ISSN 2277-2685
IJESR/April-June. 2025/ Vol-15/Issue-2s/586-591
Andraju Jyothi *et. al.*, / International Journal of Engineering & Science Research

Crop Prediction using Machine Learning

Andraju Jyothi
PG scholar, Department of MCA, DNR College, Bhimavaram, Andhra Pradesh.
K.RAMBABU
(Assistant Professor), Master of Computer Applications, DNR college, Bhimavaram, Andhra Pradesh.

Abstract: ***Crop prediction plays a crucial role in modern agriculture by helping farmers make informed decisions about what crops to plant, ensuring optimal yields, and reducing resource wastage. This study explores the application of machine learning algorithms, specifically Random Forest, Decision Tree, and Passive-Aggressive algorithms, for predicting the best-suited crop based on various environmental and soil parameters. The input features considered for prediction include temperature, humidity, pH, rainfall, and soil nutrients (Nitrogen, Phosphorus, Potassium), while the output is the recommended crop name. A dataset consisting of these parameters was used to train and evaluate the models. The performance of each algorithm was compared based on their accuracy in correctly predicting the appropriate crop. Results indicate that machine learning models, especially Random Forest, show promising results in crop prediction by effectively utilizing environmental and soil data to provide accurate recommendations. This approach offers a scalable solution for precision agriculture, helping farmers optimize crop selection, improve productivity, and manage resources more efficiently.***

Keywords: ***Crop Prediction, Machine Learning, Rainfall, temperature, Humidity.***

I. INTRODUCTION

Agriculture is one of the important occupations practiced in India. It is the broadest economic sector and plays a most important role in the overall development of the country. More than 60% of the land in the country is used for agriculture in order to suffice the needs of 1.3 billion people Thus adopting new agriculture technologies is very important. This will be leads the farmers of our country towards profit [1].

Prior crop prediction and yield prediction was performed on the basis of farmers experience on a particular location. They will prefer the prior or neighborhood or more trend crop in the surrounding region only for their land and they don't have enough of knowledge about soil nutrients content such as nitrogen, phosphorus, potassium in the land.

Being this as the current situation without the rotation of the crop and apply an inadequate amount of nutrients to soil it leads to reduce in the yield and soil pollution (soil acidification) and damages the top layer. Considering all these problems takes into the account we designed the system using a machine learning for betterment of the farmer. Machine learning (ML) is a game changer for agriculture sector.

Machine learning is the part of artificial intelligence, has emerged together with bigdata technologies and high-performance computing to create new opportunities for data intensive science in the multi-disciplinary agrotechnology domain. In the Agriculture field machine learning for instance is not a mysterious trick or magic, it is a set of well define model that collect specific data and apply specific algorithms to achieve expected results [7]. The designed system will recommend the most suitable crop for particular land.

Based on weather parameter and soil content such as Rainfall, Temperature, Humidity and pH. They are collected from V C Farm Mandya, Government website and weather department. The system takes the required input from the farmers or sensors such as Temperature, Humidity and pH. This all-inputs data applies to machine learning predictive algorithms like Support Vector Machine (SVM) [5] and Decision tree [6] to identify the pattern among data and then process it as per input conditions.

The system recommends the crop for the farmer and also recommends the amount of nutrients to be add for the predicted crop. The system has some other specification like displaying approximated yield in q/acre, required seed for cultivation in kg/acre and the market price of the crop.

Ashwani kumar Kushwaha [2] describes crop yield prediction methods and a suggest suitable crop so that it will improve the profit for the farmer and quality of the agriculture sector. In this paper for crop yield prediction they obtain large volume data, it's been called as big data (soil and weather data) using Hadoop platform and agro algorithm. Hence based repository data will predict the suitability crop for particular condition and improvement crop quality.

Girish L [3] describe the crop yield and rain fall prediction using a machine learning method. In this paper they gone through a different machine learning approaches for the prediction of rainfall and crop yield and also mention the efficiency of a different machine learning algorithm like liner regression, SVM, KNN method and decision tree. In that algorithm they conclude that SVM have the highest efficiency for rainfall prediction.

Rahul katarya [4] describes the different machine learning methods used for accelerating crop yield. In this paper they gone through different artificial intelligence techniques such as machine learning algorithm, big data analysis for precision agriculture. They explain about crop recommender system using KNN, Ensemble-based Models, Neural networks, ...etc.

II. LITEARTURE SURVEY

[1] Niketa Gandhi et al., 2016

This study investigates the use of **data mining and visualization** for forecasting **rice crop yield** in tropical wet and dry zones of India. It employs classification algorithms like **J48, LADTree, IBk, and LWL** in WEKA, analyzing the relationship between climatic variables and yield. Results indicate that **J48 and LADTree** performed best in terms of accuracy, sensitivity, and specificity, while **LWL** was the least effective. Interestingly, lower precipitation and higher temperatures were found to correlate positively with yield in the selected climatic zones, showing how environmental trends influence agricultural outcomes [1].

[2] K.E. Eswari & L. Vinitha

This paper uses **Bayesian Network Classification** to predict crop yields in Tamil Nadu, focusing on crops such as rice, coconut, and black pepper. The model evaluates **temperature and rainfall** influences while calculating MAE, RMSE, and related metrics. It highlights how ICT-based predictive models can aid farmers in making better cropping decisions, especially under varying climate conditions. The study affirms the critical role of **timely information and intelligent classification systems** in enhancing agricultural productivity and sustainability [2].

[3] Shruti Mishra et al., 2018

The authors utilize **Random Forest algorithms** for rapid crop yield prediction, emphasizing India's agricultural dependency and the challenges posed by climate variability. The study asserts that **data mining techniques** help identify adaptable crop choices based on local soil and weather data, which can support strategic decisions on storage and marketing. By analyzing these parameters, the research demonstrates that **machine-driven insights** are vital in forecasting yield and reducing economic losses due to poor crop selection or climate fluctuations [3].

[4] Anna Chlingaryana et al., 2018

This review explores **machine learning approaches** for estimating crop yield and **nitrogen status** using **remote sensing (RS)**. The study spans 15 years of research, highlighting ML's strength in handling large, non-linear datasets for agricultural analysis. It notes the increasing integration of **multi-sensor platforms** and hybrid models in **precision agriculture (PA)**, aiming to reduce costs, improve efficiency, and support decision-making. ML models like CNNs and decision trees have shown potential in nitrogen optimization, yield mapping, and crop health monitoring [4].

[5] Dakshayini Patil et al., 2017

This work provides an overview of **rice yield prediction using data mining**, with a focus on soil and weather parameters like **pH, EC, and NPK values**. The study applies the **K-Nearest Neighbor (KNN)** algorithm and finds it effective for forecasting yields in Karnataka. The approach aims to minimize crop losses and improve storage planning, especially in the face of unpredictable weather. The authors advocate that integrating

data-driven insights into farming can significantly help farmers in decision-making and increase productivity [5].

[6] Prof. D.S. Zingade et al.

This systematic review analyzes **50 machine learning-based** and **30 deep learning-based** studies on crop yield prediction. It identifies **temperature, rainfall, and soil type** as dominant features and highlights **Artificial Neural Networks (ANN)** and **Convolutional Neural Networks (CNN)** as the most used algorithms. The paper emphasizes the complex nature of crop yield modeling, noting that better data, feature engineering, and **hybrid ML approaches** are key to enhancing prediction accuracy. It reflects the growing consensus on ML's transformative impact in agriculture [6].

[7] Ashwani Kumar Kushwaha & Sweta Bhattacharya

This study introduces a novel **"Agro algorithm"** for yield prediction using **big data analytics on the Hadoop platform**. It focuses on soil suitability and weather patterns to suggest the best crops for a given region, aiming to enhance yield and economic returns. The paper underlines the role of **Hadoop in processing large agricultural datasets** and supports precision agriculture through climate-aware and soil-specific recommendations. The model integrates environmental and disease data, proposing a scalable solution for improving agricultural quality and sustainability [7].

III. PROPOSED METHOD

The Proposed system will predict the most suitable crop for particular land based on soil contents and weather parameters such as Temperature, Humidity, soil PH and Rainfall.

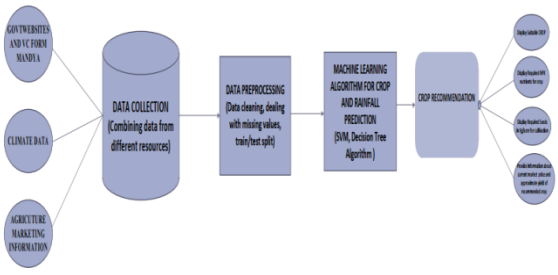

Figure 1. Architecture of the proposed system.

The Architecture of the proposed system consists of various blocks as shown in the fig (1) as follows

3.1. Data Collection: -

Data collection is the most efficient method for collecting and measure the data from different resources like govt websites, VC Form Mandya, APMC website .... etc. To get an approximate dataset for the system. This dataset must contain the following attributes i)Soil PH ii) Temperature iii) Humidity v) Crop data vi) NPK values, those parameters will consider for crop prediction.

3.2. Data Preprocessing: -

After collecting datasets from various resources. Dataset must be preprocessing before training to the model. The data preprocessing can be done by various stages, begins with reading the collected dataset the process continues to data cleaning. In data cleaning the datasets contain some redundant attributes, those attributes are not considering for crop prediction.

So, we have to drop unwanted attributes and datasets containing some missing values we need to drop these missing values or fill with unwanted nan values in order to get better accuracy. Then define the target for a model. After data cleaning the dataset will be split into training and test set by using sklearn library.

3.3. Machine Learning Algorithm for Prediction: -

Machine learning predictive algorithms has highly optimized estimation has to be likely outcome based on trained data. Predictive analytics is the use of data, statistical algorithms and machine learning techniques to identify the likelihood of future outcomes based on historical data.

The goal is to go beyond knowing what has happened to providing a best assessment of what will happen in the future. In our system we used supervised machine learning algorithm having subcategories as classification and regression. Classification algorithm will be most suitable for our system.

➢ **Crop prediction: -** Decision tree algorithm, Random Forest Algorithm, Passive Aggressive Algorithm.

A. Crop Prediction:

Crop prediction process being with the loading the external crop datasets. Once the dataset read then pre-processing will be done by various stages as discussed in Data Pre-processing section. After the data pre-processing, train the models using Decision tree classifier into training set. For a prediction of the crop, we consider a various factor such as temperature, humidity, soil PH and predicted crop.

Those are the input parameter for a system that can be entered by manually or taken from the sensors. Predicted rainfall and input parameter values will be appended in a list. The Decision tree algorithm will predict the crop based on list data.

3.4. Crop Recommendation:

Based on predicted soil contents and weather parameters the system will recommend the most suitable crop for cultivation. This system also provides details about required fertilizers like Nitrogen(N), Phosphorus (P) and potassium(K) in Kg per hectare and display the required seed for a cultivation in Kg per acre for recommended crop.

This system as contain some other feature such as display the current market price and approximated yield in quintal per acre for recommended crop. Those all details will helps to farmers for choosing the most profitable crop.

IV. RESULTS

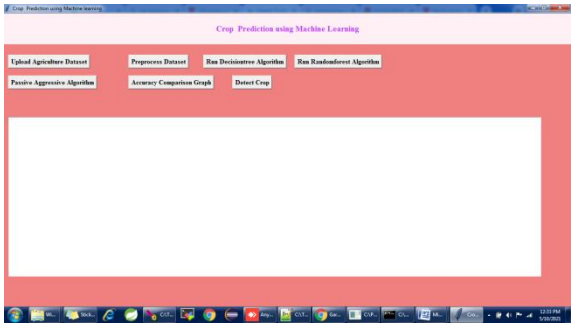

Fig.5.1 Overall GUI for Proposed Method

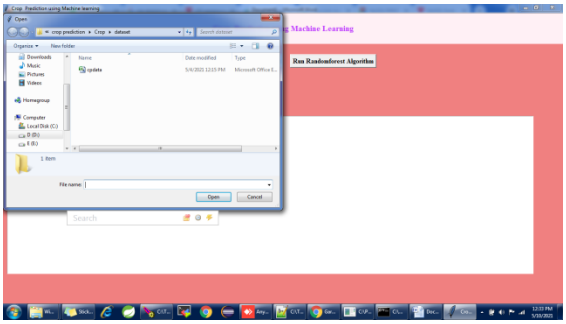

Fig. 5.2 Selection of dataset

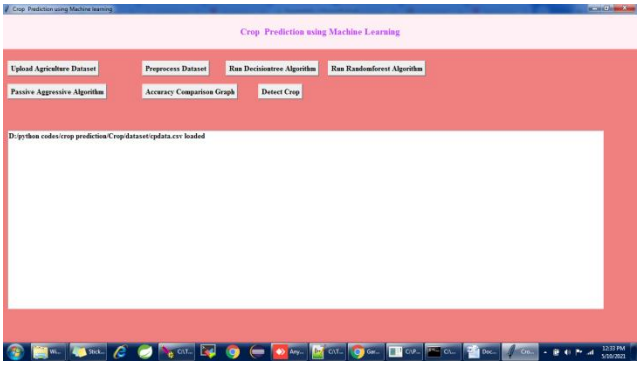

Fig. 5.3 Dataset Loaded is printed in text box

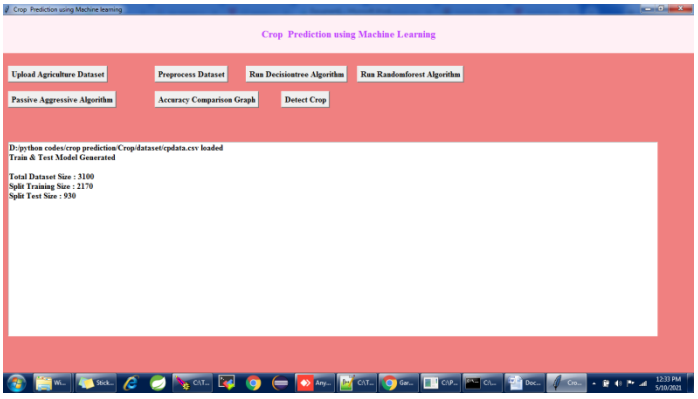

Fig. 5.4 Dataset Preprocessing (Dataset Splitting)

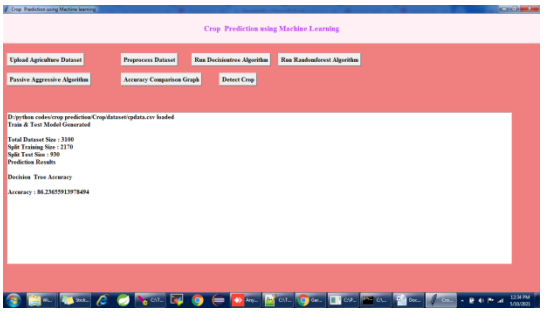

Fig. Accuracy of DT classifier

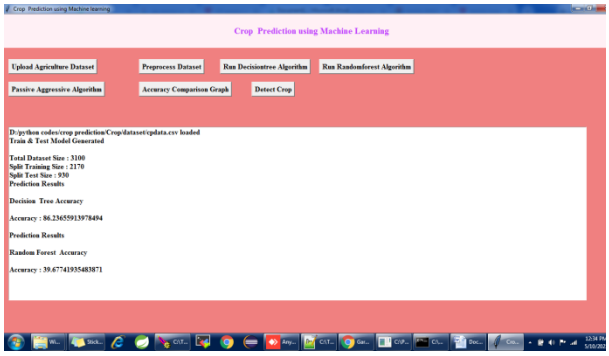

Fig. Random Forest Classifier Accuracy

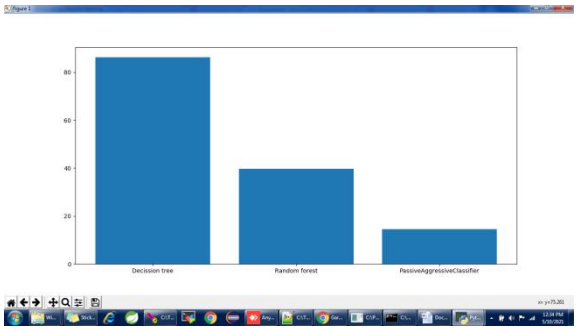

Fig. Comparison Graph of Three algorithms

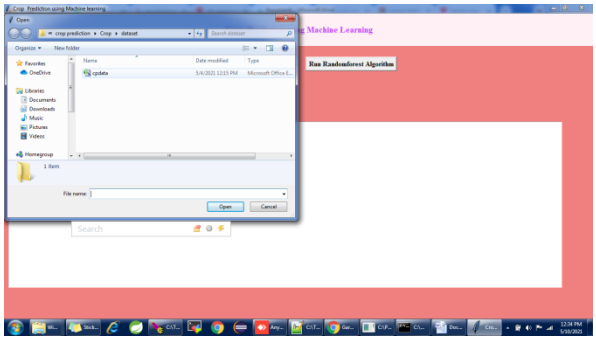

Fig. Crop prediction based on test sample data

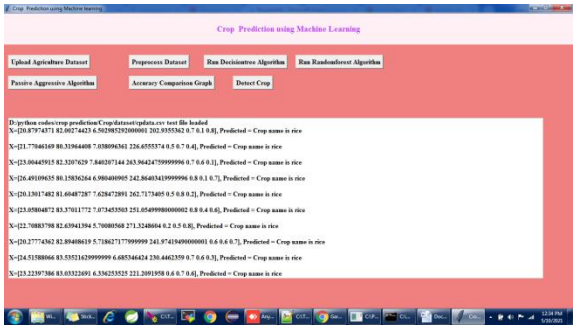

Fig. prediction results for given test data

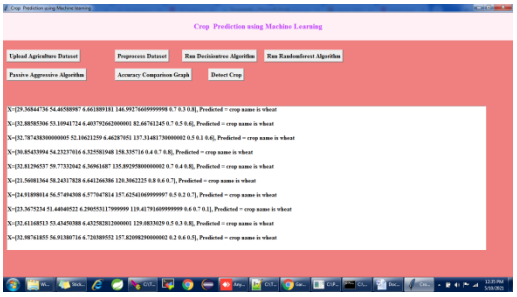

Fig. prediction results for given test data

V. CONCLUSION

Presently our farmers are not effectively using technology and analysis, so there may be a chance of wrong selection of crop for cultivation that will reduce their income. To reduce those type of loses we have developed a farmer friendly system with GUI, that will predict which would be the best suitable crop for particular land and this system will also provide information about required nutrients to add up, required seeds for cultivation, expected yield and market price. So, this makes the farmers to take right decision in selecting the crop for cultivation such that agricultural sector will be developed by innovative idea.

REFERENCES

[1] Niketa Gandhi et al, " Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India Using Data Mining Techniques", IEEE International Conference on Advances in Computer Applications (ICACA), 2016.

[2] K.E. Eswari. L.Vinitha. " Crop Yield Prediction in Tamil Nadu Using Baysian Network ", International Journal of Intellectual Advancements and Research in Engineering Computations, Volume-6, Issue-2,ISSN: 2348-2079.

[3] Shruti Mishra, Priyanka Paygude, Snehal Chaudhary, Sonali Idate “Use of Data Mining in Crop Yield Prediction” IEEE Xplore Compliant - Part Number: CFP18J06-ART, ISBN:978-1-5386-0807-4; DVD Part Number: CFP18J06DVD, ISBN:978-1-5386-0806-7.

[4] Anna Chlingaryana, Salah Sukkarieha, Brett Whelanb — Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review , Computers and Electronics in Agriculture 151 (2018) 61–69,Elisver,2018.

[5] Dakshayini Patil et al," Rice Crop Yield Prediction using Data Mining Techniques: An Overview", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 7, Issue 5, May 2017.

[6] Prof. D.S. Zingade, Omkar Buchade, Nilesh Mehta, Shubham Ghodekar, Chandan Mehta "Crop Prediction System using Machine Learning".

[7] Ashwani kumar Kushwaha, Swetabhattachrya "crop yield prediction using agro algorithm in hatoop".

[8] Girish L, Gangadhar S, Bharath T R, Balaji K S, Abhishek K T "Crop Yield and Rainfall Prediction in Tumakuru District using Machine Learning".

[9] Rahul Katarya, Ashutosh Raturi, Abhinav Mehndiratta, Abhinav Thapper "Impact of Machine Learning Techniques in Precision Agriculture".

[10] Pijush Samui, Venkata Ravibabu Mandla, Arun Krishna and Tarun Teja "Prediction of Rainfall Using Support Vector Machine and Relevance Vector Machine".