Grenze International Journal of Engineering and Technology, June Issue

Financial Time Series Data Forecasting for Event Handling

Pabitra Kumar Tripathy[1] and Sanjaya Kumar[2]
[1-2]Kalinga University, Raipur Chhattisgarh

***Abstract*—The volatile nature of the financial market makes it difficult to develop a reliable forecasting system for financial time series data, but it becomes unreliable when the financial market fluctuates unexpectedly due to events such as demonetization, earthquakes, or other government financial decisions. Financial market analysis based on various events is very crucial task due to drastic ups and down and more error prone especially for next day ahead forecasting. Financial time series data are affected by various factors and events, some of which influence prices of financial market directly and others that do so indirectly. Financial time series data can be affected by two different types of events: expected and unexpected events. Expected events are those which are almost known to everyone, such as a presidential election. On the other hand, an unexpected event occurs suddenly and without warning, such as natural disasters, wars, demonetization and other political activities. This study focuses on the creation of a resilient financial time series forecasting system based on ANN approaches.**

I. Development oF Robust Forecasting Model for US Precedential Election

Sudden fluctuations in the stock market always need to be incorporated by stock price forecasting models. The traditional statistical models can't adapt or incorporate these variations. On the other hand, ML techniques like ANNs are self-capable to adjust these fluctuations and able to produce comparatively better outputs. This section explores intelligent techniques to check the effectiveness of models developed for US presidential election.

II. Problem Statement

This section explores two Artificial Neural Network (ANN) techniques: Radial Basis Function Network (RBFN) and Error Back Propagation Network (BPN) for stock data forecasting with a particular reference to event analysis. The primary motive of the proposed research work is to check the robustness and efficiency of ANN (Laboissiere et al., 2015; Nawi et al., 2013) based models, developed from historical stock data before occurring events and validated through stock data after the events. Researchers are focusing on developing forecasting models based on different preprocessing and ANN techniques. Authors have also analyzed various events using many statistical techniques. Many authors have investigated the effect of the different events on the financial market (Chien et al., 2014; J. Z. Wang et al., 2011) of different countries. This work analyzes the US Presidential election using ANN techniques, however, very few papers related to event analysis using ANN techniques was found. Very few papers are available in which intelligent techniques were used to analyze the effect of events. These forecasting models are able to adjust the stock trading changes during the event and are robust enough to incorporate sudden fluctuations of the stock market with a high value of Mean Absolute Percentage Error (MAPE).

Grenze ID: 01.GIJET.8.2.506_2

III. FORECASTING MODEL

In this event analysis, about six years' time series data from January 2011 to March 2017 of DOW30 and NASDAQ100 are collected from the financial site: www.yahoofinance.com. These data also contain data of presidential elections in the USA in 2012 and 2016. Data are normalized by dividing each sample with the highest stock price to achieve smooth convergence during learning through nonlinear time series data as explained in

$$X_{new} = \frac{x}{x_{max}}$$

To check the robustness and efficiency of the developed models, these stock data are divided into three parts: training, testing, and validation. The stock data before the event is divided into training (80%) and testing (20%) to develop forecasting models and data after the event is considered for model validation as shown in Table I.

TABLE I: DATA PARTITION

Partition Name	**Data Range**	
	DOW30	**NASDAQ100**
Training	11-Jan-2016 to 10-Nov-2020	11-Jan-2016 to 10-Nov-2020
Testing	o 08-Nov- 2021	11-Nov-2020 to 08-Nov-2021
Validation	o 15-Mar- 2021	09-Nov-2020 to 15-Mar-2021

A. Result Analysis

Experimental work is carried out with self-written MATLAB code for RBFN and BPN under the Window 10 environment. Back Propagation Network (BPN) is one of the machines learning technique to train the model for time series forecasting. The predefined functions newff() is used to create a two-layer feed-forward network with number of neurons (20) in its hidden layer for a given a matrix of input vectors p and target vectors t, presented in equation 1.1 and Figure 1.2 given below:

net = newff (p,t, 20); (1.1)

The above equation 1.1 is used to add neurons to hidden layer and to simulate the work of a radial basis network until it meets the specified mean squared error goal.

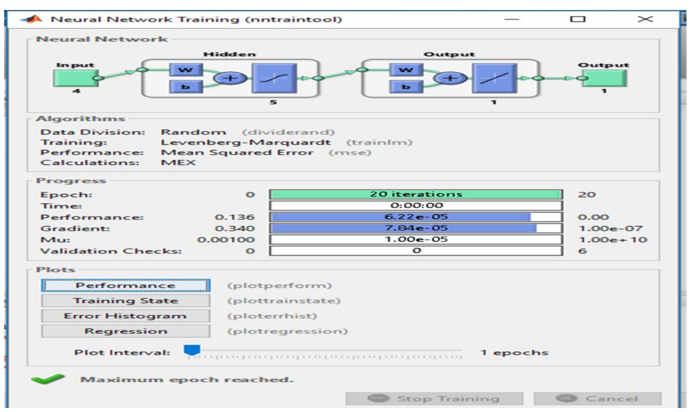


Fig. 1.2: MATLAB generated training model of neural network

A predefined function of RBFN in MATLAB as newrb() is used to create basic RBFN architecture by adding neurons to the hidden layer and also simulates RBFN until it meets the specified mean squared error (MSE) goal by feeding input and target vectors as parameters of the function. Similarly, BPN was also constructed with MATLAB function newff() and was simulated with its default parameters like learning rate, momentum, etc.

Normalized training and testing samples were presented to RBFN and BPN to train and test the models; further validation data were used for N-days ahead forecasting such as: 1-Day, 3-Days, 5-Days, and 7-Days. Predicted values for both indices, DOW30 and NASDAQ100, are compared with actual next-day-close prices regarding Mean Absolute Percentage Error (MAPE).

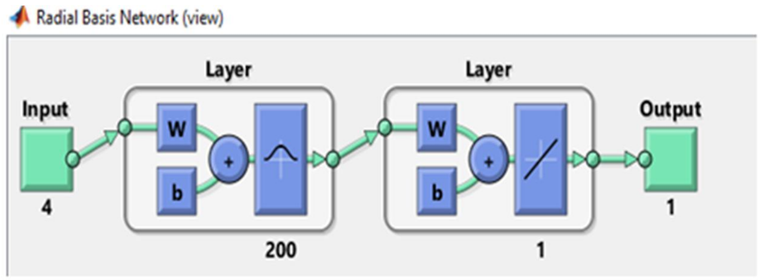


Fig. 1.3: MATLAB generated view of RBFN through newrb() function

Results are presented in Table II and Table III for DOW30 and NASDAQ100 respectively. Results are also analyzed deeply to check the overall performance of the models in three different viewpoints as shown in Table :II

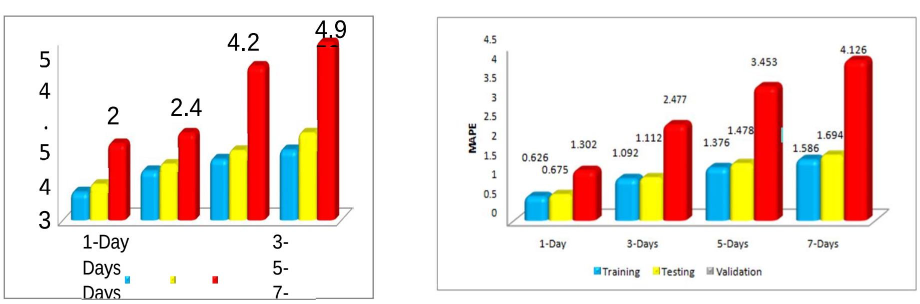
TABLE II. COMPARATIVE MAPE ON TRAINING, TESTING AND VALIDATION SAMPLES FOR DOW30 INDEX DATA WITH N-DAYS AHEAD FORECASTING

Day(s) ahead forecasting	RBFN			BPN		
	Training	Testing	Validation	Training	Testing	Validation
1-Day	0.626	0.675	1.302	0.625	0.645	1.280
3-Days	1.092	1.112	2.477	1.084	1.118	2.825
5-Days	1.376	1.478	3.454	1.377	1.658	4.218
7-Days	1.586	1.694	4.126	1.565	1.800	4.765

TABLE III: COMPARATIVE MAPE ON TRAINING, TESTING AND VALIDATION SAMPLES FOR NASDAQ100 INDEX DATA WITH N-DAYS AHEAD FORECASTING

Day(s) ahead forecasting	RBFN			BPN		
	Training	Testing	Validation	Training	Testing	Validation
1-Day	0.775	0.978	2.123	0.782	1.036	2.517
3-Days	1.367	1.539	2.417	1.427	2.358	5.972
5-Days	1.699	1.924	4.279	1.655	2.578	6.600
7-Days	1.945	2.398	4.953	1.772	2.650	6.587

Case 1: **Comparative analysis in terms of data partitions-** Typically performance of the models varies at different stages of training, testing, and validation. The performance of the models in the case of known target samples (Training stage) is always higher than the models with unseen samples (Testing and validation stage). The performance of the models decreases for the higher value of N in the case of N- Days forecasting. Comparative graphs as shown in Figure 1.4 and 1.5 reflects these facts where training MAPE is lower than testing MAPE followed by validation
MAPE for both the indices using both the techniques. On the other hand, MAPE is continuously increasing while the value of N is increasing.



(a) (b)

Fig. 1.4: Comparative analysis in terms of data partition using RBFN for (a) DOW30 (b) NASDAQ100

Case 2: Comparative analysis of RBFN and BPN- Comparative MAPE regarding N-days ahead fore casting for event related data (Validation Data)is shown in Figure 1.5 which reflects that the RBFN is performing better to incorporate event related changes in the stock market than BPN. It has been observed in many cases that BPN sometimes perform better than RBFN and it all depends upon the nonlinear trends of stock data.

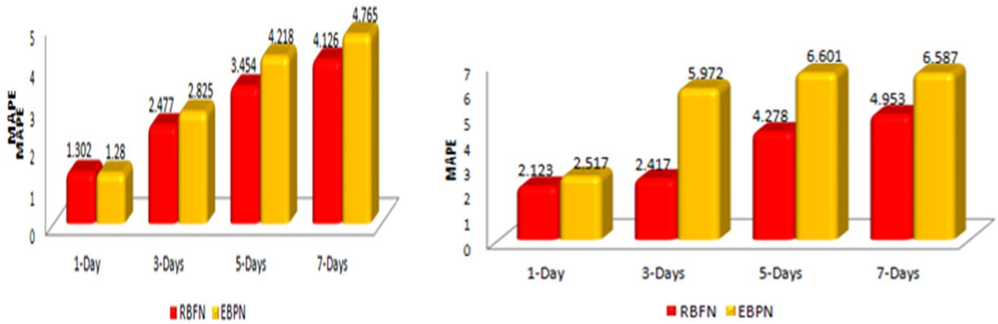


Fig. 1.5: Comparative graph in between RBFN and BPN for (a) DOW30 (b) NASDAQ100

Case 3: Comparative analysis of DOW30 and NASDAQ100 stock trends during event- This section analyzes how the developed models are self-sustainable and capable during any expected or unexpected events. As shown in Figure 1.6, the models are performing better in the case of DOW30 Index data for both RBFN and BPN as compare to NASDAQ100 for N-days ahead forecasting but in both the cases models are capable of adjusting to sudden fluctuations of the stock market.

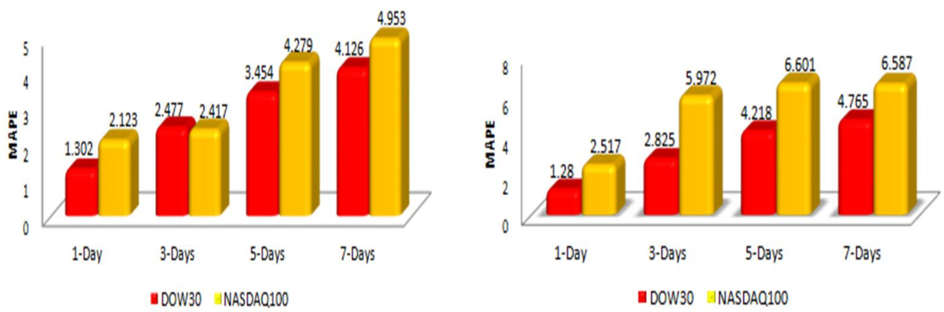


Fig. 1.6: Comparative graph of stock data of DOW30 and NASDAQ100 for (a) RBFN and BPN

IV. Conclusion

This article focus on development of a robust financial time series forecasting system based on ANN techniques to incorporate any unexpected event like demonetization in India and Presidential election in US occurred from time to time. Sudden fluctuations in the stock market always need to be incorporated by stock price forecasting models. The traditional statistical models can't adapt or incorporate these variations, on the other hand, ML techniques like ANNs are self-capable to adjust these fluctuations and able to produce comparatively better outputs. This work analyzes how ML based models are robust when events related data are presented. Experimental results drawn with the help of self written MATLAB codes stabilizes the following facts:

- This research work considered two ANN techniques where RBFN is outperforming BPN in all cases.
- RBFN model with 10-fold cross validation is best among all other models and can be considered as robust financial time series forecasting system as MAPE obtained at testing stages are better, which concludes that models developed with validation set has the capability to incorporate the sudden fall of stock market and FX rate due to unexpected event like demonetization.

References

[1] Khashei, M., Torbat, S., &Rahimi, Z. H. (2017). An Enhanced Neural-based Bi- Component Hybrid Model for Foreign Exchange Rate. Turkish Journal of Forecasting, 01(1), 16–29.

[2] Weng, B., Lu, L., Wang, X., Megahed, F. M., & Martinez, W. (2018). Predicting short-term stock prices using ensemble methods and online data sources. Expert Systems with Applications, 112, 258–273.
[3] Panda, C., &Narasimhan, V. (2007). Forecasting exchange rate better with artificial neural network. Journal of Policy Modeling, 29(2), 227–236.
[4] Hota, H. . S., Shrivas, A. K., &Hota, R. (2018). ScienceDirectAn Ensemble Model for Detecting Phishing Attack with Proposed Remove-Replace Feature Selection Technique. International Conference on Computational Intelligence and Data Science (ICCIDS 2018), 132, 900–907.
[5] Saigal, S., &Mehrotra, D. (2012). Performance Comparison of Time Series Data Using Predictive Data Mining Techniques. Advances in Information Mining, 4(1),57–66.
[6] Schapire, R. E. (1999). A Brief Introduction to Boosting Analyzing the training error. Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, 1–6.
[7] Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., &Soman, K. P. (2017). Stock price forecasting using LSTM, RNN and CNN-sliding window model. 2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017, 2017, 1643–1647.