MAULANA AZAD
NATIONAL INSTITUTE OF TECHNOLOGY
BHOPAL – 462003 (INDIA)

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

KISAAN SAMADHAN

Minor Project Report

Submitted by

PIYUSH MISHRA	201112453
ANKIT ROHILA	201112448
AYUSH CHOUHAN	201112409
AYUSH JAIN	201112073

Under the Guidance of
Dr. R.K. Pateriya
Session: 2022-2023

MAULANA AZAD
NATIONAL INSTITUTE OF TECHNOLOGY
BHOPAL – 462003 (INDIA)

CERTIFICATE

This is to certify that the Project Report entitled, "KISAAN SAMADHAN" submitted by

PIYUSH MISHRA	201112453
ANKIT ROHILA	201112448
AYUSH CHOUHAN	201112409
AYUSH JAIN	201112073

to Maulana Azad National Institute of Technology, Bhopal , India, is a record of bonafide Project work carried out by him/her under my/our supervision and guidance and is worthy of consideration for the partial fulfilment of the award of the degree of Bachelor of Technology in Computer Science & Engineering of the Institute.

Dr. R.K. Pateriya
(Project Mentor)

DECLARATION

We, hereby declare that the following report which is being presented in the Minor Project Documentation Entitled "Kisaan Samadhan" is authentic documentation of our original work to the best of our knowledge. The following project and its report, in part or whole, have not been presented or submitted by us for any purpose in any other institute or organisation. Any contribution made to the research by others, with whom we have worked at Maulana Azad National Institute of Technology, Bhopal, or elsewhere, is explicitly acknowledged in the report.

Student Name	**Scholar Number**	**Signature**
PIYUSH MISHRA	201112453	
ANKIT ROHILA	201112448	
AYUSH CHOUHAN	201112409	
AYUSH JAIN	201112073	

ACKNOWLEDGEMENT

With due respect, we express our deep gratitude to our respected guide Dr. R.K. Pateriya, for his valuable help and guidance. We are thankful for the encouragement that he has given us in completing this project successfully.

We must mention the fact that the report of the project could not have been accomplished without the suggestions and advice of our project guide Dr. R.K. Pateriya and project coordinators Dr. Sanyam Shukla and Dr. Namita Tiwari.

We are also grateful to our respected director Dr. K.K. Shukla for permitting us to utilize all the necessary facilities of the college.

We are also thankful to all the other faculty, staff members, and laboratory attendants of our department for their kind cooperation and help. Last but certainly not least; we would like to express our deep appreciation towards our family members and batch mates for providing much-needed support and encouragement.

ABSTRACT

Agriculture is the most important sector of the Indian economy that provides employment to almost half the population of the country's workforce. It contributes about 18.8% of Gross Domestic Product (GDP) as per annual report 2021-2022 from Department of Agriculture, Co-operation and Farmer Welfare. But now-a-days, agriculture in India is undergoing a structural change leading to a crisis situation. The only remedy to the crisis is to do all that is possible to make agriculture a profitable enterprise and attract the farmers to continue the crop production activities. Farmers are not getting profitable margin for their products in the market. Even though India is being a second-largest producer of fruits and vegetables in the world as per annual report 2022 from Department of Agriculture, Co-operation and Farmer Welfare, farmers are financially distressed due to crop loss. The important causes for this are the fact of inappropriate crops selection, crops being wasted due to disease, not getting precise price for their crop due to surplus production of the crop. By analyzing the soil and environmental factor at particular region the best crop is predicted in order to have more crop yield. Our proposed machine learning model will help the farmers to choose appropriate crops for their land according to the soil type, temperature, humidity, water level, soil fertility and customer need. It will also help the farmer to identify any disease on their crop along with providing them, a suitable market price for their crops. The objective of the project is to provide a web application that helps farmers to maximize their profit by doing the above-mentioned tasks. Various machine learning algorithm such as Support Vector Machines (SVM), Multivariate Linear Regression, KNN, ANN, Random Forest etc. are used in designing of this model. Among them whichever algorithm will give the most accuracy, we will use that in our proposed system.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1. Introduction

From ancient days, agriculture is considered as the main source of supply to satisfy the daily needs of human lives. It is also considered a primary occupation, and also one of the India's major industrial sectors. The farmers are ought to follow a traditional naked eye observation and yielded healthy crops without the involvement of chemicals for animals and also to their cultivation land in order to keep healthy diversity.

But nowadays, weather conditions are being rapidly changing against the elemental assets to deplete the food and increase the security. Farmers are unaware of which crop to grow, and what is the right time and place to start due to uncertainty in climatic conditions. The usage of various fertilizers is also uncertain due to changes in seasonal climatic conditions and basic assets such as soil, water, and air. In this scenario, the crop yield rate is steadily declining.

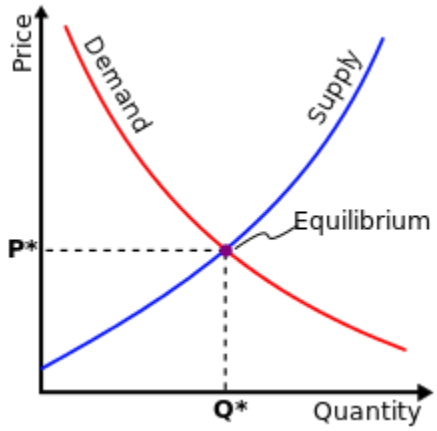

Fig1. Shows demand and supply in market

In topical years, there has been a significant variation in the rates of agricultural products like fruits, flowers, food-grains and vegetables, etc. As shown in Fig 1, the price variation is mainly due to the mismatch in demand and supply of these agricultural products. When the price of any commodity is set too high then the suppliers (farmers) tries to produce more goods to make more profit. When the price of any commodity is set too high then the consumers will tend to purchase less because of high rate, and the suppliers (farmers)incurs the loss. Conversely, if the supply is less for any commodities, as consumers have to compete with one other to buy the less supplied goods, results in increased price for the commodity, making consumers suffer with the high price. As there is no synchronization in production of and demand for the agricultural commodities either farmer fail to get good market prices for their products or consumer suffers high prices due to less production

Plant diseases are a severe threat to the entire production. Crop diseases are a major threat to food security, but their rapid identification remains difficult in many parts of the world due to the lack of the necessary infrastructure. It is essential for farmers to effectively deal with them and check them with the help of timely prevention. Also, the farmer don't get their deserved price for their production in the market.

Approximately 80 percent of farmers come from rural areas, and if the revenue from crop production goes down, their lifestyle would be influenced by the farms at industry level. This makes sense to farmers in India to show some special concern towards effective and precision farming to increase their profits. So, the deployment of one of the recent advancements in technology such as, Machine learning is one among the answer for predicting the crop with relation to atmospheric & soil parameter of the agricultural land. Since, now-a-day's ecological conditions aren't predictable like decades ago. It is changing day by day due to globalization. Hence, the farmers are facing difficulties in forecasting the weather and crops based on climate data.

The proposed system focusses on building a web-based application in which the user (farmer) will be provided with functionalities upon entering the website. The user (farmer) can get a crop prediction based on either environmental factor such as rainfall, soil, temperature etc, or based on customer requirement and market demand. Other functionalities include crop disease prediction and its solution, crop price predictor and best time to add fertilizer forecasting.

Once the crop is foretold, it will facilitate the farmers to predict the affordable crop for their individual land. Then, the farmers are guided with a detailed analysis to make them to understand that what quite seeds we will tend to sow in land to induce higher yielding. Within the past preceding data, crop prediction was calculated by analysing farmer's previous expertise on climatic condition. So, the correct data regarding history of climatic condition is a vital factor for creating selections in choosing crops. Therefore, our work proposes a thought to predict the affordable crop for the given input parameter for the poor farmers using machine learning. Thereby this proposed work will suggest the farmers with effective solutions for more profitable cultivation to maximize their profits.

Chapter 2. Literature Review and Survey

In this section, we review some of the significant works done in the agriculture field for crop prediction based on environmental factor and customer-based demands. We also review some of the research paper describing crop disease prediction techniques and crop price prediction techniques.

One of the early works developed a dedicated website to assess the impact of weather parameters on crop production in the identified districts of Madhya Pradesh. The districts were selected on the basis of the region covered by the crop. Based on these criteria, the first five top districts with a maximum crop area were chosen. The basis of the crops selected for the study was on prevailing crops in the selected districts. The crops picked included maize, soybean, wheat and paddy, for which the yield for a continuous period of 20 years of knowledge, were tabulated. The accuracy of the established model ranged from 76% to 90% for the chosen crops with an average accuracy of 82%.[1]

In [2], the authors used the sliding window non-linear regression method to suggest crop yield and price by evaluating patterns from the past data. The authors studied several districts of the state of Tamil Nadu. The authors proposed a system designed in such way that it suggests the best crop choices for a farmer to do farming. The designed system had done demand level classification. The demand level classification means the demand for the crops is predicted by classifying the dataset of change in the market prices of crops. In the proposed system the authors also did the text to speech conversion.

Sanjay Jharkharia, et al. studied the applicability of ARIMA models in the wholesale vegetable market to forecast the demand for vegetables to support the farmers and wholesalers. To achieve the same the authors have implemented the ARIMA model to the collected data for onion sales in Ahmedabad market and applied developed models to forecast the demand for the onions in Ahmedabad market. [3]

Arun Kumar & et al., "Efficient Crop Yield Prediction Using Machine Learning Algorithms". In this study, the classification of crop yields was performed to batch using Artificial neural networks based on yield productivity. And it will define the range of productivity. Regression is carried out to obtain the real crop yield and the expected cost. Nitin Singh & Saurabh Chaturvedi, "Weather forecasting using machine learning" is consider for collecting historical weather data from various weather stations to forecast weather conditions for future. [4]

Crop yield estimation using classification techniques estimates the crop yield and selects the most suitable crop for cultivation using data mining techniques thereby improving the value and gain of farming area. The disadvantages in this system are, that the methods to meet the demand and conveying the suggestions to farmers are not feasible. [5]

Shima Ramesh, Ramchandra Hebbar, PV Vinod," Plant Disease Detection Using Machine Learning" used Random Forest to detect whether it is diseased or healthy leaf. [6]

S.S. Sannakki and V. S. Rajpurohit, proposed a "Classification of Pomegranate Diseases Based on Back Propagation Neural Network" which mainly works on the method of Segment the defected area and color and texture are used as the features. Here they used neural network classifier for the classification. The main advantage is it Converts to L*a*b to extract chromaticity layers of the image and Categorization is found to be 97.30% accurate. The main disadvantage is that it is used only for the limited crops. [7]

Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa, proposed "Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease" Color histograms are extracted and transformed from RGB to HSV, RGB to L*a*b. They used nearest neighbors, Decision tree, random forest, extremely randomized tree, Naïve Bayes and SV classifier. In seven classifiers extremely, randomized trees yield a very high score, provide real time information provide flexibility to the application [8]

Ganesh Sastry Kakaraparthi, B.V.A.N.S.S. Prabhakar Rao project aims to solve crop price forecasting problem in an efficient way to ensure guaranteed benefits to poor farmers. It uses machine learning techniques different data came out with better solutions. This solution uses Decision Tree regression techniques to predict crop values using trained data from certified datasets. [9]

Awanit Kumar, Shiv Kumar suggested methods for predicting crop production for the coming year. The system employs fuzzy logic-based prediction methods. Fuzzy logic is a rule-based forecasting logic in which rules are applied to the land for agriculture, rainfall, and crop forecasting. K-means can be used to analyze the received data set by the system [10]

Chapter 3. Gaps Identified

There is a lack of study which provides a complete solution to farmer problems. The studies which are currently available, provide solution to only one of the many existing problems faced by an Indian farmer. Our work emphasizes on providing a complete solution to any problem faced by the farmers from the time of seed selection to harvesting of their crops to selling of these harvested crops.

Despite many solutions that have been recently proposed, there are still open challenges in creating a user-friendly application with respect to crop recommendation and price prediction. The solution proposed here aims to solve these limitations, by developing a user-friendly application that considers the parameters like rainfall, temperature, soil type etc. that directly affect cultivation. The main objective is to obtain a better variety of crops that can be grown over the season so that the farmer can maximize his profit. The proposed system would help to minimize the difficulties faced by farmers in choosing a crop and maximize the yield.

Many solutions that are proposed depends purely on historical data and do not use Machine Learning Algorithm to predict the crop yield. Since the climatic conditions as well as the demand market is continuously changing, these solutions will not give correct or comprehensive results. Also, we will be keeping a dynamic dataset, which will be updated regularly to get accurate prediction according to the time in which we are using the model. Our disease prediction model dataset updates itself as soon as new data is fed to it. So, it is a self-learning model.

Chapter 4. Proposed Work

In our proposed work we are using different functionalities to provide solution to farmers problem. A web application will be used to provide connectivity to the farmers. GPS helps to identify the user location. After logging in on the website, the farmers can use these four functionalities –

1) Get a prediction of most beneficial crop to grow based on soil and area of land. The user can also get an idea of what is the right time to use fertilizer in the crops for maximum yield. This feature will also provide the prediction of yield of a crop entered by the farmer.
2) Get a prediction of most beneficial crop to grow to based on customer demand and market need. It will also suggest alternate crops to grow if the demand of the crop entered by farmer is not satisfactory.
3) Detect an anomaly on the crops grown by the farmer. The farmer will have to upload images of their crop, if an anomaly (disease) is detected, then the farmer will be provided a complete solution to eliminate that disease. This function will also predict the correct ages of the crops.
4) Get a price prediction for their crops based on the current market so that the farmers can maximize their benefits.

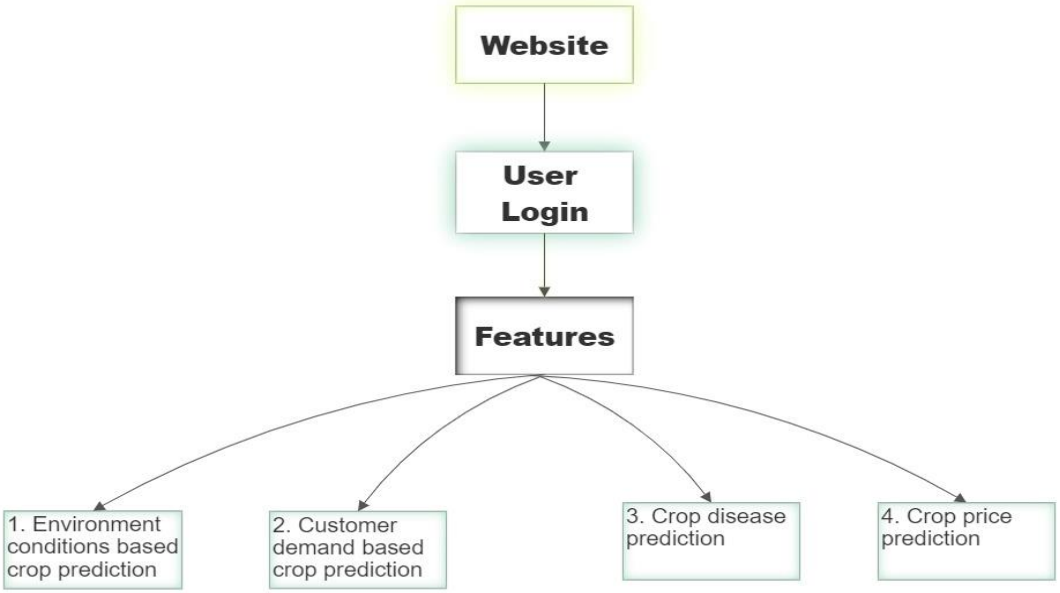

Fig2. Block diagram represents working of model

Chapter 5. Methodology

To create our website, we will be using Node JS at the backend and MongoDB to store our data. For frontend HTML, CSS and JavaScript will be used along with React JS for frontend framework.

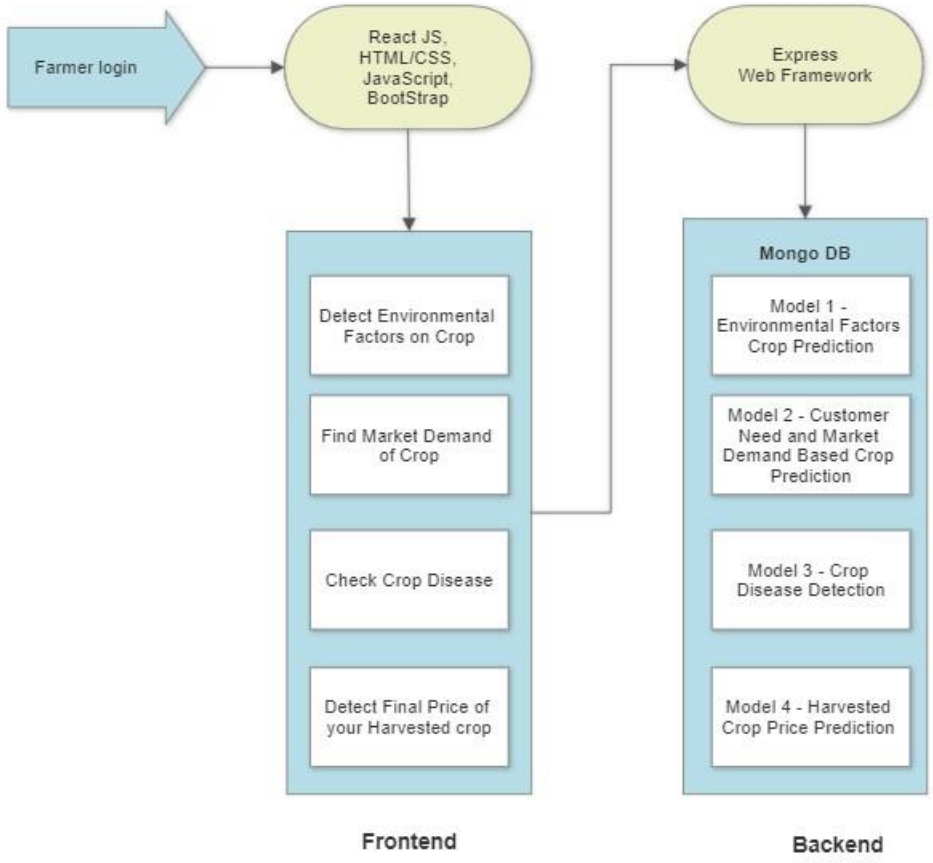

Fig3. Shows website model

For providing the four main functionalities we are using 4 machine learning models, which are based on various machine learning algorithms.

A)- MODEL 1 – ENVIRONMENTAL FACTORS BASED CROP PREDCITON

The solution to the problem is to provide a smart user-friendly recommender system to the farmers. To predict the crop yield, we are selecting five Machine Learning algorithms - Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Multivariate Linear Regression (MLR), and K-Nearest Neighbor (KNN). Among them whichever algorithm will give the most accuracy, we will use that in our proposed system. Additionally, the system will also suggest the best time to use the fertilizers to boost up the yield. Machine learning algorithms allow choosing the most profitable crop list or predicting the crop yield for a user-selected crop. Prediction of the crop yield for specific regions by executing various Machine Learning algorithms, with a comparison of error rate and accuracy.

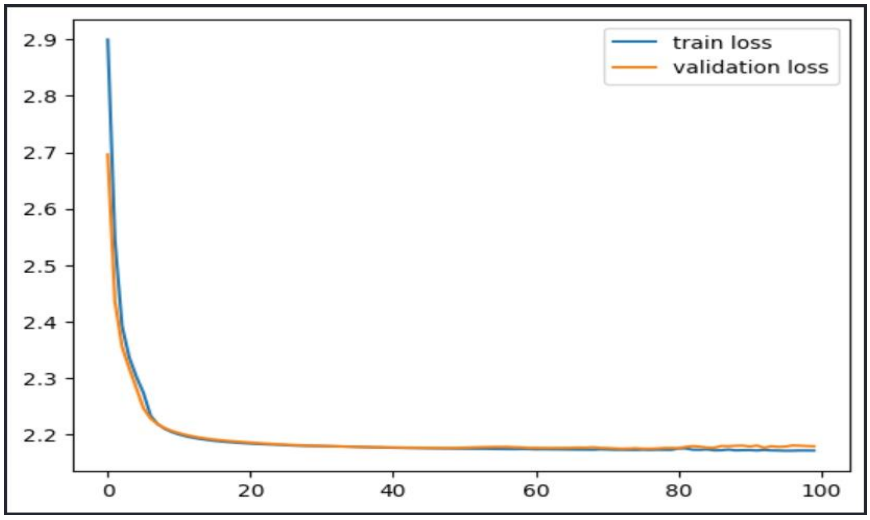

Fig4. Train loss v/s validation loss

The model predicts the crop yield for the data sets of the given region. Integrating agriculture and ML will contribute to more enhancements in the agriculture sector by increasing the yields and optimizing the resources involved. Historical data is collected from various reliable sources like data.gov.in, kaggle.com, and indianwaterportal.com. The data sets are collected for Maharashtra and Karnataka regions. The data has various attributes like state, district, year, season, type of crop, an area under cultivation, soil, production, etc.

Figure 4 is the system architecture of the proposed model. It's a web app that has two modules – the prediction module and the fertilizer module. The farmer needs to register with the app through the registration process. Once the registration is complete, the farmer can use the web application services. The prediction module predicts the crop yield using the selected attributes from the data sets for the specific crop. The predict module also suggests the farmer with the highest yield crops. The fertilizer module guides the farmer for the right time to use the fertilizer.

Figure 5 illustrates the flow chart of the proposed system. It describes the whole process starting with the registration and various services provided by the mobile application.

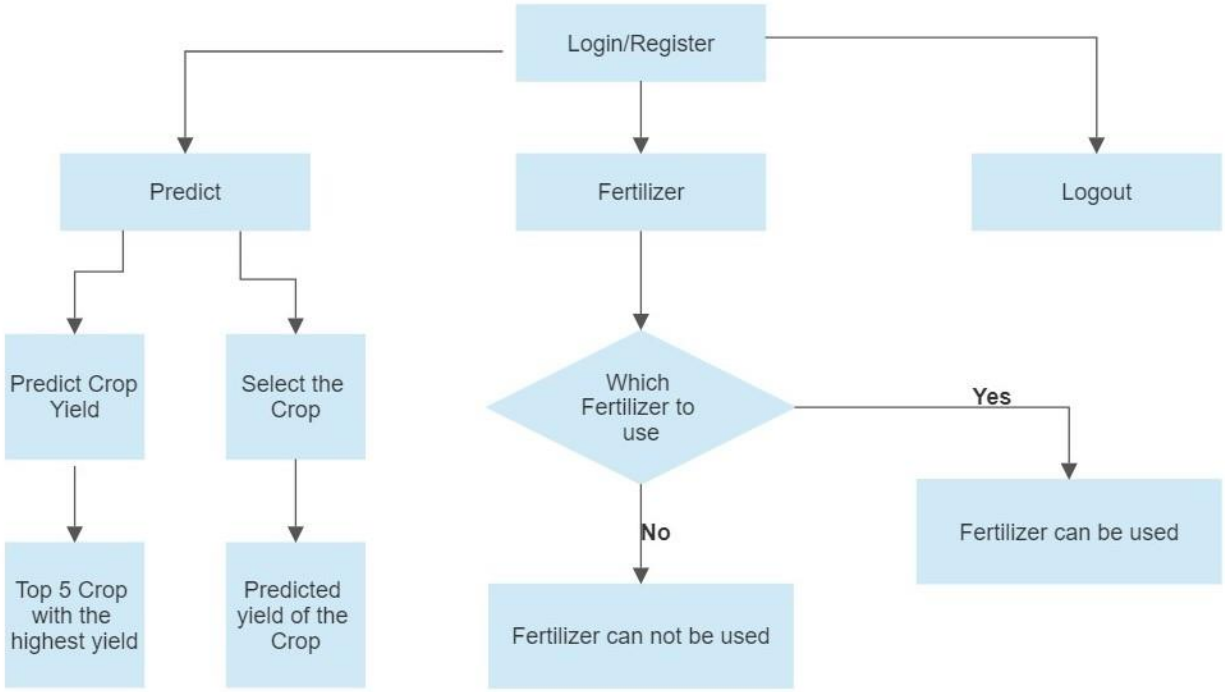

Fig5. Illustrate flow chart of proposed model -1

If the farmer is not sure about the crop to be planned this year, he can use the crop recommender system. In the crop recommender system, the farmer must provide only soil type and area. The system lists the crops with their predicted yield. This makes farmers easy to decide on a crop to be planted.

The timing of applying the fertilizer is very crucial. The farmer's effort and money will get wasted if the rain comes down too early. The proposed fertilizer usage service will guide the farmer on when to use the fertilizer. The model predicts the rain for the specific location for the next 14 days with Open Weather API. If the rainfall is more than 1.25 mm then it recommends as ‘not safe’ to use the fertilizers.

The machine learning approach is used for crop yield prediction. The patterns and correlations are discovered using ML approach. The model is trained using historical data sets where the past experience is used to represent the outcome.

B)- MODEL 2 – CUSTOMER NEED AND MARKET DEMAND BASED CROP PREDCITON

After predicting the suitable crop, we proceeded to yield prediction and customer demand-based forecasting. Yield prediction plays a crucial role in helping states estimate the crop output in a given year, which can help control market prices. This study focuses on forecasting crop yield in advance by utilizing machine learning techniques to analyse various factors, such as location, season, and

crop type, on the collected dataset available on data.gov.in. The yield prediction dataset consists of attributes such as State Name, District Name, Crop Year, Season, Crop, Area, and Production.

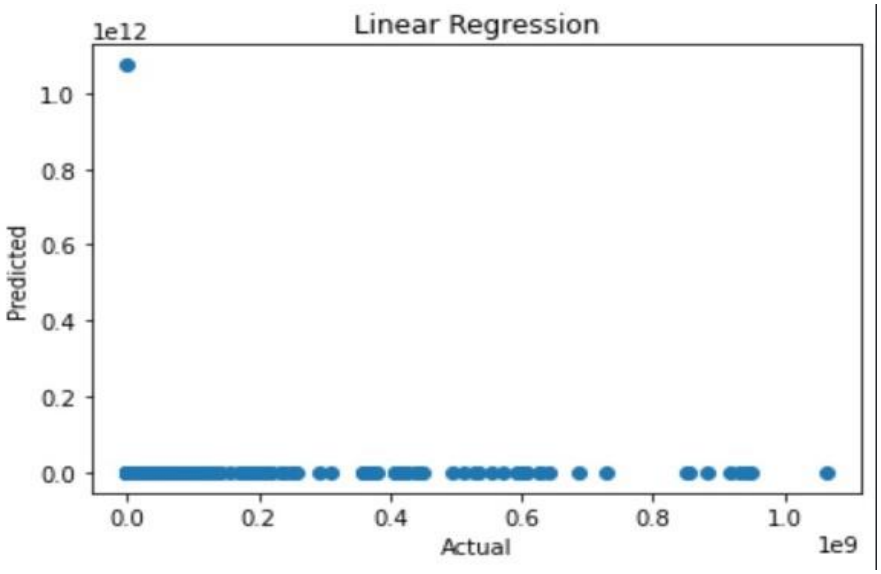

Fig6. Accuracy graph for linear regression

We employed linear regression, random forest, support vector regression, and decision tree algorithms to analyse which algorithm provided the best results for this problem. However, we found that the linear regression algorithm was not accurate for this type of prediction. Therefore, we evaluated the performance of the other algorithms and found that the random forest algorithm had a higher accuracy level, ranging from 85% to 90%. Despite its high accuracy, the random forest algorithm was relatively slow in terms of processing time.

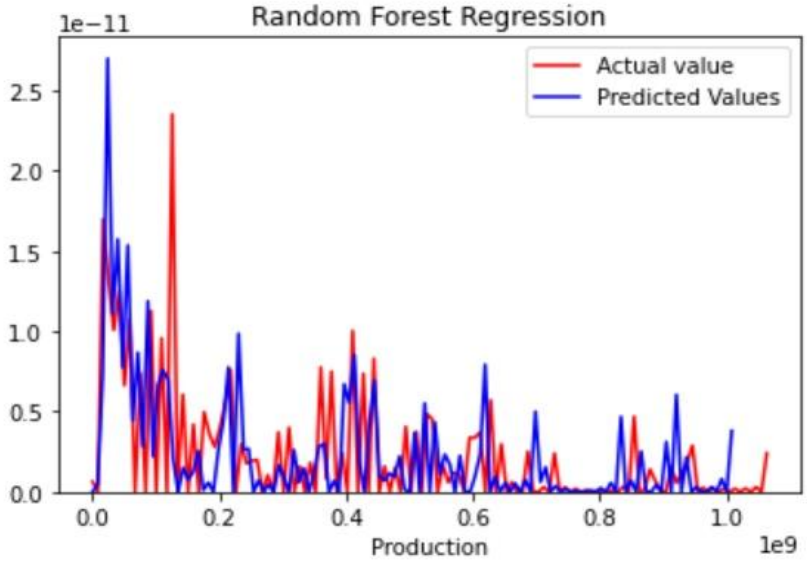

Fig7. Accuracy graph for random forest regression

Overall, our study highlights the importance of yield prediction for managing market prices and how machine learning techniques can be utilized to provide accurate yield predictions. By analyzing

various factors, such as location, season, and crop type, this model can predict crop yields in advance, which can help farmers plan for the upcoming season and improve their overall crop output.

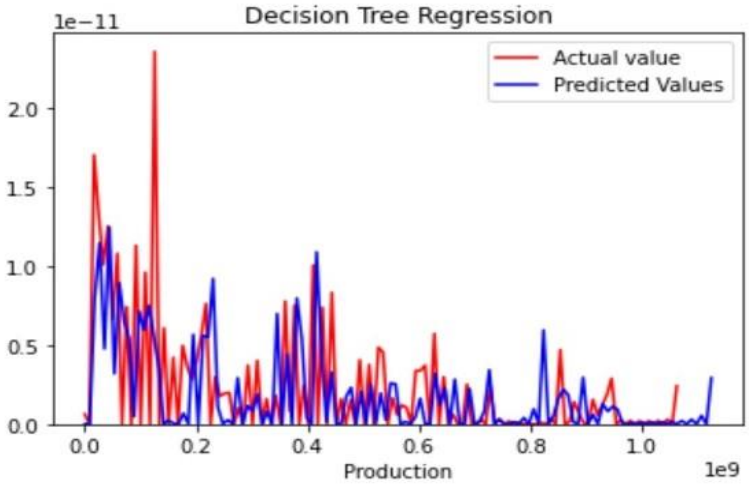

Fig8. Accuracy graph for decision tree regression

To bridge the gap between the demand and supply of food crops, it is essential to accurately forecast the demand for various food commodities and guide the farmers accordingly. This ensures that the farmers cultivate crops that meet the actual demand of the society, thus preventing any wastage of resources and minimizing losses for both consumers and farmers. With the emergence of big data analytics, it is possible to provide effective solutions to such problems. By analysing vast amounts of data, including weather patterns, market trends, and consumer behaviour, farmers can make informed decisions about which crops to grow and when to grow them. This not only helps in reducing the mismatch in demand and supply but also ensures that the farmers maximize their profits and contribute to the overall growth of the agricultural sector.

This report presents a novel forecasting model aimed at reducing the gap between the demand and supply of crops. The model predicts the demand for various food commodities, providing guidance to farmers on selecting appropriate crops to satisfy the demand. By doing so, it helps reduce the mismatch between demand and supply. To develop the forecasting model, we have collected agricultural data from the previous year, which includes information on crop areas, types of crops cultivated, soil characteristics, yields, and overall crop consumption. This data is then used to predict the future demand for crops, which enables farmers to make informed decisions on crop selection and cultivation.

The system applies data mining techniques to classify the collected dataset and predict the demand of crops based on changes in market prices. The dataset is classified into various categories such as high demand, moderate demand, and low demand. The system then recommends the appropriate crops to be grown to meet the demand for each category. This helps farmers to make informed

decisions about crop selection and avoid overproduction of crops that have low demand in the market. By utilizing these data-driven insights, the system can help reduce the waste of resources and maximize profitability for farmers.

The system classifies the data into three categories based on the change in the market prices of the crops.

The system classifies the data into three categories based on the change in the market prices of the crops.

1. Category is Excess, which is characterized by a sudden decrease in the price or less consumption of the crop compared to the cultivation. This indicates that the supply of the crop is more than the demand, leading to a surplus in the market.
2. Category is Scarce, which is characterized by a sudden increase in the price or less consumption of the crop compared to the cultivation. This indicates that the demand for the crop is more than the supply, leading to a shortage in the market.
3. Category is Neutral, which is characterized by no change in the price and a standard price is maintained throughout. This indicates that the supply and demand of the crop are in balance, leading to a stable market. By classifying the data into these categories, the system can provide guidance to farmers in selecting and growing the appropriate crops to meet the demand and avoid the gap between demand and supply.

If the demand is low, the system suggests an alternate crop, and if it is high, the recommended crop is the one with the corresponding high demand grade. To calculate demand, the formula take into account the outlier, which indicates a sudden increase in crop price that is not within the same price range as the other time periods. The calculated demand is based on the total crop cultivated, total crop consumed, and change in price, which is the difference between the outlier price and the sample mean price.

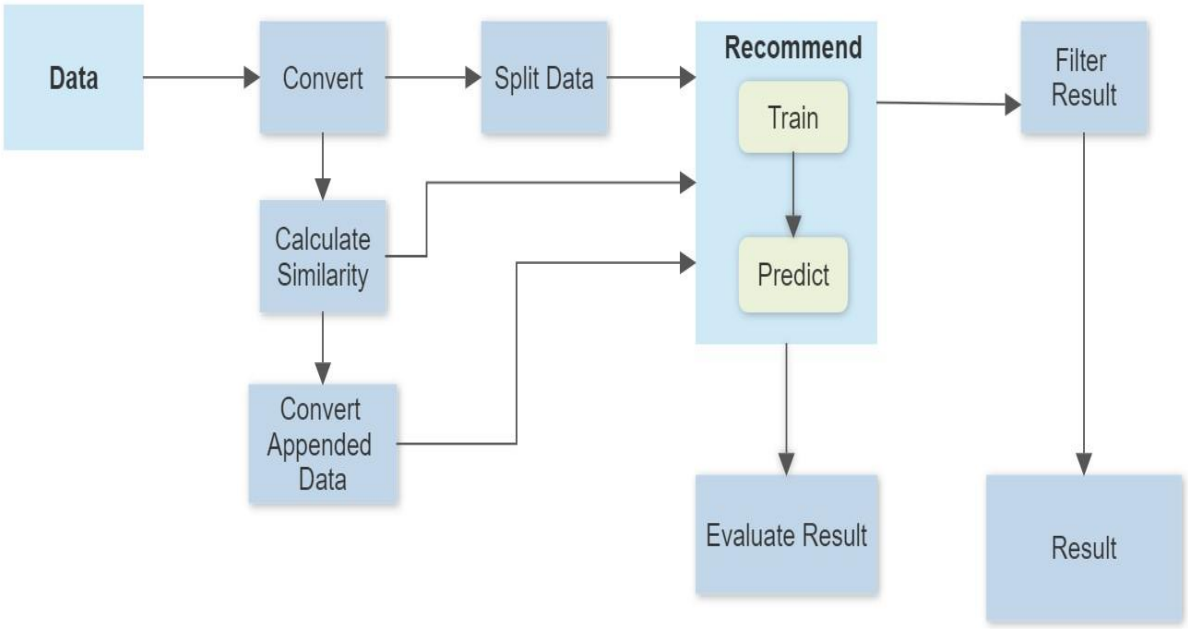

Fig9. Illustrate flow chart of proposed model -2

C)- MODEL 3 – CROP DISEASE DETECTION

Crop disease detection is a critical aspect of modern agriculture, which can assist farmers in improving crop yields, reducing economic losses, and promoting sustainable farming practices. Crop diseases can have a significant impact on agriculture, leading to reduced yields, lower crop quality, and decreased profits for farmers. Early detection and accurate diagnosis of crop diseases are essential for farmers to take timely action to manage and control the spread of the disease, protecting their investment and reducing crop losses.

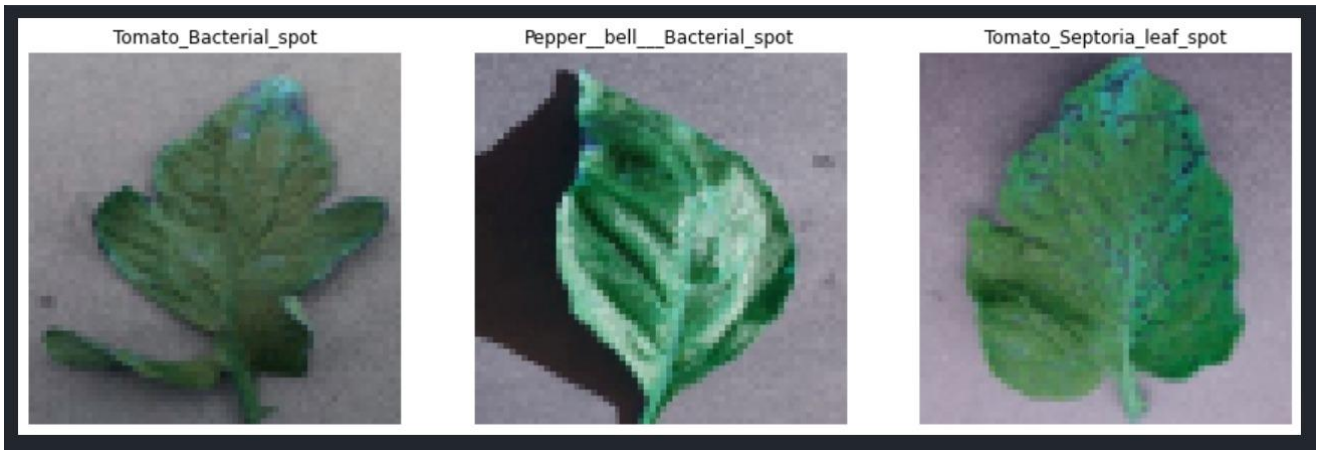

Fig10. Instance of dataset

Furthermore, excessive use of pesticides and other chemical treatments can be costly and can have negative impacts on the environment. Early detection of crop diseases can help farmers use more

targeted and effective treatments, reducing the need for excessive use of pesticides and other chemical treatments. This, in turn, promotes sustainable farming practices that prioritize environmental stewardship and reduce the negative impact of farming on the environment.

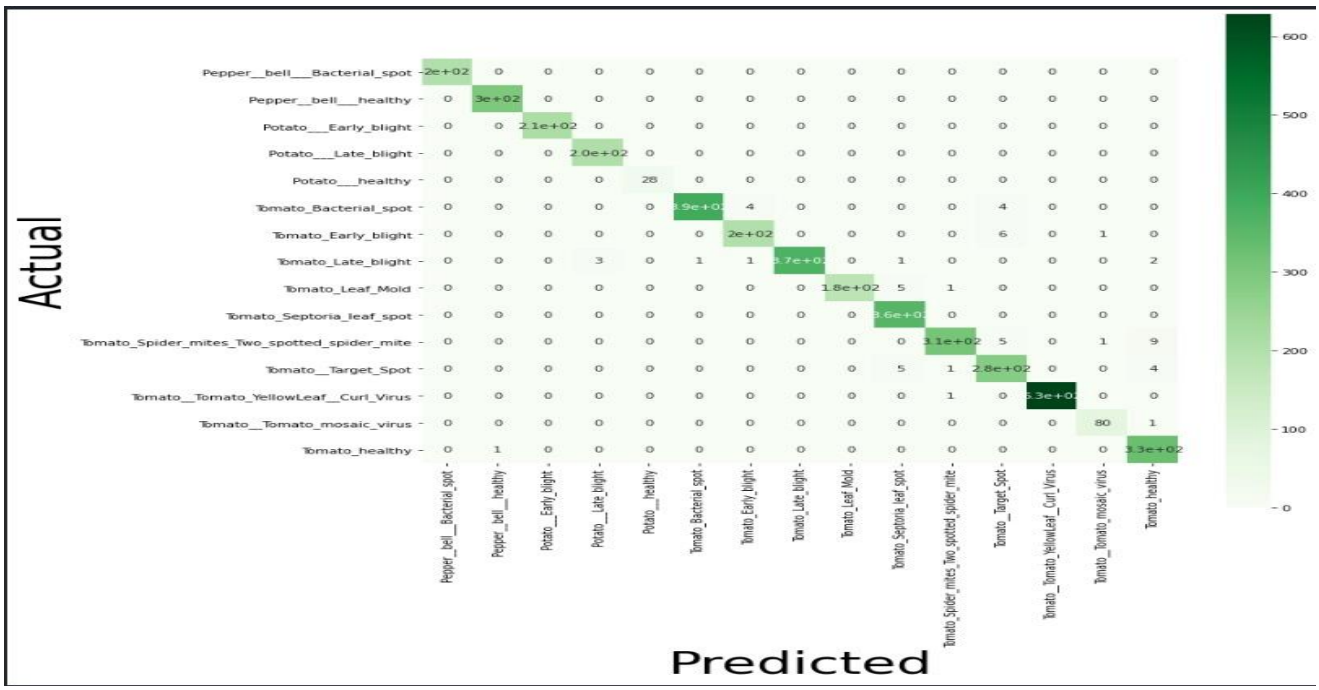

Fig11. Confusion matrix

Moreover, effective crop disease detection and management can contribute to global food security by improving crop productivity and reducing crop losses due to disease. With the global population projected to reach 9.7 billion by 2050, it is crucial to ensure that crops are protected from disease and that crop yields are maximized to meet the growing demand for food.

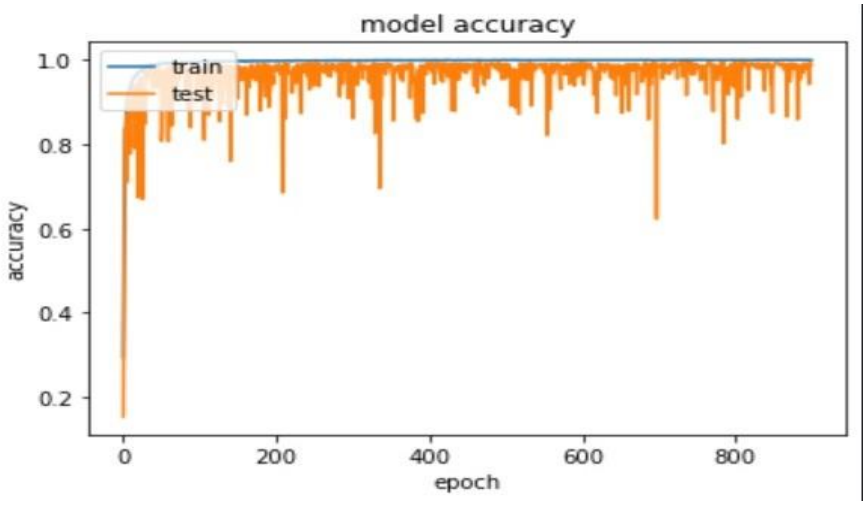

Fig12. Model accuracy

The use of deep learning models such as DenseNet for plant disease detection and classification has shown promising results. The DenseNet model is trained on a preprocessed dataset using the

backpropagation algorithm and an appropriate optimizer to achieve high accuracy in predicting plant diseases. The success of the prediction largely depends on the quality of the dataset and the accuracy of the model training. Therefore, it is crucial to pay attention to data collection, pre processing, and model training to achieve high accuracy in predicting plant diseases using DenseNet.

Transfer learning is a useful technique that reduces the volume of training data, training time, and processing costs for creating deep learning models. It enables the learning from one model to be applied to another, making it possible to reuse features across layers, increasing the efficiency of the architecture's parameters, and allowing for greater variation and better performance in later layers.

The DenseNet model, which employs transfer learning to automatically extract features and leverages their weights learned on the ImageNet dataset, is a powerful tool for reducing calculation workload and enhancing model performance. The architecture of DenseNet201 allows for the construction of simple and straightforward models, which can be adapted to a wider range of plants and diseases. Nonlinear transformations such as convolution, pooling, rectified linear units, and batch normalization are implemented in each layer of the DenseNet network, making it an effective tool for crop disease detection.

In conclusion, crop disease detection is an essential component of modern agriculture that can help farmers increase crop yields, reduce economic losses, promote sustainable farming practices, and contribute to global food security. The use of deep learning models such as DenseNet and transfer learning techniques has shown promising results in crop disease detection and classification. It is important for farmers to stay up-to-date with the latest crop disease detection techniques and tools to ensure healthy and productive crops for future generations. Future research can explore the use of other CNNs that have already undergone multi-classification training and update the developed tool to be more accurate in the treatment section, tak into account crop conditions and weather data for better diagnosing plant diseases, particularly in India

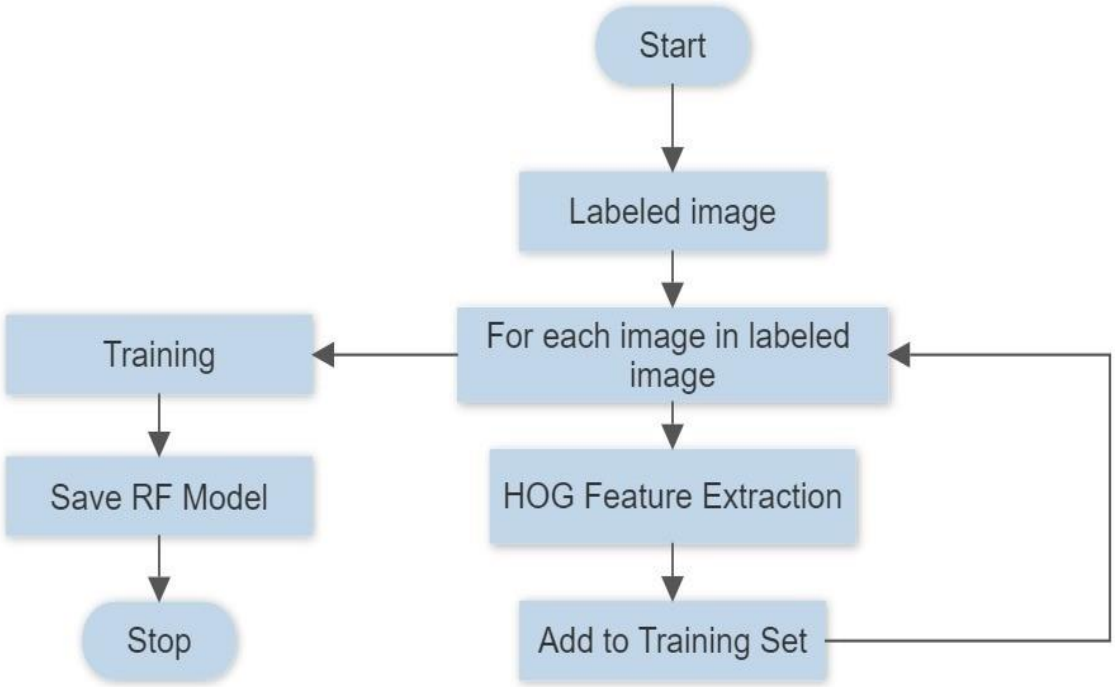

Fig13. Illustrate flow chart of proposed model -3

D) MODEL 4 – HARVESTED CROP PRICE PREDICTION

Our crop price forecasting model is designed to revolutionize the way farmers predict crop prices, offering them an efficient and effective solution to ensure maximum benefits. By utilizing machine learning techniques, we have been able to generate improved results using various data sets. Our model employs Decision Tree regression techniques to predict crop values, based on trained data from certified datasets, ultimately leading to increased productivity. The model also take into account various environmental conditions that could potentially affect crop growth. The web application provides easily accessible results for poor farmers, ultimately helping them make informed decisions. Unlike traditional technical analysis methods, our machine learning-driven price predictions offer a unique way of combining technical and fundamental analysis methods. Additionally, our approach differs in that it take into account real-time data and incorporates non-linear relationships between variables, resulting in more accurate and reliable predictions.

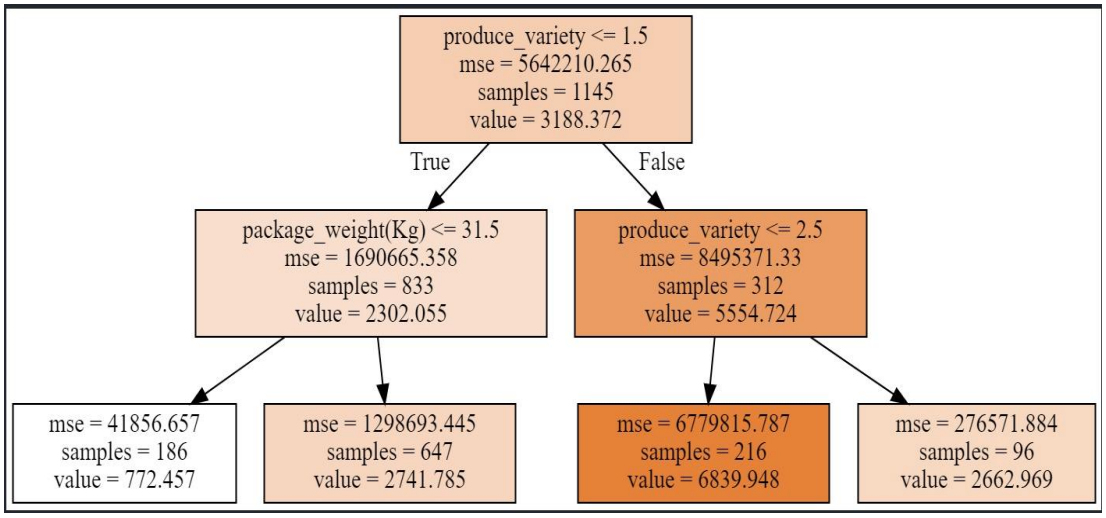

Fig14. Real decision tree

The innovative approach of machine learning-driven price prediction offers a multitude of advantages over traditional methods. One significant advantage is its ability to account for multiple price determinants, which greatly enhances the accuracy of predictions. In addition, this approach provides localized predictions at the level of a specific region or even a farm, making them highly relevant and actionable for farmers. The combination of technical and fundamental analysis further improves accuracy by providing short-term predictions through technical analysis and long-term forecasts through fundamental analysis. This integration of both analyses offers a comprehensive understanding of external and internal factors that affect commodity prices, leading to more accurate and reliable predictions. Ultimately, this results in better decision-making and higher profits for farmers.

The dataset has been imported from Kaggle (Prices for Food Crops_Commodities_2012_to_2015). We find that our Linear regression model has an accuracy of around 25\% Linear regression can be prone to underfitting the data. Hence if you build a model using linear regression and you find that both the test accuracy and the training accuracy are low then this would likely be due to underfitting.

We find that our model has an improved accuracy of around 97\% with use of Random Forest Regression. Random forest works by builds multiple decision trees and merges them together. Often it's more accurate and stable in prediction. Random decision forests correct the decision tree’s habit of overfitting to their training set.

In Decision Trees, the process of predicting the record's class label begins at the top of the tree and follows a path down to the leaf nodes. At each node, the value of the record's attribute is compared

to the value of the attribute used to split the data. Based on the comparison, the corresponding branch is followed to the next node until a leaf node is reached, which provides the predicted class label.

When constructing a Decision Tree, selecting the best attribute to use as the root node or at different levels of the tree can be a challenging task, especially when dealing with datasets that contain numerous attributes. Randomly selecting nodes as the root may result in poor performance and low accuracy. To overcome this, information gain is used as a criterion to evaluate the importance of attributes for classification, while the Gini index is utilized for continuous attributes.

In the case of crop price prediction, Decision Tree Regression is employed since crop prices are seasonal and vary with time. The dataset contains rainfall and WPI as the parameters for crop price prediction. By using Decision Tree Regression and considering relevant parameters, our model aims to provide accurate and reliable predictions to help farmers make informed decisions.

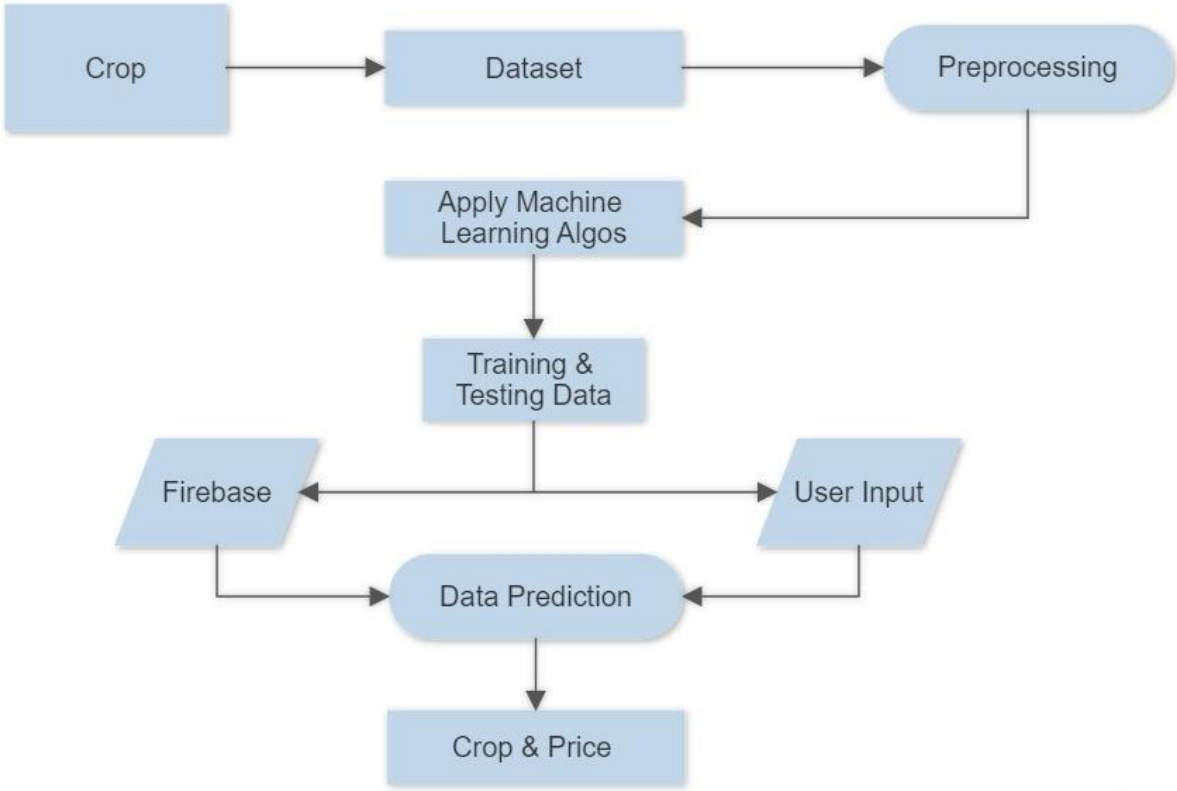

Fig15. Illustrate flow chart of proposed model -4

Chapter 6. Results and Analysis

Model Name	Best Algorithm from found	Accuracy
Crop Prediction	Deep Neural Network	97%
Yield Prediction	Random Forest	98.6%
Crop Disease Detection	DenseNet	95%
Crop Price Prediction	Random Forest Regression	97%

Table 1. Maximum accuracy of different models

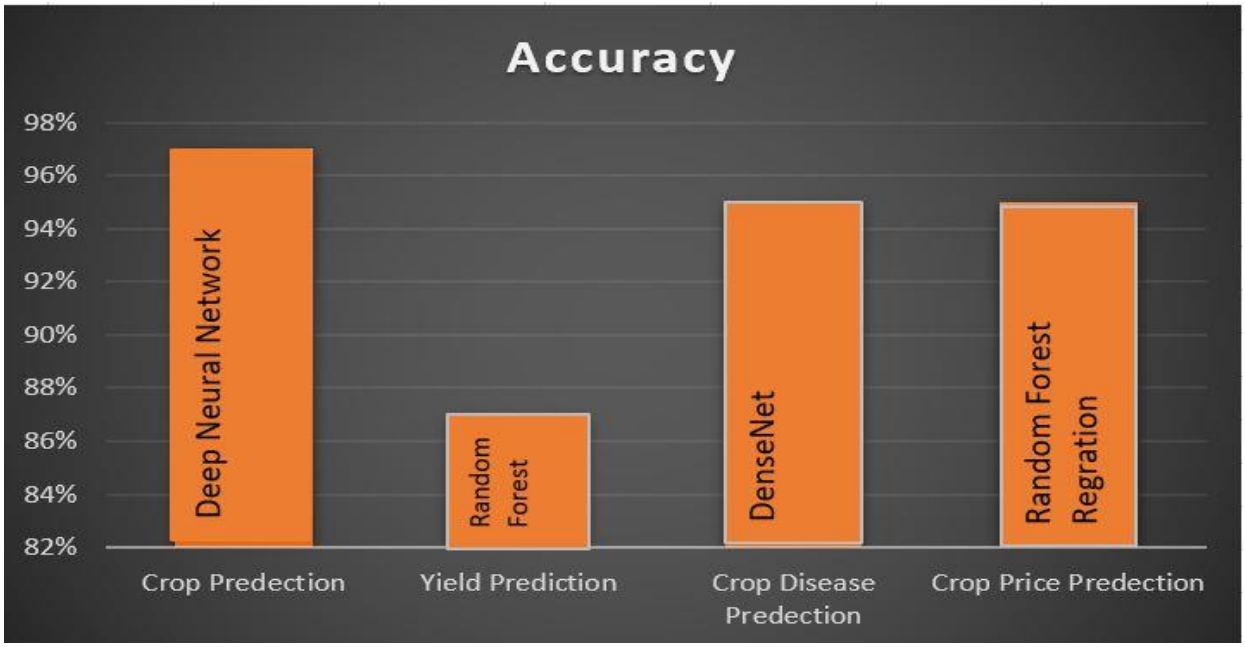

Fig16. Bar chart shows accuracy of models

Chapter 7. Conclusion

To improve profitability in an increasingly competitive and volatile environment, farmers need to have a deep understanding of their profit drivers and tactical management plans in place. .The main emphasis of our project is to help the farmer increase their profits. On completion of this project out model will be able to –

1) Predict the crops based on environmental conditions such as soil, water, rainfall, sunlight.
2) Predict the crops based on customer need and market demand.
3) Predict the correct time to add fertilizers in the crop to increase their yield.
4) Provide a detailed analysis of the future yield of the predicted crops.
5) Recognize any diseases on the growing crops and provide a detailed solution to cure those disease so that the yield of the crop is not affected.
6) Help farmer get a fair price of their crops in the market by predicting a excepted price of their crops based on the current market rates.

Machine Learning based prediction provides a unique way of combining technical and fundamental analysis methods. While technical analysis solely looks at historical price, Machine learning driven price predictions differ in two main aspects from traditional approaches. It can account for multiple price determinants thus improving accuracy. It provides localized predictions thus increasing their usability price, fundamental analysis consists of understanding external and internal factors that influence the prices of a certain commodity. Our proposed models uses various machine learning algorithm such as KNN , Random Forest , ANN, Decision Tree Regression to improve the accuracy of the predictions .

References

[1]. Veenadhari S, Misra B, Singh CD, “Machine learning approach for forecasting crop yield based on climatic parameters”, In 2014 International Conference on Computer Communication and Informatics, 2014 Jan 3 (pp. 1-5). IEEE.

[2]. M.Kalimuthu, “Crop Prediction using Machine Learning”, In Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020) IEEE

[3]. G. M. Naidu, "Applicability of arima models in wholesale vegetable market", *Int. J. Agric. Stat. Sci.*, vol. 11, no. 1, pp. 69-72, 2015.

[4]. Arun Kumar, Naveen Kumar and Vishal Vats, "Efficient Crop Yield Prediction Using Machine Learning Algorithms", *International Research Journal of Engineering and Technology (IRJET)*, vol. 05, no. 2395-0072, June 2018, ISSN 2395-0056.

[5].R.Sujatha and P.Isakki, ”Crop yield estimation using classification techniques” Computational Intelligence and Communication Networks (CICN), International Conference 7-9 Jan. 2016 [15] Tng Zhang, "

[6]. Shima Ramesh, Ramchandra Hebbar, PV Vinod,” Plant Disease Detection Using Machine Learning” In 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control.

[7]. S. S. Sannakki and V. S. Rajpurohit,” Classification of Pomegranate Diseases Based on Back Propagation Neural Network,” International Research Journal of Engineering and Technology (IRJET), Vol2 Issue: 02 | May-2015

[8]. Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa, proposed “Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease” Preceding of the 1’st international conference on the use of mobile ICT in Africa ,2014.

[9].Md. Tahmid Shakoor, Karishma Rahman, Sumaiya Nasrin Rayta, Amitabha Chakrabarty.2017. Agricultural Production Output Prediction Using Supervised Machine Learning Techniques.978- 1-5386-3831-6/17/$31.00 ©2017 IEEE.

[10].G Srivatsa Sharma, Shah Nawaz Mandal, Shruti Kulkarni, Monica R Mundada, Meeradevi.2018. Predictive Analysis to Improve Crop Yield Using a Neural Network Model. 978-1- 5386-5314- 2/18/$31.00 ©2018 IEEE.