini I. Patil, Ramesh A. Medar, Vinod Desai, KLS Gogte Institute of Technology, Belagavi, India 2Angadi Institute of Technology and Management, Belagavi, India, DOI : https://doi.org/10.32628/IJSRSET20736

[3] Performance Evaluation of Machine Learning Techniques for Mustard Crop Yield Prediction from Soil Analysis, Vaishali Pandith, Haneet Kour, Surjeet Singh Department of Computer Science , Jatinder Manhas, and ce and IT, University of Jammu.

[4] A Hybrid Approach For Crop Yield Prediction Using Machine Learning And Deep Learning Algorithms, Sonal Agarwal1 and Sandhya Tarar2 School of ICT, Gautam Buddha University, Greater Noida, India, doi:10.1088/1742-6596/1714/1/012012