Agriculture Commodity Price Forecasting using Soft Computing Techniques: A Survey

Rubi Kambo [1]
Department of Computer Science and Engineering
Bhilai Institute of Technology, Durg, C.G, India
rubi.kambo7@gmail.com

Dr. (Mrs.) Jyothi Pillai [2]
Department of Computer Science and Application
Bhilai Institute of Technology, Durg, C.G, India
jyothipillai71@gmail.com

Dr. (Mrs.) Sunita Soni [3]
Department of Computer Science and Application
Bhilai Institute of Technology, Durg, C.G, India
sunitasoni74@gmail.com

***Abstract*— Forecasting the price of agriculture commodity such as vegetables, fruits, cereals etc. is important related to economic concerned, farmer perspective, Agriculturist and Industrialist. Price forecasting help famers and the entire user of Agriculture to take effective decision regarding mandi price or selling price of their crop, which crop to grow to earn profit, ultimately improve the condition and income of famer and also helps policy maker for agriculture decision. Neural Network approaches are applied in the field of agriculture for price forecasting in both short term and long terms. Large amount of data related to commodity price, daily market price, arrival price is available. Neural approach with fuzzy can be used and also neuro fuzzy system may help in future for future price forecasting of commodity.**

Keywords*— *Agriculture commodity; Time series data; price forecasting; Artificial Neural Network; fuzzy

I. INTRODUCTION

In recent year, agriculture commodity price affects the economics of our country. Increasingly demand of crops or fruits in India or outside the country affect the price of commodity. Private Agriculture Industry or Farmer growth not only depends on the production or yielding of crop but also on the price of commodity. Factors such as time horizon, temperature, season, policies, productivity, and transportation cost etc., affect the price. Agriculture commodity price changes drastically and fastly and affecting the life of farmer and human being. Future price forecasting of agriculture commodity is not an easy task. Price fluctuation, price analysis, trend analysis in price can be done so that early information regarding crop price help farmer in decision related activity for a crop to sowed in field or to sell product in market.

Agriculture is one of the fields that contribute to Gross Domestic Product. New tools and techniques are developing that help famer providing early information related to crop, policies etc. So, for the risk related to crop price ,or pricing information analysis, price forecasting, price fluctuation of a crop are the different concerned area where one can applied the Soft computing technique, fuzzy approach, data mining technique to identify the truth or hidden pattern from agriculture product market price.

Data Mining is the process by which one can identify the hidden pattern, correlation among the data and also helps in forecasting the future trends. Data mining techniques can be applied in the field of agriculture for forecasting the price of product by predictive model task. Data mining forecasting model task include regression analysis, classification, time series data analysis and prediction.

Forecasting can be done by using historical day or past data about price and based on that future price can be forecast.

Regression analysis helps in mapping the data value to real value prediction variable. Mapping done by function and require learning that is one can use the past value for fitting to the function and can predict the future value.

Commodity price varies with time. Time points such as daily, weekly, yearly price value of commodity are obtained for forecasting future price based on time series data.

KDD is the process of finding useful and hidden information or patterns in data. In KDD process, input is data and output is information. While in Data mining algorithm is used to retrieve the information or pattern that are identified in KDD process.

One can also perform temporal analysis on agriculture price data of product varying with time. Association rule may involve temporal aspect and relationship. Based on time series future values can be predicted.

Neural Network is an information processing system inspired by the way biological neurons work and process information. NN helps in deriving information or extracting patterns from complicated or imprecise data.

Many researchers use neural network approach for forecasting either short term or long term price forecasting of vegetable commodity. Back propogation or recurrent neural network can be used for forecasting.

In this survey report, an idea of using soft computing approach or big data analytic approach for forecasting analytics of prices of agriculture commodities. Price

forecasting helps farmer and agriculturist to know about the growth or trend or price value of product before selling to mandi also help the policy maker for creation of different policy by generating rules according to price analysis. By analysing the price value, demand or supply as well as cultivated area related to crop helps farmer for decision making activity and ultimately helps in economic growth of our country.

II. Literature Survey

A. Soft Computing Techniques in Agriculture

From a year ago, artificial neural network is used by researchers for analysis, where its computational activity is similar to biological neurons.Neural network helps in dealing with problem related to finance economic.

Researchers Kohzadi et al. [l], for price forecasting of commodity compare neural network model with time series model. They used price data of US live cattle and wheat of forty years taken from USDA. Multi Layer Feed forward neural approach a supervised leraning technique is compared with ARIMA. Slinding window and walk forward approach is used. They suggested that NN perform better than ARIMA and Mean square error is about 27 and 56 percent lower than ARIMA. Also other measure shows by a neural network such as absolute mean error and mean absolute percent error were also lower.Neural netwokr has the capabiltiy to show major turning points for both wheat and cattle.

Aliev et al. [2], proposed an effective technique where for learning in neural network genetic approach used and also fuzzy inputs and fuzzy weights are feed in the network and also output produces is also a fuzzy. For quality evaluation and comparision ,simulated form of Fuzzy regression is explained by author and method is shows the usefulness of algorithm. Alpha cut and fuzzy arithmetic are used for learning. In oligopolistic environment auhtor demostrate their learning method for predicting fuzzy value profit. Genteic algorithm easily locate the global minimum error performance of fuzzy is easily located by Genetic algorithm where as alpha cut method fails.

In boiler industry pricing decision also play a great role and also assist in agriculture decision or also to pollutary association and also price help government for proper management of resources. Shih et al. [3], proposed a model for price prediction based on weighted case-based reasoning approach. Three cases such as equal wighted,unequal weigth and linear weighted CBR are investigated and compare.Most suitable weight for features are selected by genetic algorithm. For prediction model data related to imported chicken,economic index and production data from Taiwan agriculture is used. Result of proposed research shows that CBR approach performance is better than other predcition approach such as linear regression,regression tree,CART and neural net.Result also give information that colourful boiler and chiks price affect the price.

Zou et al. [4], for forecasting china food grain price, explore and compare the neural network and time series model. Researcher gives a new concept to create a model by combining two model such as time series model ARIMA and neural network approach such as back propagation using Equal weigth method i.e using arithmetic average of individual forecast as it is an easy method. Two kinds of evaluation criteria such as quantitative evaluation such as MSE,MAPE,and MAE and turning point is evaluated by advance version of Mertons test used by researcher. The result shows that forecasting performance of combine model is better then individual in terms of error evaluation measurement. Network structure such as 2*4*1 is selected to model price series and ARIMA(1,1,0) is realtively best as per researcher findings.Also suggested that ANN is best suited model for capturing profit and turning points and better then traditional ARIMA and also accuracy achieve by combine model is better than individual one.

In financial market, back propagation is used for discovering nonlinearity in financial data. But back propagation suffer from problem of low converges and is not robust method so the researchers Haofei et al.[5], for price froecasting learning task in nerual nets bring in concept of multi stage optimization in back propagation. In their paper they state that problem of backpropagation can be overcome by MSOA and performance of forecasting in terms of error and directional evalution measurement is better achieved by MSOA. Average Training time of MSOA is 4.25s and of Back propagation is 7.83s respectively. Covergence speed also improved by MSOA. Avg value of MSE of MSOA and BP algorithm is $2.11*10^{-4}$ and $6.22*10^{-4}$. And the max of MSE of MSOA i.e $1.24*10^{-3}$ is lower than that of BP whose value is $2.71*10^{-3}$.They compare the ARIMA,BP and MSOA and finds MSOA is best.

Yu et al. [6],uses neural network as a meta learning technique for designing a time series forecasting model to increase prediction accuracy while data consist of noise. Researchers do metamodelling by performing data partition and sampling for creation of different subset of training based on that base learning model is created and Fuzzy neural network is used in study for both as base learner and meta learner. They used PCA technique for model selection and pruning and then meta model is created from selected model.They concluded that prediction performance of nonlinear metamodelling technique is better when compare to single time series forecasting models such as ARIMA, FNN and SVM and other linear metamodelling technique such as simple averaging, simple MSE, Stack dregression ,variance weighting for financial time series data.netwrok perform well with pruning using PCA.

Gan-Qiong Li et al. [7], proposed a forecasting model where short term price forecasting of tomato has been done by three layer Feed forward neural network and result compared with time series ARIMA model and shows that neural network performs better for price prediction of one day with accuracy is about 90 % or week before price prediciton with accuracy about 80 %.

Riberio et al. [8], proposed a hybrid commodity price froecasting model for sugar price prediction using feed forward neural network such as multilayer perceptron model and Kalman filter. Data from brazilian and Indian market is

taken for study. Kalman filter is used to consider price as stochastic process and also include future price in forecasting and also minimise the error measure. ANN is used for Exogenous variable analysis and also ANN applied to Kalman filter result to get improvement in result.

Jha et al. [9], focuses on direction of change i.e prediction of turning point as important factor for any commodity price froecasting. They develop a ANN based decision support system included linear and nonlinear method such as ARIMA and ANN. Multi layer feed forward Time delay neural network has been employed in study.They use data of two crops of oilseed such as soyabean and rapeseed-mustard for prediction. Augmented Dickey –fuller test for different time series is used and also McLeod and Li for nonlinearity test is used. Out of twenty four model 2:3s:1L perform better than other computing model for soyabean series.

Wang et al. [10], design a system structure for short term price prediciton of agriculture product such as cucumber, tomato, cabbage, potato etc. of Beijing and they uses JAVA, Spring, Struts, and other open source project. Model is designed for market price transmission, price forecast and early warning and trend analysis. Proposed framework consist of database, system support, application service and system user layer. Effective volatile factor are estimated by multi linear or multi variate regression model with assist of Eview software. For product price risk evaluation method of value at risk is used by embedding easyfit software.For price prdeition both exponential smoothing and autoregressive moving average method is used.

Xiong et al. [11], proposed a new way that uses both linear and nonlinear approach and extend it to forecast interval of agriculture commodity future pricesof cotton and corn of china market by using Vector error correction model and multi-output support vector regression. VECM-MSVR forecast accuracy is better than individual VECM and MSVR and also with ARIMA-MSVR as VECM forecast the upper and lower bound of Interval valued timeseries simulataneously.For longer prediciton horizons such as h=3,5 the MSVR forecast is better than VECM.

Subhasree et al. [12], predict the next day price of vegetable using time series data. Machine learning algorithm such as Radial basis fucntion, back propagation neural network and genetic based neural network is used for study and comparative result concluded that predictive accuracy achieved by genetic based neural network is about 89% and error rate is 0.11 which is better than both back propagation whose accuracy is about 79 % and error rate is 0.21 and radial basis function accuracy is 52 % and error rate of 0.48.

Ahumada et al. [13], discussed that the forecasting accuracy of individual food price model can be improved by considering their cross dependency.Food price of Soyabean,corn and wheat are strongly correlated.They estimated Equlibrium correction model for each food price and them studied the residual cross correlation for interdependency and thus perfom a joint modelling. Four different econometric model are studied such as EqCM, DEqCM, DVAR and randomwalk econometric model are tested for each food price with different time horizon and scheme such as recursive or fixed.And they concluded that EqCM(Single for wheat and joint for corn and soyabean) perfrom best for time horizon h=4.

For forecasting price of agricultural product in brazil, researchers Pinheiro and senna [14] perform multivariate analysis of price of products and also uses neural network application for forecasting price i.e, combine the ANN model with multivariate analysis. They study on time series of price for product. Product chooses based on the export volume growth of product. Daily time series data is converted into weekly data. In their study, Forecast for ANN-MSSA and ANN are compared with 12 week data to final week sample. In this study the difference between ANN-MSSA and the ANN model is given by the treatment to separate noise from the original time series. Empirical data states that performance of MSSA-ANN methodology is better than that of ANN model. Study combines MSSA to decompose the time series with ANN model and good alternative for forecasting price of different commodity. Result will be useful to formulate and implement policies directed to agriculture sector. Predictive model performance is evaluated by means of square error. Thus helps in financial planning.

Wolfert et al. [15], discussed in their review paper about smart farming, use of ICT in farm management and also about the application of big data in smart farming. They discussed that big data helps in analytics of farm data and provide insight information about yield forecasting, product future price, business processes in farming. They also discussed about pull factor such as business driver, public drivers and push factor such as IOT, precision agriculture in smart farming. They concluded that big data analytics and its application will change the future of farming .

Shastry et al. [16], discussed that agriculture data set mainly contain continuous feature and for classifying dataset with multiple classes they used Multi-Class SVM. As SVM is popular as classification method, it uses concept of hyper plane for defining boundary for decision in high dimension feature space or input space. Generally mapping of input space to high dimension space is done by kernel for non linear problem in training data. Hybrid Kernel is used by author and they states that the choice of kernel highly affect the accuracy of classification as single kernel unable to take data from different sources and one drawback of single kernel is that it may have bias. Author uses genetic algorithm and gradient descent algorithm for selecting the optimal value for parameters c and γ as both affect the performance of SVM. They build ten kernels by combining linear, quadratic, polynomial, RBF, MLP on training data. They use GA for parameter selection not by Gradient descendent. For evaluation of method they use different performance metric. Author compare H-SVM model with NB, K-NN and SVM and found that H-SVM model is generic classifier suitable for any class of agricultural data not restricted to soil or crop data. They found that hybrid kernel that is quadratic and RBF kernel both show high performance for real world data and used as standard for agricultural multi class dataset. RBF kernel is local kernel that extracts the local feature of sample but unable to extract global features, so for global feature

extraction author uses quadratic kernel. Thus both kernel were combine and give the prominent result.

III. Conclusion

In the aforementioned work a study of different forecasting methodologies is done in agricultural market environment. To maximize profit in agriculture marketing, there is need to forecast the future price of commodity among the farmers and other participant of market. Multiple factor such as seasonality, trend, cycle, holidays, economic index etc., affects the price of commodity, for that reason price forecasting is not a simpler task.

It can be concluded that Forecasting require filtering, smoothing of data due to missing and noisy data, storage and retrieval of historical data, model creation for processing past data and predicting future trend or value. Performance measure such as MSE, MAPE is used for evaluating the performance of forecasting model.

There are several linear and nonlinear method individually or jointly used for forecasting such as ARIMA, Hybrid ARIMA-GARCH, and back propagation, Feed forward neural network, radial basis, genetic is used for optimization etc. are used by the researcher for forecasting.

Accordingly, in future other soft computing, Evolutionary approach with fuzzy can be used for better forecasting and also one can employ deep learning with big data for agriculture commodity price forecasting.

References

[1] Nowrouz Kohzadi, Milton S. Boyd, Bahman Kermanshahi, and Iebling Kaastra, "A comparison of artificial neural network and timeseries model for forecasting commodity prices," *NeuroComputing*, pp. 169-181, 1996.

[2] Rafik A. Aliev, Bijan Fazlollahi, and Rustam M. Vahidov, "Genetic algorithm-based learning of fuzzy neural networks.Part 1:feed-forward fuzzy neural network," *Fuzzy Sets and Systems*, pp. 351-358, 1998-2001.

[3] M. L. Shih, B. w. Huang, Nan-Hsing Chiu, C. Chiu, and W. Y. Hu, "Farm price prediction using case based reasoning approach-A case of broiler industry in Taiwan," *Computer and Electronics in Agriculture*, pp. 70-75, 2009.

[4] H. F. Zou, G. P. Xia, H. Y. Wang, and F. T. Yang, "An investingation and comparision of artificial neural network and time series model for chinese food grain price forecasting.," *NEURO COMPUTING:Elsevier*, pp. 2913-2923, january 2007.

[5] Zou Haofei, Xia Guoping, Yang Fangting, and Yang Han, "A neural network model based on the multi stage optimization approach for short term price forecasting in china," *Expert system with Application*, pp. 347-356, 2007.

[6] Lean Yu, Shouyang Wang, and Kin Lai Keung, "A neural network based nonlinear metamodeling approach to financial time series forecasting.," *Applied Soft Computing*, vol. 9, pp. 563-574, August 2009.

[7] Gan-qiong Li, Shi-wei Xu, and Zhe-min Li, "Short-Term price forecasting for Agro-products using Artificial Neural Networks," in *Elsevier*, 2010, pp. 278-287.

[8] Celma Q. Ribeiro and Sydnei M. Oliveria, "A hybrid commodity price forecasting model applie dto the sugar-alcohol sector," *Autralian Journal of Agricultural and Resource Economics*, vol. 55, no. 2, pp. 180-198, April 2011.

[9] Girish K. Jha and Kanchan Sinha, "Agricultural Price Forecasting Using Neural Network Model: an Innovative Information Delivery System," *Agricultural Economics Research Review*, vol. 26, no. 2, pp. 229-239, July-December 2013.

[10] Chuan Wang, Anping Zhao, and Yousen Zhao, "Design and Implementation of Agricultural Product Prices Short Term Forecasting System," in *Springer*, 2013, pp. 15-27.

[11] Tao Xiong, Chongguang Li, Zhongyi Hu, and Lu Zhang, "A combination method for interval forecasting of agricultural commodity furture prices," *Knowledge-Based System*, pp. 92-102, january 2015.

[12] M. Subhasree and Mrs. C. Arun Priya, "Forecasting Vegetable price using Time Series Data," *International Journal of Advanced Research*, vol. 3, pp. 535-541, february 2016.

[13] H. Ahumada and M. Cornejo, "Forecasting Food Prices : The case of Corn, Soyabean and Wheat," *International Journal of Forecasting*, pp. 838-848, 2016.

[14] Carlos Alberto Orge Pinheiro and Valter de Senna, "Multivariate analysis and neural network applications to price forcasting in the brazilian agricultural market," *Ciencia Rural,Santa Maria*, vol. 47, pp. 1-7, 2017.

[15] Sjaak Wolfert, Lan Ge, Cor Verdouw, and Marc- Jeroen Bogaardt, "BigData in Smart Farming : A review," *Agricultural System*, pp. 69-80, feburary 2017.

[16] K. Aditya Shastry, H. A. Sanjay, and G. Deexith, "Quadratic-radial-basis-function-kernel for classifying multi-class agricultural datsets with continuous attributes," *Applied Soft Computing*, pp. 65-74, april 2017.