International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 1 | ISSN : 2456-3307

Survey on Price Prediction Techniques

Sindhuja T[1*], Sakhi Chawda[1], K. P. K. Devan[2]

[1]UG Students, Computer Science and Engineering, Easwari Engineering College, Anna University, Chennai, India

[2]Associate Professor, Computer Science and Engineering, Easwari Engineering College, Anna University, Chennai, India

ABSTRACT

Agriculture is the major occupation of India. The farmers who are the backbone of the country are suffering in utter poverty. One of the main reasons being that they are not totally aware of the business in the market. As a result of which they fall prey to the tricky dealers who convince them to sell their harvest at a price much lower than actually it should be. The systems are currently available to the educated people who deal with the economics of the agriculture. They predict the trend and causes, and finally publish a survey in the newspapers which brings the scenario in the eyes of educated public but not the actual sufferers. The academic papers taken as references do a comparative study for the efficiency of the prediction algorithm in terms of vegetable prices as a dataset. The algorithm that they used is in terms of neural networks and will help in prediction over a short period of time. Analyzing data over a time period regularly will lead to various insights and conclusions. Hence, this system suggests a time-series approach to develop a forecast model and predict, by considering the prices over a period of time.

Keywords : Time-Series, Prediction, Forecasting.

I. INTRODUCTION

The majority of population in India earns their daily living by cultivating crops. This is popular in the rural part of India where the people are unaware of the situations in the cities. The most affected in the process of business of crops are the poor farmers who are exploited because they are not aware of the happenings of the outside world. The fixing of prices by the dealers who bring the produce to the market is mostly unfair as they tend to have fixed price of the produce they buy irrespective of their price in the market outside. Technology has now evolved a lot. Everyone now has access to it. It can be used to educate and update the unaware regarding the changes in the outside world.

II. LITERATURE SURVEY

[1] Explains the challenges of forecasting crude oil price due to its high volatility. Since crude oil is the world's leading fuel, and its prices has a global impact on the environment and the economy the prediction method used should be yield high accuracy hence, a new machine learning paradigm called stream learning is used. In stream learning the model is continuously updated as it captures the changing pattern in oil price with small overhead constant. To prove the efficiency of stream learning, the model is compared to three different models: (1) heuristic approaches; (2) econometric models; and (3) machine learning techniques. Heuristics approach included professional and survey forecasts, which are mainly based on professional knowledge, judgments, opinion and intuition. Econometric model included autoregressive moving average (ARMA) models and

vector autoregressive (VAR) models. Machine learning techniques included artificial neural networks (ANN) and support vector machine (SVM). A new stream learning approach was developed to handle applications where continuous data streams are generated from non-stationary processes. Like any machine learning data set, the data is divided into training set and test set. For such an approach it is assumed that both the data sets are homogeneous. Hence, stream learning takes into account recent data history and is updateable.

[2] explains about the most popular measure for accurate forecasting-mean absolute percentage error(MAPE).A significant comparison is done between two forecasting measures, mean absolute percentage error and mean arctangent absolute percentage error(MAAPE) which was developed by looking at mean absolute percentage error from a different perspective. Essentially, MAAPE is a slope as an angle, while MAPE is a slope as a ratio. MAAPE preserves the idea of MAPE but overcomes the problem of division by zero by using bounded influences. MAPE measures the error in terms of percentage, it calculates the number of unsigned percentage errors. Although MAPE is used highly due to its properties of f scale-independency and interpretability it has its issues too, it produces infinite or undefined values for zero or close-to-zero which is a common occurrence in some fields, a very large percentage error is given if the actual values are very small(one or less than one). The MAPE, as a percentage, only makes sense for values where divisions and ratios make sense. It doesn't make sense to calculate percentages of temperatures, for instance, so you shouldn't use the MAPE to calculate the accuracy of a temperature forecast. MAPE allows us to compare forecasts of different series in different scales.

[3] Electricity Price Forecasting (EPF) looks ahead and speculates the direction EPF will or should take in the next decade or so. Five major topics are discussed about EPF (1) Fundamental price drivers and input variables, (2)Beyond point forecasts, (3) Combining forecasts, (4) Multivariate factor models and (5)The need for an EPF-competition. A key point in EPF is the appropriate selection of input variables. On the one hand, the electricity price exhibits seasonality at the daily and weekly levels, and the annual level to some extent. On the other hand the electricity spot price is dependent on a large set of fundamental drivers, including system loads (demand, consumption figures), weather variables (temperatures, wind speed, precipitation, solar radiation), fuel costs (oil and natural gas, and to a lesser extent coal) etc. In Beyond point forecasts we look into interval forecasts which is associated with random variables, density forecasts and threshold forecasting. Combining forecasts combines several forecasts together which shows significantly better results than individual forecasts. There is a need for EPF-competition due to different datasets, different software implementations of the forecasting models and different error measures, but also to the lack of statistical rigor in many studies.

[4] Using data mining techniques for stock prediction to help financial investors to make subjective decision. One such technique is artificial neural network (ANN). In ANN the use of technical analysis variables for stock prediction is prevalent. It also explains a hybridized approach which integrates the use of variables of technical and fundamental analysis of the stock market for better prediction and improvement of existing approaches. ANN is gaining heavy attention due to its ability to learn and detect relationship among nonlinear variables. It also does a deep analysis of large sets of data that usually have a tendency to fluctuate within a short span of time. One of the major drawbacks of ANN is too many hidden node and it consumes a lot of time. The focus here is to improve the prediction by using a hybridized method. The technical analysis variables are the core stock market indices which include current stock price, opening price, closing price, volume, highest price and lowest price., while the fundamental analysis variables are company

performance indices such as price per annual earning, rumor/news, book value and financial status etc.

[5] explains the working of an Artificial Neural Networks(ANN) with respect to the forecasting of load demand of electricity. This technique is used for short-term load forecast. The prediction was done on the active hourly variations of power collected from Renigunta substation A.P., India over a period of one month. For training the ANN, active powers were taken as the input quantities and obtained respective active powers for the corresponding day as the output. The algorithm comprises of a network with 5 nodes as input, 1 output node and 21 hidden nodes. The process of training involved setting up the nodes' weights with random values between 0.5 to 5.5. Since ANN is a Back-Propagation Network (BPN), for every input vector an output vector is obtained and the error between the desired value and the network output is calculated. This error is used to adjust the weights of the network in a way to minimize the error. The error is back propagated until the error value obtained is considerably low. The issues with ANN is that it requires a fine-tuned architecture to achieve better accuracy.

[6] forecasts short-term prices for electricity using fuzzy neural networks (FNN). This FNN comprises of a hypercubic training mechanism with feed-forward architecture. It is designed by combining fuzzy logic with a learning algorithm that models no stationary behavior and outliers for the prices presented. The FNN comprises of Neural Network (NN) and fuzzification method. A feed forward architecture is used to represent NN that has input layer, hidden layer and output layer. The input features are propagated to the hidden layers which obtains a vector of weighted inputs. The learning process of FNN involves two tasks by adjusting the parameters: the classification of input vector space based on training samples into hypercubes and, implementing each functional relationship in one class. A comparative study was carried out among DR, TF, MLP, RBF, ARIMA, wavelet-ARIMA and FNN. The performance of FNN was greater than other models. But for FNN, the relationship among various factors has to be assessed in order to determine an accurate forecast, which in this case is not considered.

[7] It is a review on the past 25 years of time series forecasting. It explains over 12 forecasting methods along with their issues and how one overcomes the issue of the other. Starting with Exponential smoothing which smoothens time series data using exponential window function, its drawback is It produces forecasts that lag behind the actual trend. ARIMA model while linear exponential smoothing models are all special cases of ARIMA models, the non-linear exponential smoothing models have no equivalent ARIMA counterparts. There are also many ARIMA models that have no exponential smoothing counterparts. Seasonality, the oldest approach to handling seasonality in time series is to extract it using a seasonal decomposition procedure such as the X-11 method. Seasonality generally cannot be identified until the trend is known, however a good estimate of the trend cannot be made until the series has been seasonally adjusted. Therefore X11 uses an iterative approach to estimate the components of a time series. As a default, it assumes a multiplicative model. State space and structural models and the Kalman filter, used in the early 1980's where statisticians used state space model for time series. State space models provide a unifying framework in which any linear time series model can be written. As years went on several other models developed such as nonlinear models, Long memory models, ARCH/GARCH models, Count data forecasting, Combining and Prediction intervals and densities.

[8] Explains the advantages of using exponentially weighted moving average that is used to smoothen random functions that has the following desirable properties: (1) declining weight is put on older data, (2) it is extremely easy to compute, and (3) minimum data is required. This paper utilizes these desirable

properties both to smooth current random fluctuations and to revise continuously seasonal and trend adjustments. The flexibility of the method combined with its economy of computation and data requirements make it especially suitable for industrial situations in which a large number of forecasts are needed for sales of individual products. The Moving Average model takes a data set with variation and creates another data set with less variation or a smoothed data set by aggregating several periods of data. The routines are designed to remove seasonal and random noise variation within time series. Applying this routine repeatedly would result in removal of cyclic variation and left with combination of trend and some cyclic behavior. The smoothing effect of the moving average model provides for a "cleaner" data set, which may or may not help in estimating the future level of a variable. The concurrent disadvantage of the greater sensitivity of the EMA is that it is more vulnerable to false signals and getting whipsawed back and forth.

III. COMPARATIVE ANALYSIS

S.No	TITLE	AUTHOR(S)	APPROACHES	RESULTS	ISSUES
1	A new approach for crude oil price prediction based on stream learning	Shuang Gao, Yalin Lei	Artificial Neural Networks(ANN) SVM, VAR, ARMA	Handles continuous data streams from non-stationary processes	Data taken into consideration is homogeneous
2	A New Metric Of Absolute Percentage Error For Intermittent Demand Forecasts	Sungil Kima, Heeyoung Kimb	Mean Absolute Percentage Error (MAPE) Mean Absolute ArcTangent Percentage Error (MAAPE)	MAAPE can only overcome division by zero error. MAPE is a widely used metric.	MAPE cannot be used for domains like temperature.
3	A look into the future of 'Electricity Price Forecasting'	Rafał Weron	Beyond Drive Forecasts, Point Drive Forecasts, Combining Forecasts	Combining forecasts will lead to better results	The datasets for forecasts lack statistical rigor.
4	Stock Price Prediction using Neural Network with Hybridized Market Indicators	Adebiyi Ayodele A., Ayo Charles K., Adebiyi Marion O., and Otokiti Sunday O.	Artificial Neural Network (ANN)	Prediction of stock prices by identifying relationship among non-linear variables.	ANN involves a longer training time.
5	Load Forecasting by a novel technique using ANN	T. Gowri Manohar and V. C. Veera Reddy	Artificial Neural Network (ANN)	Forecasting of load demand of electricity.	ANN requires a fine-tuned architecture for better results.
6	Day-Ahead Price Forecasting of Electricity Markets by a New Fuzzy Neural Network	Nima Amjady	ARIMA, wavelet-ARIMA, Fuzzy Neural Network (FNN)	The performance of FNN is higher than others.	Relationship among all factors has to be considered.
7	25 years of Time series Forecasting	Jan G. De Gooijer a,1, Rob J. Hyndman b	Exponential Smoothing, ARIMA	ARIMA is a better counterpart.	The non-linearity in data cannot be handled.
8	Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages	Charles C. Holt	ARIMA, Exponential Moving Average (EMA)	EMA removes seasonal and random noise for a better forecast	EMA responds to false signals.

IV. FUTURE SCOPE

The time-series approach comprises of various algorithms such as HoltWinters, ARIMA, ARMA, Recurrent Neural Network and so on. Based on the number of data points available and the scalability of the application that are available any one of the best suitable technique can be implemented to forecast the future market prices.

V. CONCLUSION

A survey on various methods of price prediction has been presented. The survey mainly comprised of price predictions of stock, electricity and crude oil. Various techniques used in price predictions were studied and the results of each technique had been formulated. The techniques mainly consisted of ANN, neuro-fuzzy logic and ARIMA. The performance constraints and the limitations have also been taken into account. An accurate prediction is obtained by the selection of best model to forecast the future prices as further learning of the model is improved. This paper comprises of forecasting techniques for prices using deep learning and time series techniques. This paper will provide a run-through regarding the existing forecasting techniques and help in deciding a better solution to the problem.

VI. REFERENCES

[1]. Shuang Gao, Yalin Lei, A new approach for crude oil price prediction based on stream learning, GeoScience Frontiers, Elsevier, 2017

[2]. Sungil Kima, Heeyoung Kimb, A New Metric Of Absolute Percentage Error For Intermittent Demand Forecasts, International Journal of Forecasting, Elsevier, 2016

[3]. Rafal Weron, A look into the future of 'electricity price forecasting', International Journal of Forecasting, Elsevier, 2014

[4]. Adebiyi Ayodele A., Ayo Charles K., Adebiyi Marion O., and Otokiti Sunday O., Stock Price Prediction using Neural Network with Hybridized Market Indicators, Journal of Emerging Trends in Computing and Information Sciences, 2012

[5]. T. Gowri Manohar and V. C. Veera Reddy, Load Forecasting by a novel technique using ANN, ARPN, 2008

[6]. Nima Amjady, Day-Ahead Price Forecasting of Electricity Markets by a New Fuzzy Neural Network, IEEE Transactions On Power Systems, Vol. 21, NO. 2, May 2006

[7]. Jan G. De Gooijer, Rob J. Hyndman, A look into the future of 'electricity price forecasting', International Journal of Forecasting, Elsevier, 2006

[8]. Charles C. Holt, Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages, International Journal of Forecasting, Elsevier, 2004