

Agriculture Commodity Price Forecasting using ML Techniques

Varun R, Neema N., Sahana H. P., Sathvik A., Mohammed Muddasir

Abstract: India is mainly an agricultural country the farmer is an important part of agriculture. Agriculture mainly depends on him. Even then the farmers cannot predict prices for their commodities because prediction of prices plays a major challenge. Several characteristics are taken into account so that the crop price forecast is accurate. We consider the attributes of the Mysore region to make it a real-time application framework. Price prediction is a big issue for farmers who are not aware of the market prices. Forecasting price of agriculture commodities helps the agriculturist and also the agriculture department of mysore region to make decisions. The new model predicts the accuracy for the agricultural yields and it also avoids the role of middle man.

Keywords : Price Prediction, Data Mining, Naïve Bayesian Classifier, k-means, Artificial Neural Networks, Support Vector Machine, Prediction, Extended Kalman filter, Wavelet, Error Analysis;

I. INTRODUCTION

The agriculture involves cultivation of crops and live stock commodities. Even though large areas in India are of irrigational land, the survey says that only 1/3rd of cropped part is irrigated. But productivity is extremely low [1]. Some of the major problems our farmers facing today are yield prediction, price prediction and profit prediction. The yield primarily depends on weather conditions such as rainfall, humidity, ph, moisture, wind and etc. All these factors influence the yield of each and every crop that are cultivated. Using Data mining techniques, we can bring changes in agriculture development. The volume of data that has been found under Indian Agriculture is vast enough. Once if these data becomes information, it would be highly useful for many purposes. The mining of descriptive data and statistical data can be of 2 different types. In Machine learning functions have cumulative properties of records where computer helps to get the value considering the proven evidences or of the previous data result patterns [3]. Yield forecasting is a major issue in agriculture. Any farmer wants to know how much harvest he expects. In the data mining procedure, training data is obtained from previous market price, and the collected data is used from a training perspective that has been leveraged [10].

Revised Manuscript Received on December 07, 2019.

Varun R, Dept. of Information Science and Engineering, VVCE Mysuru. Email: varunramson@gmail.com

Neema N, Dept. of Information Science and Engineering, VVCE Mysuru. Email: neeman629@gmail.com

Sahana H P, Dept. of Information Science and Engineering, VVCE Mysuru. Email: hpsahana98@gmail.com

Sathvik A, Dept. of Information Science and Engineering, VVCE Mysuru. Email: asathvik.gowda@gmail.com

Mohammed Muddasir, Dept. of Information Science and Engineering, VVCE Mysuru. Email: mohammed.muddasir@vvce.ac.in

II. DATA MINING TECHNIQUES

This paper present the data mining technique in agricultural domain. In order to classify the unknown samples the classification techniques used are Neural Networks and Support Vector Machines[4]. Artificial Neural Network have been demonstrated to be powerful tool for modeling and prediction, to increase their effectiveness. Naïve Bayesian Classifiers are a collection of classification algorithm based on Bayes' Theorem.

The Kalman filter attempts to understand the general problem of estimating the status of a system under continuous time control with linear differential function n RX.

In this fundamental analysis, the first part is some economic and financial parameters used. They are placed in the system model after these variables are moved to the structure and changed to prices. The second part is Mar's general trend [11]. The BP-neural network system used to forecast the value on a weekly or monthly basis demonstrates the proportion of errors produced by back propagation using neural network [12].

The well-organized guidelines for processing of information are one of the most effective strategies of finding the right model for the large amount of data. The different machine learning algorithms include Apriori's Algorithm, Partition, Hashing, Hemset Counting, FP Growth, SEAR, Spear, Eclat&Declat, MaxEclat [5].

A statistical method called the extended Kalman filter is used in short-term forecasting. The value of the grain market is projected by using the wavelet forecast system [13]. The expanded process of Kalman filter information fusion makes the combining of numerous measures used in basic analysis such as economic factors with technical analysis and general trends of shift [11].

Regression is one such data mining technique where learning acts as a method to get the price of the crop. The regression task will be considered as classification tasks and task with defined class labels [6].

The information contained is minimal when it comes to agriculture on the basis of the available date. There are two techniques classifying and clustering. Classifying is known some sample of data where we get to know about the unknown samples. A collection of unknown samples which can be broken into cluster is known as clustering. The different application of K-means and K-nearest algorithm are finding the changes in weather or the amount of pollution. Neural networks help in mapping data under machine learning. The information will be split into the set of training examples, validation and a test data [7].

III. RESULTS AND DISCUSSION

System makes use of “Data Science” algorithms for price prediction. Price forecasting framework is a farm-oriented software that is an effective tool for forecasting prices because it is a software in real time designed for the departments of agriculture and peasants which is an algorithm for the estimation of agricultural commodity price. System predicts price based on the attributes such as date, yield, max trade, min trade, rain fall, speed of wind, humidity, max & min temperature, cloud coverage, pesticides, harvesting, fertilizers, soil and sunlight. System uses “Visual Studio” as front-end technology and SQL Server as back-end technology. Some of the previous year data(MSP) along with its production cost is shown in fig.1.

Table- I: Msp vs production cost (rs/quintal) [3]

Crop	Production cost		% Rise	MSP		%Rise
	2017-18	2018-19		2017-18	2018-19	
Wheat	817	866	6.00	1735	1840	6.05
Barley	844	860	1.78	1410	1440	2.13
Chana	2461	2637	7.15	4400	4620	5.00
Masur	2366	2532	7.02	4250	4475	5.29
Mustard	2123	2212	4.19	4000	4200	5.00
Sunflower	3125	3294	5.41	4100	4945	20.61

In the above table we can view the variations in price of the crops. The price may vary due to the various conditions such as weather, soil, production cost, pesticides and so on. The invested amount by farmer may not be obtained in every year.

In our website we can make the farmers to enter the details of the commodities and prices the customer can view the details and buy here agriculture department can manage details and based on the predicted value they can fix the amount of the particular crop.

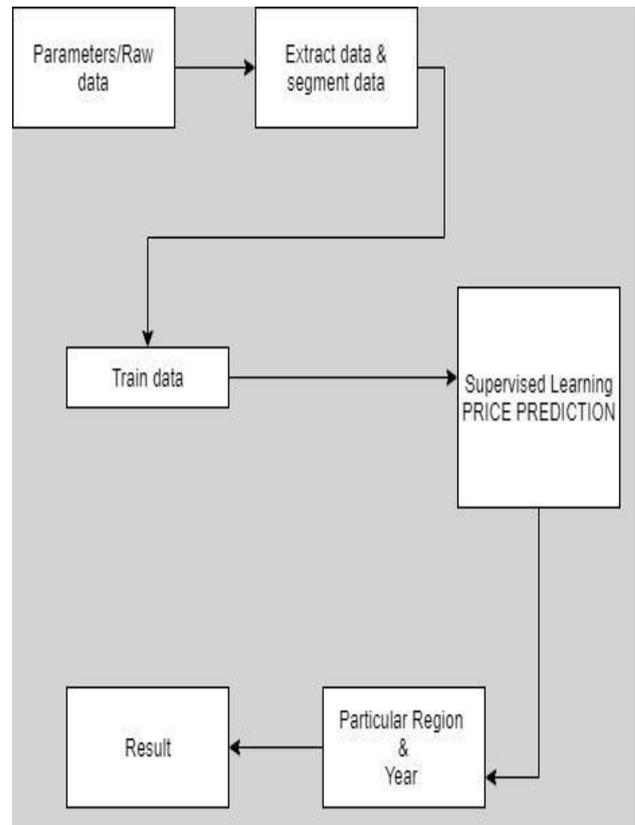


Fig.1 Working diagram of price prediction

Supervised Learning

Naïve Bayesian Algorithm Steps

Step1: Take the Raw data and extract the data which is required

Step 2: We need to get the probability of all the entities. [ex: sunny, windy, soil]

Step 3: Now Applying the below formula

$$P = [n_c + (m*p)] / (n + m)$$

Where:

- n_c = Count of the parameter when the price is same
- m = Number of parameters taken
- p = probability obtained
- n = Total number of prices
- P = price estimated

Step 4: The result obtained will be assigned to a particular region and for a particular year

Sample Example (Paddy)

Parameters – P1, P2, P3 [m=3]

Outcomes – 50rs, 80rs [p=1/2=0.5]

One can understand how the previous dataset is organised in the below table according to the set of parameters used. Based on the price and other factors the data has been classified and now one has to apply the below formula to get the accurate price.

Table- II: The training dataset of few data

Data	P1 (X,Y,Z)	P2(A,B,C)	P3(P,Q,R)	Prize
1	X	A	P	50rs
2	X	B	Q	50rs
3	Y	B	P	80rs
4	Z	A	R	50rs
5	Z	C	R	80rs

New Record (next year - 2019)

Features – P1- X, P2 – A, P3 - R

Crop Prize– 50 rs / 80 rs?

$$P = [n_c + (m*p)] / (n + m)$$

Table- III: Values obtained for 50rs and 80rs

50rs	80rs
<p>X</p> $P = [n_c + (m*p)] / (n+m)$ <p>n=2, n_c=2,m=3,p=0.5</p> $p = [2+(3*0.5)] / (2+3)$ <p>p=0.7</p>	<p>X</p> $P = [n_c + (m*p)] / (n+m)$ <p>n=2, n_c=0,m=3,p=0.5</p> $p = [0+(3*0.5)] / (2+3)$ <p>p=0.3</p>
<p>A</p> $P = [n_c + (m*p)] / (n+m)$ <p>n=2, n_c=2,m=3,p=0.5</p> $p = [2+(3*0.5)] / (2+3)$ <p>p=0.7</p>	<p>A</p> $P = [n_c + (m*p)] / (n+m)$ <p>n=2, n_c=0,m=3,p=0.5</p> $p = [2+(3*0.5)] / (2+3)$ <p>p=0.3</p>
<p>R</p> $P = [n_c + (m*p)] / (n+m)$ <p>n=2, n_c=1,m=3,p=0.5</p> $p = [1+(3*0.5)] / (2+3)$ <p>p=0.5</p>	<p>R</p> $P = [n_c + (m*p)] / (n+m)$ <p>n=2, n_c=1,m=3,p=0.5</p> $p = [1+(3*0.5)] / (2+3)$ <p>p=0.5</p>

Since 50rs > 80rs

So this new 2019 is classified to 50rs

Farmers don't get the sum they need. They are facing many losses for different reasons. To increase their gain the previous product value is inserted in the process provides the expected sum so that the farmer can use this information when increasing the plant. This system takes previous information on average and forecasts the current market price.

Previously used Formula

x1, x2, x3 are the previous price
Average price = $x1+x2+x3, \dots + xn/n$

ANNs cannot give a clear picture equal to many statistical model it cannot be a black box [8].

IV. RELATED WORK

This paper attempts to predict the prices of the crops based on the previous year data sets. The analysis would help the farmers, the customers.

In this report, we attempted, on the basis of the new fusion method called the Kalman extended filter, to prediction inventory price of the stock market in the short term and allow simple inventory price analysis to be used for technical analyzes. They have mentioned the various technologies that can be used in agriculture. The algorithms were related [2]. Linear and non-linear projections aid to forecast the value and also to increase accuracy. According with ARIMA, BP-neural systems are much stronger [14]. Prices of commodities have long been an important factor in accounting for global uncertainty. The prediction of increases in commodity prices is an important factor for future policymakers. With through the use of the future market, the issue of how much details these contracts provide about future spot price changes has been posed. Analysis of errors shows that the wavelet is very useful and effective [13].

Agriculture stands to be the most important use and use of knowledge for agriculture will change the process of decision making [9].

The price is influenced by different variables like its characteristics, demand, seasonal trends, other commodities price offers from numerous suppliers. Price prediction can be formulated as a regression task. Regression analysis lets researchers determine how much these predictors influence a target variable. In regression, a target variable is always numeric.

V. CONCLUSION

Our model allows the farmer to anticipate the best yield in terms of quality and quantity. The farmers also get the desired or estimated price for their yield. In our proposed model there is no role of middle man and therefore there can be no loss for both the farmers and customers. The farmer can directly sell his agricultural commodities through e-commerce website;

the customer can login to that website and purchase those commodities according to his needs. The Agriculture Department can help the farmers in uploading his agricultural commodities to the e-commerce website.

REFERENCES

1. M. C. S. Geetha, "A Survey on Data Mining Techniques in Agriculture", International Journal of Innovative Research in Computer and Communication Engineering., vol. 3,no. 2, 2015
2. E. Manjula, S. Djodiltachoumy, "Analysis of Data Mining Techniques for Agriculture Data", *Int. J. Comput. Sci. Eng. Comm.*, vol. 4, no. 2, pp. 1311–1313, 2016
3. Aman Vohra, Nitin Pandey,S.K. Khatri, " Decision Making Support System for Prediction of Prices in Agricultural Commodity", in 2019 IEEE Available:<https://farmlead.com/blog/graincents/oct-2018-chickpeas-harvest-smaller-india/>. [Accessed: 28-Oct-2019].
4. Manpreet Kaur, Heena Gulati, Harish Kundra, "Data Mining in Agriculture on Crop Price Prediction: Techniques and



- Applications”, *Int. J. Comput. Appl.*, vol. 99, no. 12, pp. 0975-8887, 2011.
5. Neha Sharma, “A Review on Yield Prediction of Various Techniques and Features”, *Int. J. Scientific Research & Eng. Trends.*, vol 5, no. 3, pp. 2395-556X, 2019.
 6. Hatel Patel, Dharmendra Patel, “A Brief survey of Data Mining Techniques Applied to Agricultural Data”, *Int. J. Comput. Appl.*, vol. 95, no. 9, pp. 0975-8887, 2014.
 7. Georg Rub, Rudolf Kruse, Martin Schneider, Peter Wagner, “Data Mining with Neural Networks for Wheat Yield Prediction”, Springer-Verlag Berlin Heidelberg 2008, pp. 47-56,2008.
 8. Rama Krishna Singh, Prajneshu, “Artificial Neural Network Methodology for Modelling and Forecasting Maize Crop Yield”, *Agriculture Economics Research Review.*, vol. 21, pp. 5-10, 2008.
 9. Preeti Tiwari, Piyush Shukla, “Artificial Neural Network-Based Crop Yield Prediction Using NDVI, SPI, VCI”, at Springer Nature Singapore Pte Ltd.2020.
 10. D Ramesh, B Vishnu Vardhan, “Data Mining Techniques and Applications to Agricultural Yield Data”, *International Journal of Advances Research in Computer and Communication Engineering.*, vol. 2, no. 9, 2013.
 11. Beitollah Akbari Moghaddam, Hassan Haleh, Saeed Ebrahimijam, ”Forecasting Trend and Stock Price with Adaptive Extended Kalman Filter Data Fusion”, 2011 International Conference on Economics and Finance Research IPEDR vol. 4, 2011.
 12. G. M. Nasira, N. Hemaageetha, “Vegetable Price Prediction Using Data Mining Classification Technique”, Under International Conference on Pattern Recognition, Informatics & Medical Engineering, March 21-23,2012.
 13. M. Khalid, Mariam Sultana, Faheem Zaidi, “Prediction of Agriculture Commodities Price Returns Using ARMA and Wavelet”, *Journal of Natural Sciences Research.*, vol. 4, no. 23, pp. 2224-3186, 2014.
 14. Lu Ye, Xiaoli Qin, Yuping Li, Liu Weihong Liang, ”Vegetables Price Forecasting in Hainan Province Based on Linear and Nonlinear Combination Model”, *IEEE*, 2016.

AUTHORS PROFILE



Varun R. holds B.E in Information Science and Engineering from Vidyavardhaka College of Engineering.



Neema N holds her engineering degree in Information Science and Engineering from Vidyavardhaka College of Engineering.



Sahana H P completed her B.E. in Information Science and Engineering from Vidyavardhaka College of Engineering.



Sathvik A. holds his B.E. in Information Science and Engineering from Vidyavardhaka College of Engineering.



Mr.Mohammed Muddasir Asst Professor, Department of Information Science and Engineering in Vidyavardhaka College of Engineering, Karnataka, India.