2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February, 2019

Indices prediction of Bangladeshi stock by using time series forecasting and performance analysis

Md. Mujibur Rahman Majumder
Department of ICT
Comilla University
Cumilla, Bangladesh
mujib.contact@gmail.com

Md. Imran Hossain
Department of ICT
Comilla University
Cumilla, Bangladesh
sujan_bgd@yahoo.com

Mohammad Kamrul Hasan
Department of ICT
Comilla University
Cumilla, Bangladesh
Kamrul2363@gmail.com

***Abstract*— Prediction of share index is a very mystical task for the traders who take part on stock business. Investors invest their money to earn profit from this sector. However often they face misfortune in their life due to the wrong prognosis of stock index. In this work, we have predicted the indices of Bangladeshi stock by using various stock prediction algorithms such as Feed-Forward Neural Network (FFNN), Auto Regressive Integrated Moving Average (ARIMA), Linear model & Holt- Winter approaches and analyzed the performance of these algorithms over 35 Bangladeshi stocks. Time series analysis is considered to this work and performance of algorithms is computed by calculating percentage accurate prediction. From analysis, it is found that, ARIMA (1,0,0) gives maximum prediction accuracy (82.1%) in average among all and FFNN shows best algorithm in forecasting stocks index. FFNN gives maximum accuracy in 14 out of 35 stocks.**

Keywords— Time series forecasting; Stock market prediction; Data Mining; ARIMA; Neural Network

I. INTRODUCTION

A place where various kinds of exchange materials such as shares, bonds and other monetary apparatus can be chaffered and sold out by the investor is known as S*tock exchange*.

There are 2 stock exchanges in Bangladesh -

1. Dhaka Stock Exchange (DSE).
2. Chittagong Stock Exchange (CSE).

Stock price prediction is a procedure of ordain the oncoming worth of any stock material such as bond, shares etc. The successful forecasting of a share's oncoming price is profitable business for the investors.

Prediction method can be classified into 3 categories in broad sense -

1. Fundamental analysis.
2. Technical analysis.
3. Machine learning.

Fundamental analysis is the study which concern with the company of a particular stock itself. In fundamental analysis, analysts evaluate the past story and performance of an organization. Analyst on fundamental analysis belief that if a company works well their stock price would be upward and if company don't operate well their market would be downward. In fundamental analysis, analysts also concern the global economy, domestic economy of that stock organization and company level economy related to that stock.

In technical analysis, analysts are concerned with the past price of a particular stock rather than company level analysis. "Time series analysis", most common method of forecasting, is fallen into this category. The objective of time series analysis is to build a model the forecast the future price of a stock based on past price. Time series analysis is also used in various discipline such as forecasting the weather, earthquake and rainfall at any geographic area. It is also used in statistics, pattern recognition, electroencephalography, econometrics, mathematical finance, signal processing, astronomy and communications engineering.

In machine learning, Researchers proposed several machine learning algorithms that forecast future value of stock. Genetic Algorithm (GA) and Artificial Neural Networks (ANNs) algorithm is the researcher's major interest in machine learning. Researchers has formed different types of neural network algorithm from which Feed forward network was most prominent algorithm in the past year that forecast stock prices by utilizing the back-propagation networks. Still researcher finds interest on this algorithm. Recurrent neural network (RNN) is another form of ANN, used in forecasting stock price.

Data is the primary concern in any prediction methodology. In my work, Data of stocks are collected from Dhaka Stock Exchange's data archive (*https://www.dsebd.org/data_archive.php*). From this website, Data of 35 stocks are collected for analyzing purpose, where data contains closing price and yesterday closing price (Opening price) with respect to date for a stock.

In this work, we have used following methods over DSE's data-

- Auto Regressive Integrated Moving Average (ARIMA).
- Feed- Forward Neural Network (FFNN).
- Linear model
- Holt- Winters model
- Holt- Winters exponential smoothing model

978-1-5386-9111-3/19/$31.00 ©2019 IEEE

II. Related Work

Prediction of stock price is a complicated process due to randomness of stock price. In forecasting the price of stock, Researchers has used various algorithms.

Haider [1] has used ARIMA model in forecasting DSE returns. He has used ARIMA model for predicting index series and return series and finds that ARIMA (3,1,2) are best fitted model for index series and ARMA (3,1) for return series data.

Bose et al. [2] has also used ARIMA model over Dhaka Stock Exchange (DSE) and Chittagong Stock Exchange (CSE) stocks data. They have showed that CSE had better efficiency than DSE.

Kamruzzaman et al. [3] have used Box-Jenkins method to examine the returns of DSE of Bangladesh. They found that the ARIMA (2, 0, 2) model appears best than the rest of the time series models in dealing with market return volatility.

Adebiyi et al. [4] have found that ARIMA model has a brawny potential for meager-term forecasting. Sharma et al. [5] use machine learning approaches in forecasting stock market. They are successful to build a survey on stock price by using regression approach.

Somani et al. [6] and Hassan-Nath [7] uses Hidden Markov model in predicting stock price. Somani calculated the percentage average absolute error (MAPE) between predicted and actual stock price and the formula is -

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{(p_i - a_i)}{a_i} \times 100 \dots\dots\dots\dots\dots\dots (1)$$

Where p_i is the predicted stock price, a_i is the actual price of stock in day i. n is the number of days [6]. Somani et al. were successful in predicting stock than Hassan and Nath model. Model of Hassan and Nath finds the patterns from the bygone data sets that matched with the present behavior of the stock price.

Abhishekh et al. [8] forecast the stock market by using Feed Forward Neural Network approaches with Back- Propagation. They have shown that the accuracy of prediction was 99% in case of 1m (with two layer, 10 neurons in input layer and 1 in output layer) and have found mean square error of 0.00650.

III. Some Common Stock Prediction Algorithms

A. Holt-Winters Model

Holt-Winters method is one of common methods of forecasting time series data without concernment of accompanying a parametric model. Holt-Winters smoothing is extensively used for predicting data that consist of seasonality, changing tendencies and seasonal interrelation [9]. Holt-Winters provide additive and multiplicative function for prediction. Halt- Winters also provides seasonal method which comprise of forecasting and smoothing equations.

The additive prediction function [10] of Holt-Winters is represent by -

$$\hat{Y}[a+h] = L[a] + h * T[a] + S[a + 1 + (h-1) \bmod p] \dots\dots\dots\dots (2)$$

$$L[a] = \alpha\,(Y[a] - S[a-p]) + (1-\alpha)(L[a-1] + T[a-1]) \dots\dots\dots\dots (3)$$

$$T[a] = \beta\,(L[a] - L[a-1]) + (1-\beta)L[a-1] \dots\dots\dots\dots\dots (4)$$

$$S[a] = \gamma\,(Y[a] - L[a]) + (1-\gamma)S[a-p] \dots\dots\dots\dots\dots\dots (5)$$

Where equation (2) is the forecasting equation and (3),(4),(5) represent 3 smoothing equations for additive model. L[a] is the level, T[a] is the trend and S[a] is the seasonal smoothing parameters.

The multiplicative prediction function [10] of Holt-Winters is represent by -

$$\hat{Y}[a+h] = (L[a] + h * T[a]) * S[a + 1 + (h-1) \bmod p] \dots\dots\dots\dots (6)$$

$$L[a] = \alpha\,(Y[a] / S[a-p]) + (1-\alpha)\,(L[a-1] + T[a-1]) \quad (7)$$

$$T[a] = \beta\,(L[a] - L[a-1]) + (1-\beta)T[a-1] \dots\dots\dots\dots\dots (8)$$

$$S[a] = \gamma\,(Y[a] - L[a]) + (1-\gamma)S[a-p] \dots\dots\dots\dots\dots (9)$$

Where equation (6) is the forecasting equation and (7),(8),(9) represent 3 smoothing equations for multiplicative model. L[a] is the level, T[a] is the trend and S[a] is the seasonal smoothing parameters and p is the period. Here α, β & γ smoothing parameters.

B. Autoregressive Integrated Moving Average Model

Autoregressive Integrated Moving Average (ARIMA) model is the successor of the ARMA model. ARIMA models are harmonized into time series data for predicting the coming point of the time series and for better perception of data. Seasonal and non-seasonal models are 2 types of ARIMA model, where ARIMA(p,d,q) is prevailed to non-seasonal and (p,d,q)(P,D,Q)m is prevailed to seasonal model. In non-seasonal data, p is order of auto regressive, d is the degree of distinction and moving average is denoted by q. P,D,Q refers to the same but in seasonal trend and m denotes the length of period[11].

In this paper, we have considered ARIMA first order autoregressive (1,0,0) model and the equation for first order autoregressive [12] is -

$$\hat{Y}_t = \mu + \phi_1 Y_{t-1} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (10)$$

Here ϕ_1 is slope coefficient.

C. Linear Model

Linear model is often called statistical model and this model is often used in regression purpose. In case of time series, the regression model [13] can be written by using -

$$X_t = c + \epsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \dots\dots\dots (11)$$

Where, ε_i representing small white noise, β_j is regression coefficient. X_t is linear function of past values.

D. Feed Forward Neural Network (FFNN)

Feed Forward Neural Network is the simplest types of artificial neural network that contains 3 layers (input, hidden, output layer). The characteristics of this network is that nodes of this network don't create any cycle to the network and information step to forward one direction through the hidden network nodes to the output nodes [14].

Output of FFNN neural network can be written as-

$$O = \sum_{j} [B_j g(\sum_{i} A_{ij} I_i)] \dots\dots\dots\dots\dots\dots\dots\dots (12)$$

In this equation, the weight of the i[th] input neuron to the j[th] hidden layer neuron is represented by A_{ij} and the weight of the j[th] neuron of the hidden layer to the output layer neuron represent by B_j.

Sigmoid transfer function for FFNN can be written as [14] -

$$s(x) = \frac{1}{1 + e^{-x}} \dots\dots\dots\dots\dots\dots\dots\dots (13)$$

IV. Methodology

Our working procedure consists of following steps-

1. Read pre-processed stock data: In our research, we have worked with weakly data of DSE stock. As time series forecasting predict the ordered stock data, thus we ordered the dataset and filled out missing value.

2. Convert data into time series: In order to perform time series forecasting, we have to convert original dataset into time series data which makes our data equal spacing between every 2-consecutive data.

3. Split the stock data: In order to evaluate prediction, Data of stocks are divided into 2 sets-
   - Train set
   - Test set

In our research, we have classified data from 1st January 2016 to 31 December 2017 into train set and data from 1st January 2018 to 31 July 2018 into test set. We have worked with data ranges from 1 January 2016- 31 July 2018.

4. Determine polynomial and STL trends from the train data set.

5. Apply prediction algorithms over polynomial, STL, Actual trend data: Prediction algorithms such as Feed-Forward neural network, Auto Regressive Integrated Moving Average (ARIMA)(1,0,0) order & ARIMA best fit model, Linear model, Holt-Winters and Holt-Winters exponential smoothing are applied for all kinds of polynomial, STL and actual functions data.

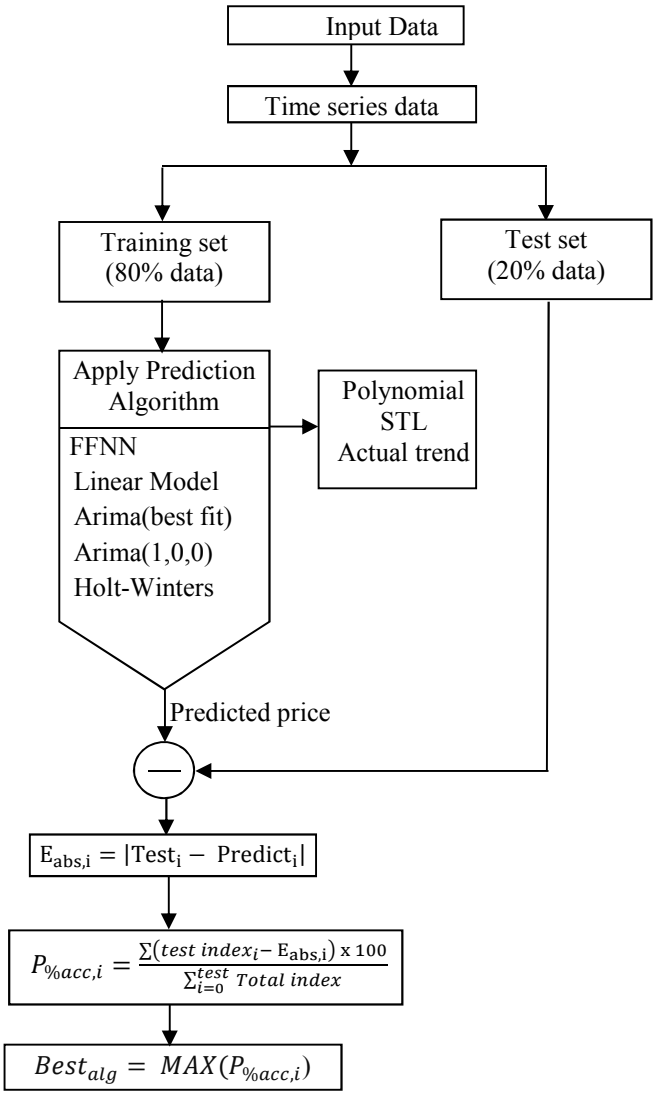

Figure 1: Flow chart of methodology steps

6. Determine best algorithm: Determining best prediction algorithm is the primary concern of this work. The determination is done by computing the errors occurring in the forecasting by various algorithms. Absolute errors can be calculated by using following equation –

$$E_{abs,i} = |Test_i - Predict_i| \dots\dots\dots\dots (14)$$

The algorithm which returns minimum error is the selected algorithm (best) for a particular share. Prediction accuracy can be computed by following equation-

$$P_{\%acc,f} = \frac{\sum(test\ index_i - \mathrm{E_{abs,i}}) \times 100}{\sum_{i=0}^{test} Total\ index} \ldots(15)$$

Here, $P_{\%acc,f}$ is the percentage accuracy per filter. For which algorithm (filter), the prediction accuracy is $P_{\%acc,f}$ is high is also the best filter for desire share. That means-

$$Best_{alg} = MAX(P_{\%acc,f}) \ldots\ldots\ldots\ldots (16)$$

V. RESULTS AND DISCUSSION

After applying model into a ABBANK share of DSE data, it returns best prediction accuracy for Feed- Forward Neural Network algorithm in polynomial trend which is about 90%.

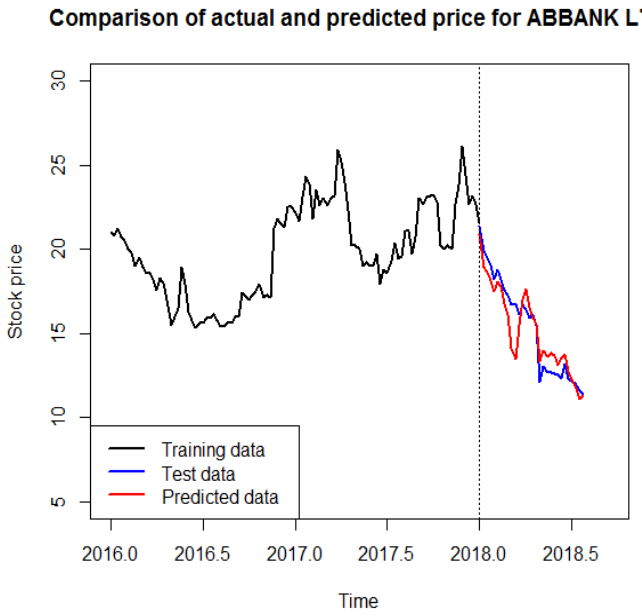

Figure 2: Plotting of actual price of ABBANK share with predicted price by using FFNN.

Again, from 35 stocks, FFNN shows best in 14 stocks, Linear model shows best in 7 stocks, Holt-Winters in exponential smoothing & ARIMA (1,0,0) shows best in 4 stocks, Holt-Winters & ARIMA (best fit model) shows best in 3 stocks.

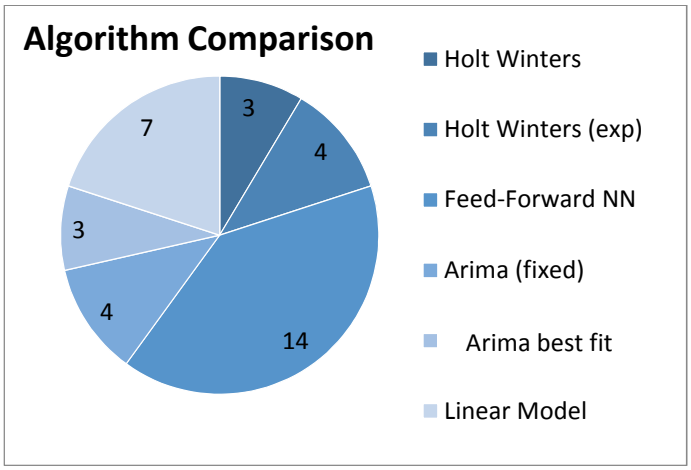

Figure 3: Comparison of algorithms over 35 DSE stocks.

In overall, the prediction accuracy is 82.1% for ARIMA (1,0,0), 75.7% for ARIMA (best fit model), 83.3% for FFNN, 75.9% for Holt-Winters, 75.4% for Holt-Winters with exponential smoothing, 81.9% for Linear model in average.

Table 1: Prediction accuracy of algorithms in different trend

Algorithm	Polynomial	STL	Actual	Average
ARIMA (1,0,0)	87.54	74.88	83.90	82.11
ARIMA (best fit)	66.41	76.61	84.06	75.70
FFNN	80.75	85.70	83.45	83.30
Holt-Winters	75.15	74.28	78.12	75.85
Holt-Winters (exp)	66.81	76.61	82.87	75.43
Linear	81.58	82.68	81.33	81.86

VI. CONCLUSION

From the analysis of this report, it can be said that, the stock of Dhaka Stock Exchange (DSE), Bangladesh falls under forecasting with time series analysis. The average prediction accuracy of all algorithms is greater than 75%. Whereas Feed-Forward neural network shows maximum average accuracy (>83%), ARIMA first order auto regressive model (1,0,0) shows about 82% accuracy in average.

Moreover, while considering a single stock, the prediction accuracy of the stock is shown around 93% and often it exceeds 95% for few stocks.

After all, Time series forecasting is really good for forecasting of Bangladeshi stock based on the analysis from this report however 100% gain couldn't achieve due to the randomness and complicated behavior of the stock prices.

REFERENCES

[1] ASM Shakil Haider, " Forecasting Dhaka Stock Exchange (DSE) return: An Autoregressive Integrated Moving Average (ARIMA) approach", 2009, North South Business Review, Volume 3, Number 1, ISSN 1991-4938.

[2] Tarun Kanti Bose , Md. Reaz Uddin, and Md. Wahidul Islam, "Measuring and Comparing the Efficiency of Dhaka Stock Exchange and Chittagong Stock Exchange", 2014, International Journal of Scientific and Research Publications, Volume 4, Issue 3,1 ISSN 2250-3153

[3] Md. Kamruzzaman, Md. Mohsan Khudri, and Md. Matiar Rahman, "Modeling and predicting stock market returns: A case study on Dhaka stock exchange of Bangladesh", 2017, Dhaka Univ. J. Sci. 65(2): 97-101

[4] Ayodele A. Adebiyi, Aderemi O. Adewumi, and Charles K. Ayo, "Stock Price Prediction Using the ARIMA Model", 2014, UKSim-AMSS 16th International Conference on Computer Modelling and Simulation.

[5] Ashish Sharma, Dashish Sharma, Dinesh Bhuriya and Upendra Singh, "Survey of Stock Market Prediction Using Machine Learning Approach", 2017, International Conference on Electronics, Communication and Aerospace Technology, ICECA 2017, doi:10.1109/iceca.2017.8212715

[6] Poonam Somani, Shreyas Talele, and Suraj Sawant "Stock Market Prediction Using Hidden Markov Model", 2014, IEEE 7th Joint International Information Technology and Artificial Intelligence Conference. doi:10.1109/itaic.2014.7065011

[7] Md. Rafiul Hassan and Baikunth Nath, "Stock Market Forecasting Using Hidden Markov Model: A New Approach",2005, 5th International Conference on Intelligent Systems Design and Applications (ISDA'05)".

[8] Kumar abhishek, Anshul Khairwa, Tej Pratap, and Surya Prakash "A stock market prediction model using Artificial Neural Network", 2012, Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12). doi:10.1109/icccnt.2012.6396089

[9] Prajakta S. Kalekar "Time series Forecasting using Holt-Winters Exponential Smoothing "

[10] http://astrostatistics.psu.edu/su07/R/html/stats/html/HoltWinters.html

[11] Hyndman, Rob J; Athanasopoulos, George. "8.9 Seasonal ARIMA models: Forecasting: principles and practice"

[12] https://people.duke.edu/~rnau/411arim.htm

[13] Hilary L. Seal, "The historical development of the Gauss linear model", 1967, Biometrika. 54 (1/2): 1–24. doi:10.1093/biomet/54.1-2.1. JSTOR 2333849.

[14] Behrouz Safarinejadian, Mohammad Amin Tajeddini, and Abdolrahman Ramezani , "Predict time series using extended, unscented, and cubature kalman filters based on feed-forward neural network algorithm",2013, 3rd International Conference on Control, Instrumentation, and Automation (ICCIA 2013) , Tehran, Iran.