Indian Journal of Agriculture Engineering (IJAE)
ISSN: 2582-9319 (Online), Volume-1 Issue-1, May 2021

Agricultural Commodities Price Prediction System using ML Algorithm

Vinutha M, Gayitri H.M

Abstract: ***Most of the India's population depends on agriculture for their source of income. But, as and when it evolved, the complexities also grew with it, which made the farmers to struggle. Few major difficulties being faced by the marginal and small farmers at present are, high investment cost, unpaid loans, lack of basic awareness about agriculture, marketing issues, drought, etc. So, the proposed work gives solution to few of these issues through a web application. It firstly gives an "Agricultural Guide" in the form of a PDF file that contains overall information starting from farm land details to marketing their cultivated product and secondly, it includes a crop price prediction system that uses the Naive Bayes algorithm. The crop price prediction system is specifically developed for Mysuru District. The dataset containing Crop name, Rainfall, Yield, Max Trade, Crop Price is used as its attributes. The web app contains agricultural departments as one of its users, that is the different regulated agricultural markets present in Mysuru, the staff of any particular department can upload the training dataset and can predict the price. The farmers can know the predicted price by contacting the respective agricultural departments. This application is developed using Visual Studio IDE and SQL Server Management Studio Express .The application is bilingual with Kannada and English languages and the Android version of the same application is provided for faster access. Hence the application helps to reduce some of the problems being faced by the farmers.***

Keywords: Agriculture, Naive Bayes, Price Prediction, Web Application.

I. INTRODUCTION

According to 2011 Census, 56% people of Karnataka depend on agriculture for their livelihood. But, small and marginal farmers are struggling very hard to lead a good life. Also, many farmers have committed suicide and are facing various problems like high costs of agricultural inputs such as seeds, fertilizers and equipments, high amount of unpaid loans, drought or water crisis, lack of awareness regarding services being provided by the government, lack of direct integration with the market and exploitation by the middleman etc. Though Government is taking measures to solve these problems, the reality is still lagging behind. The proposed work aims at providing solution to few of these issues by developing a web application.

Firstly, the application provides awareness to farmers about basic and important agricultural aspects in the form of a guide which is provided in PDF file format. It gives information regarding almost all aspects of agriculture like choosing the agricultural land, Irrigation systems, Climatic conditions for agriculture, Seeds, Fertilizers, Agriculture Marketing and few others. So, this guide helps a farmer from sowing a seed till marketing their produce. Secondly, it provides a crop price prediction facility to the farmers of Mysuru. As agricultural statistics and data varies from one region to other, Mysuru region is specifically considered in the application. The prediction system uses Naive Bayes, a Classification algorithm for predicting the price of given crop. Prediction is made by considering previous datasets with Crop name, Rainfall, Yield, Max Trade and Crop Price used as its attributes. Visual studio is used for developing the application. SQL Server Management Studio Express is used as database for storing account details of the administrator and agricultural departments.

The Website administrator, Agricultural Departments and farmers are the users of this application. The administrator manages the Agricultural Departments by giving them the initial credentials. The Agricultural Departments are the different regulated agricultural markets present in Mysuru, the staff of any particular department can upload the training dataset and predicts the price of the required crop. The farmer can access the "agricultural guide" in the application. The application is bilingual with Kannada and English languages and the Android version of the same application is provided for faster access. So, the application helps the farmers by providing the above explained facilities.

II. RELATED WORK

In this survey 2, the author finds the impact of mobile phone in improving the agricultural productivity. The author selects Delhi, Maharashtra, Rajasthan and Uttar Pradesh states for the survey and chooses 187 farmers in total. The survey includes impact on fishermen and brokers, but only the impact of mobile phones on farmers are discussed in the present context. By this research, the author have found that almost all the farmers who were interviewed reported some convenience and cost savings by using mobile-phones as basic communication devices to seek information such as input availability or to check on market prices.

Manuscript received on 24 April 2021 | Revised Manuscript received on 02 May 2021 | Manuscript Accepted on 15 May 2021 | Manuscript published on 30 May 2021
* Correspondence Author
Vinutha M*, Student, Agricultural University, Berasia, Madhya Pradesh
Gayitri H.M, Agricultural University, Berasia, Madhya Pradesh

But few differences in the mobile usage was found in Maharashtra and Uttar Pradesh, the farmers of Maharashtra reported that they had to make higher use of mobile phones to access information in general and also reported benefits like yield improvements, price realisation and increased revenues through better adjustment of supply to market demand. In contrast, farmers in Uttar Pradesh found their benefit in improvements in yields only, through the mobile usage. So, the overall summary of this survey is that most of the farmers using Mobile phones have experienced benefits in one or the other way in the agriculture sector.

The survey 7 aims to find the factors that affect the likelihood of adoption of different agriculture-related information sources by the farmers. The author surveys some 240 households from 4 different states of India and got around 17 information sources that were being used by different farmers. These sources were divided into 4 categories based on their common characteristics and the four being Face to face sources like Krishi Vigyan Kendra's, Research stations and few more; Other farmers in the same village; Traditional media like Television, Radio, Newspaper; Modern ICT Landline phone, Mobile phones, Internet. The author by analysing these data and statistics from the survey finds the factors such as age, education level of the farmer, farm size, access to household information assets and geographical disparity in each state, affect the adoption of a information source by a farmer. This survey helps in narrowing the information source to a farmer based on the different factors mentioned above.

In the next paper [8], the author does a yield prediction of rice crop in semi-arid Climatic Zones of India using different Data Mining techniques. Different Data Mining algorithms are used and the performance of each algorithm is calculated. The algorithms considered are J48, LADTree, IBk and LWL. Sensitivity, specificity, accuracy were calculated to validate the experimental results. F1 score was calculated to measure the accuracy, Mathews Correlation coefficient was calculated to measure the quality of the classification and many other standard values were calculated for the purpose of calculating the performance of different classifiers. Some ninety seven districts from various states of India are selected for this study. The parameters considered for the prediction are precipitation, minimum average and maximum temperature, soil type, area and production. The study used an open source tool called WEKA, for executing different algorithms considered. The results of the study showed that J48 and LADTree were achieved more accuracy, sensitivity and specificity, IBk and LWL showed lowest accuracy. So, this study explains clearly the performance of different classifiers in predicting the rice crop yield in the selected climatic zone and gives insights to the way in which different parameters must be used in a specific study.

III. OBJECTIVES

The objectives of the proposed system are:

1. The system provides "Agricultural Guide" in PDF format.
2. A crop price prediction facility is provided by the application.
3. The application is bilingual with English, Kannada languages.
4. The system is also provided as Android app.

IV. PROPOSED SYSTEM

The proposed model is a web application. There are three users of this application. Namely, the Website administrator, Staffs of Agricultural Departments and farmers. The administrator manages the Agricultural Departments by giving them the initial credentials. The Agricultural Department's staff can upload the previous agricultural datasets containing Crop name, Rainfall, Yield, Max Trade, Crop Price as its attributes and can calculate the predicted price of a particular crop. The staff needs to input the present rainfall, yield, max trade values to predict the price of a system. The farmer can access the "agricultural guide" in the website and needs to contact the respective agricultural department for knowing the predicted price of any crop needed.

The application is implemented using the standard 3-Tier architecture. In the Presentation Layer, the GUI interface is provided to the users of the application. HTML5, CSS and Javascript are used for implementing the web pages of the GUI that is presented to the users. Home Page, Manual Page, Contact us and Login Pages are the different web pages implemented in the application. The Login page in turn contains separate Home pages and its sub pages for Administrator and the "Agricultural departments". The second layer is the Business Logic Layer which is the implementation of the logic behind the application, the Naïve Bayes algorithm for crop price prediction is implemented in this layer using C# language. The third layer is the Data Layer which provides access for storing data to and retrieving data from the database. The account details of the administrator and the agricultural departments are stored in the SQL database and the agricultural dataset is stored in excel sheet. The Presentation and Business Logic Layers are implemented using Visual Studio tool and SQL Server Management Studio Express is used as a tool for Data layer.

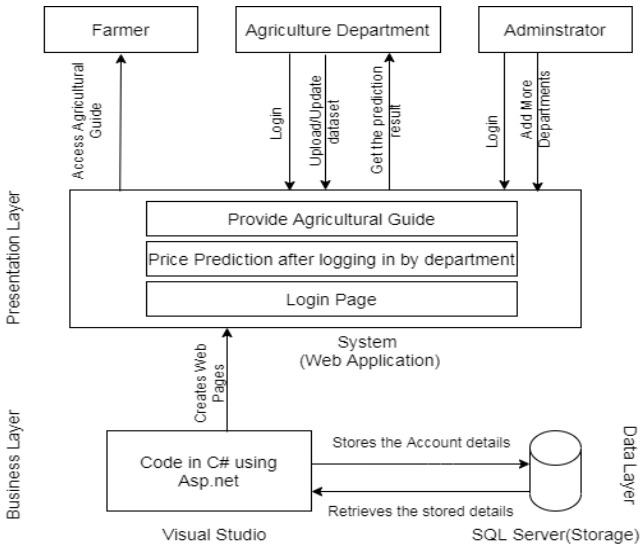


Fig : System Architecture

V. METHODOLOGY

The project implementation follows below steps.

i) Preparing the dataset :

Different data in the dataset are obtained from different and official Government's website and it is specifically taken for Mysuru region. Crop name, Rainfall, Yield, Max Trade, Price are the attributes of the dataset. The dataset is stored in excel format.

ii) Creating the database and data tables:

The account details of Administrator and Agricultural departments are stored in the SQL database Tables. So, firstly a database with name User Credentials is created under which a table with name "tblUsers" is created. The table contains 4 columns, UserID, Password, UserType, Location. The Usertype column determines whether the user is an admin or an agricultural staff. The location column is used for storing different locations of different agricultural markets.

iii) Creating the Visual Studio Project and importing the required files to the project:

Before creating a web project in Visual Studio, firstly the .Net and ASP.Net frameworks are installed using the installer. Then, "ASP.Net Web Application" project type is selected and the same is created. After the project creation, the required CSS, Jquery and JS files are imported to the project. Jquery library, Flora plugin and UI Jquery files are imported to the project along with the easing.js, move-top.js and cloud-zoom javascripts. Each file is used for fulfilling different requirements.

iv) Connecting data sources such as SQL Server Management Studio Express, MS Excel with the web application:

For allowing access to the dataset's excel sheet, the details such as data provider name and data source name should be mentioned in Visual Studio's web.config file and in case of SQL Server Management Studio Express as data source, the User ID and password of the server, with database and table name must be provided in web.config file.

v) Implementing the Web pages and the Business Logic:

There are totally 3 web pages in the website they are Home page, Manual page, Contact page and Login Pages. There are separate home pages and its sub pages for the admin and agriculture department's staff once they log in. The price prediction system is provided in staff's web page and it is implemented as below.

Naive Bayes is a Supervised Machine Learning technique that work on a idea that, the presence of one feature is not related to the presence of another feature in a dataset. The formula of the algorithm is as follows.

$$V_{NB} = \text{argmax}_{v_j \in V} P(v_j | a_1, a_2, a_3, \dots, a_n) \quad \rightarrow \text{Equation 1}$$

Where,

- V_{NB} is the final predicted or classified result.
- $P(v_j|a_1, a_2, a_3, \dots, a_n)$ is the posterior probability of any jth target, given the attribute or feature, that is from attribute a1 to an. In the present problem scenario, vj is the jth unique previous price of a given crop and "ai" comprises the 4 parameters such as Yield, Max Trade, Minimum Support Value (MSP) and Rainfall of a given crop.
- $argmax_{v_j \in V}$ of the probability is taken to obtain highest probability value and in turn to get more accurate result.

From Bayes Theorem,

$$P(v_j|a_i) = \frac{P(a_i|v_j)P(v_j)}{P(a_i)}$$

Where,

- $P(a_i|v_j)$ is the conditional probability of each attribute "ai", given the target value vj. In the present problem scenario, "ai" comprises the 4 parameters Yield, Max Trade, Minimum Support Value (MSP) and Rainfall of a given crop and vj is the previous prices of a given crop.
- $P(v_j)$ is the probability of target being equal to one of its value. In the present problem scenario, it is the price of a crop being equal to the inputted price.
- $P(a_i)$ is the probability of an attribute's value being equal to the inputted attribute value.

Replacing $P(v_j|a_i)$ with its value in equation 1, we get,

$$V_{NB} = argmax_{v_j \in V} \frac{P(a_1,a_2,a_3,\dots,a_n|v_j)P(v_j)}{P(a_i)} \quad \rightarrow \text{Equation2}$$

Naive Bayes assumption:

$$P(a_1, a_2, a_3, \dots, a_n|v_j) = \pi i\, P(a_i|v_j)$$

Hence, replacing $P(a_1, a_2, a_3, \dots, a_n|v_j)$ with its above value in equation 2, we get the formula as follows,

$$V_{NB} = argmax_{v_j \in V} P(v_j)\, \pi i\, P(a_i|v_j)$$

The final formula for calculating $P(a_i|v_j)$ is as below,

$$P(a_i|v_j) = \frac{n_c + (mp)}{n + m}$$

Where,

"nc" is the number of examples for which v =v_j and a = a_i.

"m" is the number of attributes considered. In the present problem scenario m=4 and the corresponding attributes are Yield, Max Trade, Minimum Support Price(MSP), Rainfall.

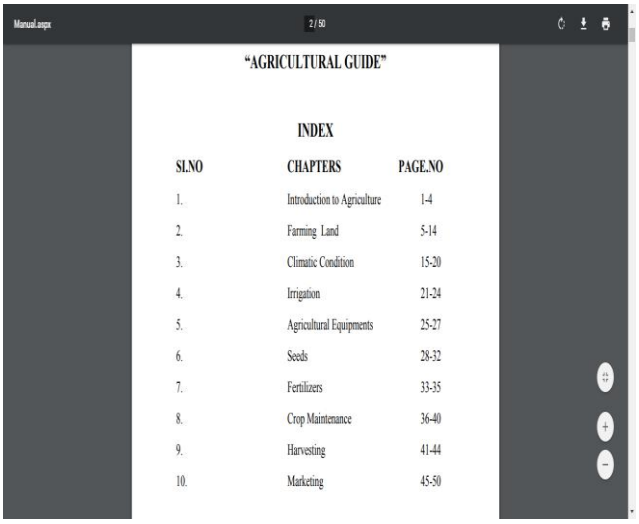
"p" is the prior estimate of $P(a_i|v_j)$.

"n" is the number of training examples for which v = vj.

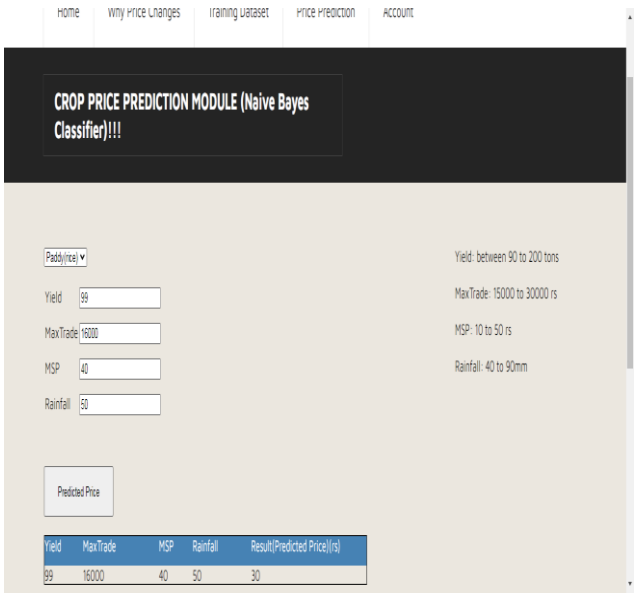
VI. RESULT AND DISCUSSION

The results are presented in this section. Below screenshot shows the contents of the "Agricultural Guide" file. As mentioned in the introduction, the guide covers all the information on all stages of agriculture.

Published By:
Lattice Science Publication



The predicted price of the paddy crop is being shown in the below screenshot. As shown below, each parameter in the dataset, that is the yield, Max Trade, MSP and Rainfall values must be within some range. This range is fixed based on the minimum and maximum values of each parameter in the dataset.



VII. CONCLUSION

As most of the population in the country are farmers, it is important to ensure their well being. Hence a web-based application is developed in the present work that helps the farmers to know the basic, important information of agriculture through a guide in a readable format. The application also provides a crop price prediction system, to help farmers predict the price of the crop before they sell so as to make their marketing strategies according to their profit. Also these implementations can also be used through the android app. So this application helps the farmers at least to some extent and hence in increasing their productivity.

REFERENCES

1. Mittal, Surabhi and Praduman Kumar (2000), "Literacy, technology adoption, factor demand and productivity: An econometric analysis", Indian Journal of Agricultural Economics, Vol. 55 No. 3, pp. 490-499.
2. Surabhi Mittal, Sanjay Gandhi, Gaurav Tripathi (2010), "Socio-Economic Impact of Mobile Phones on Indian Agriculture", Indian Council for Research on International Economic Relations, Working Paper No. 246. Available at: http://hdl.handle.net/10419/176264.
3. Jenkins, A., M. Velandia, D. M. Lambert, R. K. Roberts, J. A. Larson, B. C. English, and S. W. Martin (2011),"Factors Influencing the Selection of Precision Farming Information Sources by Cotton Producers", Agricultural and Resource Economics Review, Vol.40 No. 2, pp. 307-320.
4. A.Mankar, M Burange (2014), "Data Mining- An Evolutionary view of Agriculture", International Journal of Application or Innovation in Engineering & Management, Vol 3. No 3, pp. 102-105.
5. R. Kalpana, N. Shanthi, S. Arumugam (2014), "A survey on data mining techniques in agriculture", International Journal of Advances in Computer Science and Technology, Vol. 3 No. 8, pp. 426-431.
6. H. Patel, D. Patel (2014), "A brief survey of data mining techniques applied to agricultural data", International Journal of Computer Applications, Vol. 95 No. 9, pp. 6-8.
7. Mittal S and Mehar M (2015), " Socio-economic factors affecting adoption of modern information and communication technology by farmers in India: analysis using multivariate probit model", The Journal of Agricultural Education and Extension Vol. 22 No. 2, pp 199–212.
8. Gandhi, N., Armstrong, L. J. & Nandawadekar, M (2017), "Application of data mining techniques for predicting rice crop yield in semi-arid climatic zone of India", IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR).

Published By:
Lattice Science Publication