E-commerce Farming Using Machine Learning

Guided by **Dr. T.A.Chavan**
(Ph.D. (CS & E),M. Tech., LMISTE, Asst. Prof. in I.T., SKNCOE)
Special Thanks to **Anil Audumbar Pise**
(Ph.D. Research Scholar, University of Witwatersrand Johannesburg, South Africa)
Pranali Pawar, Purva Mahajan, Alfiya Husain, Mayuri Gurav
Department of information technology,
Smt Kashibai Navale college of Engineering,
University of Pune.

Abstract— **India is an agrarian country and its economy is largely based on crop productivity. Many times farmers don't get fair price for their product due to intermediates present in supply chain. We have developed a web application where farmers can add their product and details of their product. Customers can buy product according to their needs. Customer can buy different products from different farmers according to their quality and price. It also provides suggestions to the farmers about which alternative crops can be cultivated and how much yield will be produced. This process of suggestion is done using naïve bayes algorithms and apriori algorithm. In this algorithm inputs such as temperature, climate, soil type is taken from farmers. Using trained dataset and algorithms result is obtained. Also by using naïve bayes algorithm we can also obtain yield of the product and duration needed to produce. This will help farmers for taking decision of cultivation. Farmers need to take good care of the product while growing. There are many diseases which can affect plant and destroy the crop. Farmers need to pay attention and protect the plant from being affected and also should be aware of the diseases so that it can recover quickly. In this we will be providing disease detection and some measures to cure. This application also provides government schemes available for farmer so that they can take advantage of it. This system will provide fair price to customers and as well as farmers will gain fair price for their product.**

Keywords— Climate, agricultural productivity, prediction, classification.

I. INTRODUCTION

Machine Learning deals with finding previously unknown patterns and trends in databases and using that information to build predictive models. As we step forward into the modern era of technology, we may find many engineering related applications which are useful for society and people. This is the world of technology where people use smart phones for completing their daily tasks like shopping, paying bills, managing work and much more. The idea of this project is to break the supply chain of indirect sales, so that the farmer can be connected directly to the customer and the selling can be done accordingly.

Prediction of crop and their yield can also be found out using naïve bayes and apriori algorithm. The external factors that affect the crop yield are Climatic, Edaphic, Biotic, Physiographic, and Socio-economic factors. Soil temperature, Soil mineral matter, Soil organic matter, Soil organisms, Soil reactions affects the plant growth. Soil moisture is the available water content in the soil pores, which hold the water content. As in most cases, the water available in the soil pores is responsible for the yield of the plant. The chemical and biological activities in the soil are dependent on the soil moisture. Two different machine learning classifiers were used to classify the crop yield with respect to the crop selection. Machine learning classifier such as naïve bayes classifier and apriori classifier were used. Among all two classifiers Naïve Bayes Classifier showed better accuracy when compared with other classifiers.

Plants are attacked by numerous kinds of diseases which target different parts of the plant body such as leaf, stem, seed, fruit and so on. Diseases are specific to specific parts of the plant body. Leaves can be addressed as the core part of the plant. If the plant leaf is prone to the disease it directly effects the plant life cycle. In order to handle these diseases bravely it is essential to put forward a system which identifies and classifies the disease automatically. To solve this problem machine learning is used. Various machine learning techniques have made way to solve the problems, but the greatest challenge being faces is the accuracy and the robustness of the results obtained. The machine learning and image analysis sector of artificial intelligence promises some benefits to the agriculture community especially when it comes to using mobile devices which are capable of running some artificial intelligence applications both online and offline. When paired with image analysis and deep neural networks, there is promise especially in detection of plant diseases by using leaves and stems and also detecting insects well before damage is done to crops.

II. PROBLEM STATEMENT

The Crop Yield Prediction and E-marketing System is an end user support and online consultation project. Help farmer for managing his expenses and schedule: Using this Website farmer can get approximate budget for their product. Farmers can insert the data such as soil type,

temperature, area from which they can predict which crop is suitable to grow. After the prediction of crop farmer can get information such as how much the product can yield in given condition and also the duration to grow that product. By using these application farmers will also get the exact schedule depending upon the plantation he has chosen for e.g. "Suru Plantation", "Adsali Plantation", "Pre seasonal Plantation". He will also get the schedule for irrigation and fertilization.

Latest Government Policies are easily available. Farmers will get latest government policies related to agriculture. Government policies available on Ministry of Agriculture web site but for farmer always it is not possible to check government policies on site.

Provide new Strategies and Technologies for Farming. This application will provide new strategies in plantation and upcoming technologies for better production of crop.

4.Farmer can also process the disease diagnosis on crops by using ANN algorithm. He will add the input of the images of crops and then the system will extract the images and display the report and suggest predictive measure.

III. LITERATURE SURVEY

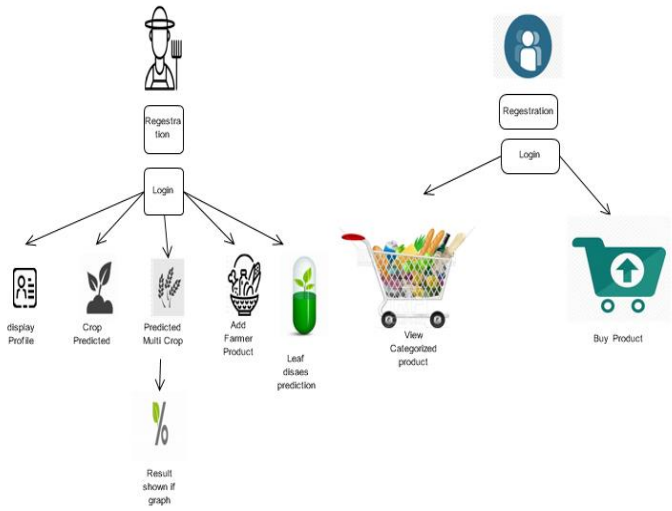
Two points are very essential for the crop to be grown that can be decided mainly on the basis of market and profit. But up till now this is inherited or traditional method used which are not 100 percent assured effectiveness. In any area, the main cropping systems are the combined results of past and present conclusions by individuals, communities or governments and their actions.

There is a lot of diversity in crops in India. In north India, there are two seasons, kharif, it's duration is July to October, and rabi whose duration is October to March. Crops produced between March and June are called as zaid. Can you make a guess why India has a diversity in crops? India has range of topography, climate and soil. Since India has tropical as well as moderate climate, crops of both the climates are found in India. There are very rare countries in the world that have variability as compared to that of India. The major significant features of Indian agriculture are sustainable farming, which is highly dependent on monsoon and animals, variation in crops and predominance of food crops. Major crops in India can be generally divided into four categories i.e. fruits crops, plantation crops, cash crops, and food crops. In Precision Farming disparities are checked, stored for management and upholding of the valuable resources using technologies to achieve and increase cultivation or produce. This can be the apparatus at the hands of agriculturalists for selecting crop, management with objective of refining return on investments while preserving natural resources.

IV METHODOLOGY

In this a user friendly web application is developed which will be useful for farmers as well as customers. Farmers will be providing their information about product and customer will be able to buy that product according to their area. Customer will search for the nearest farmer and according to needs they will place order. This system will provide fair price for the produce and will be beneficial for the farmer. In this system we have provided features like crop prediction, disease diagnosis, government schemes for the farmers.

System Architecture:



A. crop Prediction

Crop prediction provides information to farmers that which alternative crop can be cultivated under particular conditions. Inputs for crop prediction are temperature, soil type and sub type. It will also provide duration required for the growth of crop and how much yield will it give. There are two machine learning algorithm used for crop predictions which are naïve bayes algorithm and apriori algorithm. Naïve bayes algorithm provides a single output where as apriori algorithm provides multiple outputs.

Naïve Bayes algorithm

In this project naïve bayes algorithm is applied for prediction crop as well as its duration to grow and how much quantity will be grown.

One of the easiest ways of selecting the most probable hypothesis given the data that we have that we can use as our prior knowledge about the problem. Bayes' Theorem provides a way that we can calculate the probability of a hypothesis given our prior knowledge.

Bayes' Theorem is stated as:

$$P(h|d) = (P(d|h) * P(h)) / P(d)$$

Where

- **P(h|d)** is the probability of hypothesis h given the data d. This is called the posterior probability.

- **P(d|h)** is the probability of data d given that the hypothesis h was true.
- **P(h)** is the probability of hypothesis h being true (regardless of the data). This is called the prior probability of h.
- **P(d)** is the probability of the data (regardless of the hypothesis).

You can see that we are interested in calculating the posterior probability of P(h|d) from the prior probability p(h) with P(D) and P(d|h).

After calculating the posterior probability for a number of different hypotheses, you can select the hypothesis with the highest probability. This is the maximum probable hypothesis and may formally be called the maximum a posteriori (MAP) hypothesis.
This can be written as:

$$MAP(h) = max(P(h|d))$$

or

$$MAP(h) = max((P(d|h) * P(h)) / P(d))$$

or

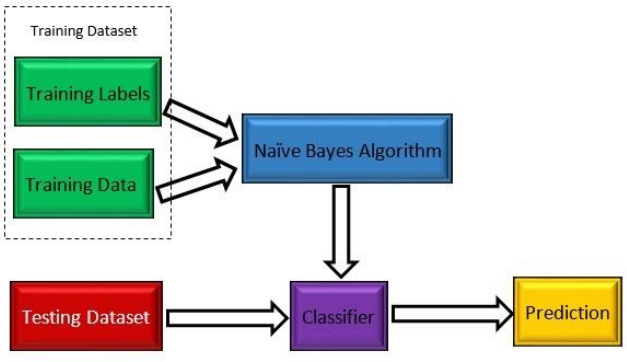
$$MAP(h) = max(P(d|h) * P(h))$$

The P(d) is a normalizing term which allows us to calculate the probability. We can drop it when we are interested in the most probable hypothesis as it is constant and only used to normalize.

Back to classification, if we have an even number of instances in each class in our training data, then the probability of each class (e.g. P(h)) will be equal. Again, this would be a constant term in our equation and we could drop it so that we end up with:

$$MAP(h) = max(P(d|h))$$

This is a useful exercise, because when reading up further on Naive Bayes you may see all of these forms of the theorem.



In our system Naïve Bayes algorithm predicted Using soil type like red soil, rainfall. this algorithm gives result crop name, crop yield and also predicted no of quantity and how much duration is required for crop growth. Using this information farmer is grow their crop according to predicted result. This useful information farmer is used in their farming for better yield and in small duration.

Apriori Algorithm

Three significant components comprise the apriori algorithm. They are as follows.

- Support
- Confidence
- Lift

Frequent Item Sets- A set of attributes is termed as frequent item set if the occurrence of the set within the database is more than a user given threshold.

Support- Support determines how often a given rule is applicable to a given data set. i.e. In our system one type of soil have multiple crops we take soil type temperature, rainfall from user.

Confidence- Confidence determines how frequently items in soil appear in transactions that contain the temperature. Calculate the frequently which crop yields more in that soil type, we calculate the probability of the temperature

SUPPORT, S(temp->soil) = Σ temp U soil/ N

Confidence, C(temp->soil) = Σ (temp U soil) /Σ(temp)

Where, temp and soil temperature and rainfall disjoint item set

Lift for two items, say temp and rainfall. It is because most frequently crops growing with these two conditions.

Support (Rainfall) = (crops with particular rainfall) / (Total crops)

= 200/2000 = 10%

a) Confidence

In our example, Confidence is the likelihood that crops yield s both rainfall and temperature. Dividing the number of transactions that include both rainfall and temperature by the total number of transactions will give the Confidence figure.

Confidence = (Transactions involving both bread and jam) / (Total Transactions involving jam)

= 100/200 = 50%

It implies that 50% of crops which are grown in same temperature and rainfall.

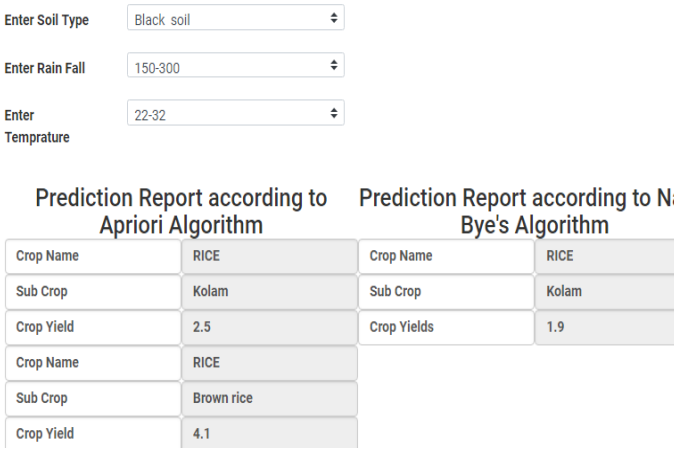
b) Lift

According to our example, Lift is the increase in the ratio of the growing crops when you match the above soil rainfall and temperature. The mathematical formula of Lift is as follows.

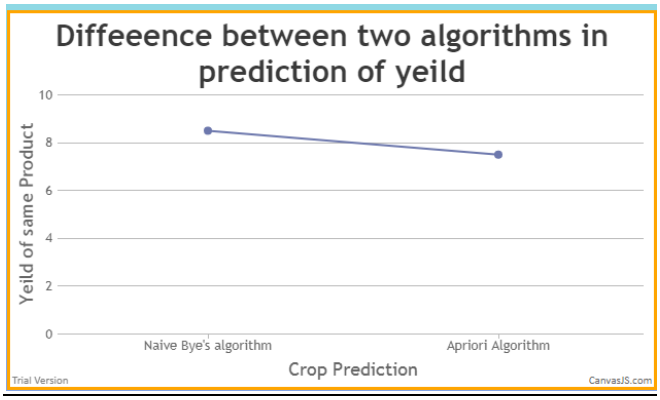
Lift = (Confidence (soil – temperature)) / (Support (rainfall))

= 50 / 10 = 5

Given result is generated multiple crop yield prediction using Apriori Algorithm this algorithm predicted more than one crops. According to the predicted result Farmer can select one or more crops planting ideas.



Given predicted result shown in graphical representation using both algorithms.



Support Vector Machine

In Our System SVM is used for classification of different type of crops/products. Support vector machines for binary or multiclass classification.

B) Disease diagnosis using leaf Images

This is the method which utilizes the techniques of image processing and neural network in a composite manner to obtain the desired goal. The proposed leaf disease identification and control prediction algorithm consist of the following steps:

A. Step 1: Image Acquisition

leaf images are captured from different regions by using digital mobile camera, Xiomi Redmi 1S, 8 Megapixel and are used for training and testing the system then the background data are removed and stored in standard jpg format.

B. Step 2: Image Pre-Processing

Image pre-processing includes the following three modules: – Cropping leaf image. – Resize. – Median filter.

C. Step 3: Image Conversion

The image conversion includes the following types of conversion for different purposes: – RGB to gray. – Gray to binary. – RGB to L*a*b* color shape.

D. Step 4: Segmentation

To extract the ROI in diseased mango, leaf the K-means clustering algorithm is used. This algorithm clusters the point nearest to the centroid. The centroid is basically the average of all the points in that cluster and has coordinate as the arithmetic mean over all points in the cluster, separately for each dimension.

E. Step 5: Feature Extraction

The following features are extracted to classify the disease: 1) Area: The actual number of pixels in the region of interest. Identify applicable funding agency here. If none, delete this text box.

2) Orientation: The angle θ (in degrees ranging from - 90 to 90 degrees) between the x-axis and the major axis of the ellipse that has the same second- moments as the region. Θ=arc (tan(√( )) (1)

3) EquivDiameter: It specifies the diameter of a circle with the same area as the region. Computed as: EquivDia=sqrt( ) (2)

4) Extent: It specifies the ratio of pixels in the region to pixels in the total bounding box. Computed as: Extent= (3)

5) Solidity: It specifies the proportion of the pixels in the convex hull that are also in the region and computed as: Solidity= (4)

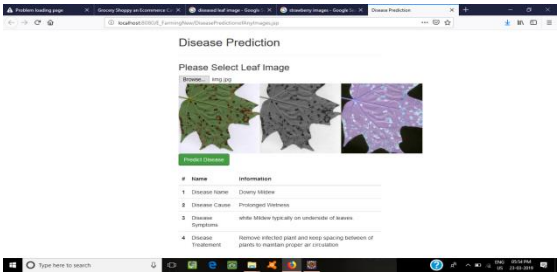
6) ConvexArea: It specifies the number of pixels in 'Convex Image'.

7) MarorAxisLength: It specifies the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the region.
8) Number of Objects: It is the number of white pixels which are disconnected to each other in binary image.

F. Step 6: Classification

For the classification the feed forward Back Propagation Neural Network classifier technique is used which consist of three layers namely input layer, a hidden layer, and an output layer.

G. Step 7: Disease Identification and Control Prediction



The BPNN assigns an appropriate mango leaf disease class i.e. red rust or bacterial leaf spot. Then it appropriate control prediction for the bacterial leaf spot or red rust is also given by the system automatically.

V .FUTURE SCOPE

The system can be enhanced further to add the following functionality:

1. Implemantation of smart irrigation system monitor soil and weather conditions ,water management ,e.t.c to automatically alter the watering schedule.

2. To ensure the organic authenticity of farm products that is the crops and vegetables which the customer buys by determining the ENOR (Economic Nitrogen Organic Rate) in each of the farm products.

VI CONCLUSION

The developed web application uses supervised and unsupervised machine learning algorithms and gives the best result based on accuracy and precision. The results of the two different algorithms used will be compared and the one giving the best, accurate and optimal output for suggesting the crop type as well as the crop yield will be selected. Thus the developed web application will help to reduce the difficulty faced by the farmers and in turn this will help in the reduction of their suicide rates. It will act as a medium to provide farmers with efficient and required information to get high yield and productivity and thus help to maximize he profits.The present study demonstrated the potential use of data mining techniques in predicting the crop yield based on the climatic input parameters. The developed webpage is user friendly and the accuracy of predictions are above 75 per cent in all the crops and districts selected in the study indicating higher accuracy of prediction. The user friendly web page developed for predicting crop yield can be used by any user their choice of crop by providing climatic data of that place. The paper speaks about the project in which we took the idea that will make every farmer reach the homes in there nearby locality or cities by the medium of this web application. In this we have used some simple database and used a reference algorithm for displaying the images on the left side termed as related product in the purchase product. We have implemented the Customer login, Farmer Login as additional features to the system making system more users friendly. By the help of this application people will be able to get fresh food to eat and will be able to explore parts of their nearby villages for picking up their purchases and exploring the place establishing relation with farmers and gaining profit by saving their money, adding profit directly to the farmer helping farmers too.

ACKNOWLEDGMENT

Inspiration, Motivation and perseverance have always played a key role in the success of my Venture. Success is never achieved single handedly, so it's my duty to acknowledge all those who have provided timely guidance and helping hands in making my project successful.

Firstly, I would like to express gratitude to my project guide **Dr. T. A. Chavan** for his invaluable guidance that he has extended to me all over the project work. Without his guidance and help this project would have been an uphill task for me. He has taken keen interest in going through my analysis and designing part correcting me whenever I made mistakes, also encouraging in doing something new and extra.
I wish to convey deep sense of gratitude to **R. H. Borhade** of Information Technology Smt. Kashibai Navale College of Engineering, Pune-41 for giving me an opportunity to carry out the project work.

.

REFERENCES

[1] Juhi Reashma, Anitha S.Pillai"Edaphic factors and crop growth using Machine learning – A Review"Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS 2017) IEEE Xplore Compliant - Part Number:CFP17M19-ART, ISBN:978-1-5386-1959-9

[2] Hemantkumar Wani,Nilima Ashtankar "An Appropriate Model Predicting Pest/Diseases of Crops Using Machine Learning Algorithms" 2017 International Conference on

Advanced Computing and Communication Systems (ICACCS -2017), Jan. 06 – 07, 2017, Coimbatore, INDIA —

[3] Sk Al Zaminur Rahman, Kaushik Chandra Mitra" Soil Classification using Machine Learning Methods and Crop Suggestion Based on Soil Series" 2018 21st International Conference of Computer and Information Technology (ICCIT), 21-23 December, 2018

[4] Boikobo Tlhobogang and Muhammad Wannous," Design of Plant Disease Detection System: A Transfer Learning Approach Work in Progress", Proceedings of IEEE International Conference on Applied System Innovation 2018 IEEE ICASI 2018- Meen, Prior & Lam (Eds)

[5] Neha Khadke, Divya More," Agriculture Decision Support System using Data Mining", 2017 International Conference on Intelligent Computing and Control (I2C2)

[6] Konlakorn Wongpatikaseree, Promprasit Kanka,Arunee Ratikan," Developing Smart Farm and Traceability System for Agricultural Products using IoT Technology", 978-1-5386-5892-5/18/$31.00 ©2018 IEEE ICIS 2018, June 6-8, 2018, Singapore

[7] Shikha Prakash, Animesh Sharma," SOIL MOISTURE PREDICTION USING MACHINE LEARNING", and Communications (ICTACC), 2017 International Conference on IEEE, 2017.

[8] Subhadra Mishra, Debahuti Mishra and Gour Hari Santra," Applications of Machine Learning Techniques in Agricultural Crop Production: A Review Paper", Indian Journal of Science and Technology, Vol 9(38), DOI: 10.17485/ijst/2016/v9i38/95032, October 2016

[9] Zhisheng Qin,D. Brenton Myers, Curtis J. Ransom ." Application of Machine Learning Methodologies for Predicting Corn Economic Optimal Nitrogen Rate", Agronomy Journal • Volume 110, Issue 6 • 2018

[10Godliver Owomugisha, Friedrich Melchert, Ernest Mwebaze," Machine Learning for diagnosis of disease in plants using spectral data" Int'l Conf. Artificial Intelligence | ICAI'18

[11] Aarti Rao Jaladi, Karishma Khithani, Pankaja Pawar, "Environmental Monitoring Using Wireless Sensor Networks(WSN) based on IOT", International Research Journal of Engineering and Technology (IRJT) Volume: 04 Issue:01Jan-2017

[12] Rushika Ghadge, Juilee Kulkarni, Pooja More, Sachee Nene, Priya R L.," Prediction of Crop Yield using Machine Learning", International Research Journal of Engineering and Technology (IRJET)e-ISSN: 2395-0056 Volume: 05 Issue: 02 | Feb-2018

[13] http://en.banglapedia.org/index.php?title=Soil_Series. Last Accessed date: 25th August, 2018