International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 3 | ISSN : 2456-3307

A Review of the Application of Data Mining Techniques for Vegetable Price Prediction

Shivani Mahida, Bhargesh Patel
Department of IT G H Patel College of Engineering and Technology Vallabh Vidyanagar, Anand, Gujarat, India

ABSTRACT

Agriculture is the backbone of Indian economy. The volatile nature of Agriculture commodity prices cause the unfortunate effect on the economy and expose the growers to high risk. Nowadays Government and other commercial sectors provide fundamental data such as market information, weather information through several services using the internet. Data mining in addition to Information technology enables automation of extracting significant data in an effort to obtain knowledge and trends. Vegetable price prediction system aims to predict vegetable prices in near future by analyzing time series data on daily arrivals and prices of selected vegetable. By having the necessary information, farmers can make better and more informed decision regarding production and sell for their produces and government can frame appropriate agriculture policies.

Keyword: Data Mining, time series analysis, price forecasting, agriculture

I. INTRODUCTION

Agriculture in India is one of the most important sectors of its economy. Data mining in agriculture is an evolving research field. Agricultural commodities remain an important source of earnings. The overall performance of macroeconomic is affected by commodity price movements.

The prices of food and agricultural commodity in India are primarily determined by domestic market forces along with the domestic price policy. The imperfections in agricultural markets influence the prices and price transmissions. Besides temporal volatility, there also exits wide spatial variability in the prices of agricultural commodities such as varying climatic conditions, availability of resources, supply and demand requirements, festivals etc.

Agricultural price forecasts are an integral component of trade and policy analysis. Proper understanding of agricultural price mechanism and their forecast helps the farmers to plan and decide about the production plans and their marketing for improved farm profit, consumers to plan their budget, traders to know the market trend and government to augment economic development in the nation.

Information technology can play important role in bringing out sustainable agricultural development[1]. Data mining in addition to Information technology enables automation of extracting significant data in an effort to obtain knowledge and trend.

II. AGRICULTURAL PRICE AND TIME SERIES MODELING

Vegetables price modeling is different from modeling of non-farm goods and services due to certain special features such as perishable nature, storage availability, demand and supply requirements. In recent years, the agricultural prices have exhibited a very high volatility. For farmers to increase agricultural production and productivity a remunerative and

stable price environment is considered to be very important.

Time series analysis is one of the tasks of predictive data mining. Data mining is an important branch of computer science that extracts hidden patterns and useful information from large sets of data. Data mining tasks can be classified in two categories, descriptive data mining and predictive data mining[2]. Descriptive modeling is a mathematical process that describes real-world events and the relationships between factors responsible for them. Predictive analytics is the process of extracting information from large sets in order to make prediction and estimates about future outcomes.

For forecasting agricultural commodity prices there are two basic approaches available namely structural model and time series models[10].

Structural model proceed from the first principles of consumer and producer theory to identify the demand and supply schedules and the equilibrium prices resulting from their intersection. Time series modeling is used to describe the hidden relationship of one or more variable by analyzing past observations of them.

III. RELATED WORK

N. Hemageetha, G M Nasira[3], have proposed Back-propagation neural network for forecasting weekly and monthly vegetable price. Neural Network is used for its characteristics such as self-adapt, self-study and high fault tolerance. Data Normalization is used to transfer the data to fit within the limit of transfer function. They have also explored prediction model that uses BPNN based on Radial Bias Function[4]. The back propagation neural network gives accuracy of 77.42% but RBF neural network gives 85.55%.

N. Hemageetha, G M Nasira[5], have conducted similar study that analysis Adaptive Neuro-Fuzzy Inference System for vegetable price forecasting. This Inference Systems model for ANFIS uses a first-order Takagi-Sugeno-Kang (TSK) system as inference to generate the if-then rule to build the model that maps input to output. The proposed ANFIS model gave total prediction accuracy of 96.91%.

YE Lu et al[6], provide a review of research based on PSO-BP Neural network to predict vegetable prices. This study concludes that PSO when used with neural network improves generalization ability of neural network. When compared with traditional BP neural network, SAD and MSE of BP neural network optimized by PSO reduce respectively by 65% and 66% indicating that PSO improves the prediction accuracy.

YouzhuLi et al[7] investigated the performance of Hybrid Neural Network and H-P Filter for forecasting vegetable prices. To improve performance of forecasting model proposed method decomposes the raw data into trend data and cyclic data using H-P filter. Both trend and cyclic data is trained separately using artificial neural network but once they are trained, these data is combined together for vegetable price predictions

M.Subhasree, C.Arun priya[8] have analyzed the neural network model based on Genetic Algorithm. GA is an evolutionary optimization algorithm. GA is used to optimize the process of BPNN by effectively selecting initial parameters of neural network. Proposed method provides better results compare to BPNN.

Purna Chandra Padhan[9] describes the applicability of Auto regressive integrated moving average (ARIMA) model for forecasting vegetable prices. ARIMA model uses only stationary series; so the original series was made stationary by applying appropriate order of differencing. The efficiency of the proposed models is judged based on the MAPE values. The outcome shows that the proposed model can forecast the onion demand with an accuracy of 71.71%

Hakan Adanacioglu, Murat Yercan[11] provides a review of research on the application of Seasonal ARIMA (SARIMA) model. SARIMA model uses Seasonal indexes to detect the fluctuations and seasonality of vegetable prices by measuring how much the average for a particular period tends to be above or below the expected value. The outcome shows that the proposed model can forecast prices with prediction accuracy of 75.65%.

Girish K. Jhaa, Kanchan Sinhab[10], have constructed a hybrid ARIMA-ANN model for vegetable price forecasting. In this study the first ARIMA model is applied on the dataset. As ARIMA model is used for linear forecasting the non-linear part of the data is considered as residuals. Residual part is then modelled by neural network. Result of this study shows that hybrid model is more accurate than any individual ARIMA or ANN model for non-linear time series data.

P. Jasinthan et al[12] used Markov chain model which analysis the market property to understand the price movement of market. Authors have created two models, where the price movement is considered as being in a state of gain or loss and large gain, or small gain or loss, or large loss. This study finds similarity of price movements among various vegetables and accordingly suggests investors to invest in the vegetable market at any time in a way which leads to a greater chance of getting more gain than loss.

Table 1. Comparison Table

Author	Year	Data	Approach	Description
N. Hemageetha, G M Nasira	2012	Tomato weekly and monthly prices from Jan 2009 to may 2011	Back-propagation Neural Networks	Neural Network is used to predict tomato prices. Data Normalization is used to transfer the data to fit within the limit of transfer function.
N. Hemageetha, G M Nasira	2013	Tomato weekly and monthly prices from Jan 2009 to may 2011	Radial Bias Function Neural Networks	RBF NN is proposed which is capable of fast learning compared to NN.
N. Hemageetha, G M Nasira	2017	Tomato weekly prices from 2009 to 2012	Adaptive Neuro-Fuzzy Inference system	ANFIS model generate if-then rule that maps input to output.
YE Lu, LI Yuping, LIANG Weihong, SONG Qidao, LIU Yanqun, QIN Xiaoli	2015	Green pepper prices from January 2012 to March 2015 including compound fertilizer, glyphosate and gasoline price	Back-propagation Neural Networks, Particle swarm intelligence	PSO an optimization algorithm is used to assign initial weights of NN.
YouzhuLi, ChongguangLi, MingyangZheng	2014	Monthly price data for five vegetables from 2012 to 2013	Hybrid model of Neural Network and H-P Filter	The linear HP filter extracts the trend and cyclical component which is learned by NN separately to produce better results.

M. Subhasree, Mrs. C. Arun priya	2015	Weekly data of tomato	Neural Network, Genetic Algorithm	Genetic Algorithm is used for optimization of neural network model. This model provides better accuracy than BPNN
Purna Chandra Padhan	2012	Annual productivity data of 34 various agricultural products from 1950 and 2010	ARIMA model	ARIMA forecasting model is applied for large stationary data and involved four different but such as identifying, fitting, estimating and forecasting the annual productivity.
Girish K. Jhaa, Kanchan Sinhab	2013	Monthly data of soybean from October, 1991 to September, 2010 and mustard from January, 1980 to December, 2010	ARIMA-ANN model	ARIMA is used to model linear part of data and its residual (considered to have non-linear property) are modeled by ANN model.
Hakan Adanacioglu, Murat Yercan	2012	Tomato prices from 2000 to 2010	Seasonal ARIMA (SARIMA) model	Seasonal indexes are used to Analysis the fluctuation in tomato prices.
P. Jasinthan, A. Laheetharan and N. Satkunanathan	2015	Daily price data of 12 vegetables from Jan 2009 to Sept 2013	Markov Chain Model	Markov model analyses the pattern of price movement by considering the prices being in a state of gain or loss.

IV. CONCLUSION

Data mining in agriculture is a novel research field. Various researches have been carried out for vegetable price forecasting. ARIMA model is one approach to predict vegetable prices by transforming non-stationary series into stationary model. Recently, Artificial Neural Network (ANN) modeling has attracted much attention as an alternative technique due to its characteristics such as non-linearity self-adapt, self-study and high fault tolerance.

V. REFERENCES

[1]. Prof. B. MiloviC, V. Radojevi, "Application of Data Mining in Agriculture", Bulgarian Journal of Agricultural Science, 2015

[2]. Pradnya P. Sondwale , 'Overview of Predictive and Descriptive Data Mining Techniques' International Journal of Advanced Research in Computer Science and Software Engineering, 2015

[3]. N. Hemageetha, G M Nasira, "Vegetable Price Prediction Using Data Mining Classification Technique", IEEE, 2012

[4]. N. Hemageetha, G M Nasira, "Radial Basis Function Model for Vegetable Price Prediction", IEEE, 2013

[5]. N. Hemageetha, G M Nasira, "Vegetable Price Prediction using Adaptive Neuro-Fuzzy Inference System", International Journal of Computer Sciences and Engineering, 2017

[6]. YE Lu, LI Yuping, LIANG Weihong, SONG Qidao, LIU Yanqun, QIN Xiaoli, "Vegetable

Price Prediction Based on PSO-BP Neural Network", IEEE, 2015

[7]. YouzhuLi, ChongguangLi, MingyangZheng, "A Hybrid Neural Network and H-P Filter Model for Short-Term Vegetable Price Forecasting", Hindawi, 2014

[8]. M. Subhasree, Mrs. C. Arun priya "Vegetable price prediction based on time series analysis", IRACST, 2015

[9]. Purna Chandra Padhan, "Application of ARIMA Model for Forecasting Agricultural Productivity in India", J. Agric. Soc. Sci., 2012

[10]. Girish K. Jhaa, Kanchan Sinhab, "Agricultural Price Forecasting Using Neural Network Model: An Innovative Information Delivery System", Agricultural Economics Research Review, 2013

[11]. Hakan Adanacioglu, Murat Yercan, "An analysis of tomato prices at wholesale level in Turkey: an application of SARIMA model", Cost and Agribusiness, 2012

[12]. P. Jasinthan, A. Laheetharan and N. Satkunanathan , "A Markov Chain Model for Vegetable Price Movement in Jaffna ", IASSL-2015