JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Stock Closing Price Prediction Using Various Deep Learning Techniques

Neha Patidar
Research Scholar, Faculty of IT and CS
Parul University,
Vadodara, Gujarat

Dr. Harshal Shah
Associate Professor, Parul Institute of Technology,
Faculty of Engineering and Technology, Parul University,
Vadodara, Gujarat

Abstract

Accurate prediction of stock market returns is a very challenging task due to volatile and non-linear nature of the financial stock markets. With the introduction of artificial intelligence and increased computational capabilities, programmed methods of pre- diction have proved to be more efficient in predicting stock prices. In this work, Artificial Neural Network and Random Forest techniques have been utilized for predicting the next day closing price for five companies belonging to different sectors of operation. The financial data: Open, High, Low and Close prices of stock are used for creating new variables which are used as inputs to the model. The models are evaluated using standard strategic indicators: RMSE and MAPE. The low values of these two indicators show that the models are efficient in predicting stock closing price.

Keywords: Random Forest Regression; Artificial Neural Network; Stock market prediction

1. Introduction

Stock market is characterized as dynamic, unpredictable and non-linear in nature. Predicting stock prices is a challenging task as it depends on various factors including but not limited to political conditions, global economy, company's financial reports and performance etc. Thus, to maximize the profit and minimize the losses, techniques to predict values of the stock in advance by analyzing the trend over the last few years, could prove to be highly useful for making stock market movements [1] [2]. Traditionally, two main approaches have been proposed for predicting the stock price of an organization. Technical analysis method uses historical price of stocks like closing and opening price, volume traded, adjacent close values etc. of the stock for predicting the future price of the stock. The second type of analysis is qualitative, which is performed on the basis of external factors like company profile, market situation, political and economic factors, and textual information in the form of financial new articles, social media and even blogs by economic analyst [3]. Now a days, advanced intelligent techniques based on either technical or fundamental analysis are used for predicting stock prices. Particularly, for stock market analysis, the data size is huge and also non-linear. To deal with this variety of data efficient model is needed that can identify the hidden patterns and complex relations in this large data set. Machine learning techniques in this area have proved to improve efficiencies by 60-86 percent as compared to the past methods [4].

Most of the previous work in this area use classical algorithms like linear regression [5], Random Walk Theory (RWT) [6], Moving Average Convergence / Divergence (MACD) [7] and also using some linear models like Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA) [8], for predicting stock prices. Recent work shows that stock market prediction can be enhanced using machine learning. Techniques such as Support Vector Machine (SVM), Random Forest (RF) [10]. Some techniques based on neural networks such as Ar- tificial Neural Network (ANN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and

deep neural networks like Long Short Term Memory (LSTM) also have shown promising results [4] [11].
ANN is capable for finding hidden features through a self learning process. These are good approximators and are able to find the input and output relationship of a very large complex dataset. Thus, ANN proves to be a good choice for predicting stock price for an organization. Selvin et al predicted stock price of NSE listed companies by a comparative analysis of different Deep learning techniques [13]. Hamzaebi et al. experimented multi-periodic stock market forecasting using iterative and directive methods like ANN model [14]. Rout et al. predicted stock market using a low complex RNN model and tested it Bombay stock exchange and S & P 500 index dataset [15]. Roman et al. applied RNN models on stock market data of five countries: Canada, Hong Kong, Japan, UK and USA, to train the networks and then these networks were used to predict the trend in stock returns [17]. In 2014, Yunus et al. applied ANN on NASDAQ to predict the closing price of stock [16]. Mizuno et al. used ANN to perform technical analysis on TOPIX dataset and its application to the buying and selling timing prediction system [18]. Some works have been proposed which use Random Forest (RF) for forecasting purposes. RF is an ensemble technique. It is normally capable of performing both regression and classification tasks. It operates by constructing multiple decision trees at training time which outputs mean regression of individual decision trees [19]. Mei et al. employed RF for accurately forecasting real time prices on New York electricity market [20]. Similar work is done by Yand et al. who performed short term load forecasting in the operation of electrical power systems using RF model [21]. Herrera et al. used RF as a predictive model for forecasting hourly urban water demand [22]. In this work, two techniques i.e. ANN and RF have been used for predicting the closing price of an organization. The models use a set of new variables created using the financial dataset with Open, High, Low and Close of a particular company. These new indicators will pay a crucial role in terms of improved accuracy of the models in predicting the next day closing price of a particular company. The effectiveness of the models is tested using two performance measures: RMSE and MAPE.
The remainder of the papers is as follows: Section 2 provides methodology of techniques applied. Section 3 discusses the result with section 4 concluding the paper.

2. Methodology

Description of Data

The historical data for the five companies has been collected from Yahoo Finance [23]. The dataset includes 10 year data from 4/5/2009 to 4/5/2019 of Nike, Goldman Sachs, Johnson and Johnson, Pfizer and JP Morgan Chase and Co. The data contains information about the stock such as High, Low, Open, Close, Adjacent close and Volume. Only the day-wise closing price of the stock has been extracted. Table 2 shows statistics of the dataset that is used for training and testing.

	Dataset	Training Dataset	Testing Dataset
Time Interval	04/05/2009 – 04/05/2019	04/06/2009- 04/03/2017	04/04/2017 – 04/05/2019

Table 1: Statistics of the dataset

New Variables

Six new variables have been created for the prediction of stock closing price. These variables have been used to train the model. The new variables are as follows:

1. Stock High minus Low price (H-L)
2. Stock Close minus Open price (O-C)
3. Stock price's seven days' moving average (7 DAYS MA)
4. Stock price's fourteen days' moving average (14 DAYS MA)
5. Stock price's twenty one days' moving average (21 DAYS MA)
6. Stock price's standard deviation for the past seven days (7 DAYS STD DEV)

Artificial Neural Network

ANN, is one of the intelligent data mining techniques that identify a fundamental trend from data and to generalize from it. ANN is capable of simulating and analysing complex patterns in unstructured data as compared to most of the conventional methods. The model uses the basic structure of Neural Network having neurons with different layers. The model works with three layers. It consists of input layer, hidden layer and the output layer. The input layer consists of new variables which are H-L, O-C, and 7 DAYS MA, 14 DAYS MA, 21 DAYS MA, 7 DAYS STD DEV and Volume [23]. The weights on each input load is multiplied and added and sent to the neurons. The hidden layer or the activation layer consists of these neurons. The total weight is calculated and is moved to the third layer which is the output layer. The output layer consists of only one neuron which will give the predicted value in terms of closing price of the stock. The Fig. 1 shows a detailed representation of ANN architecture with the new variables acting as input.

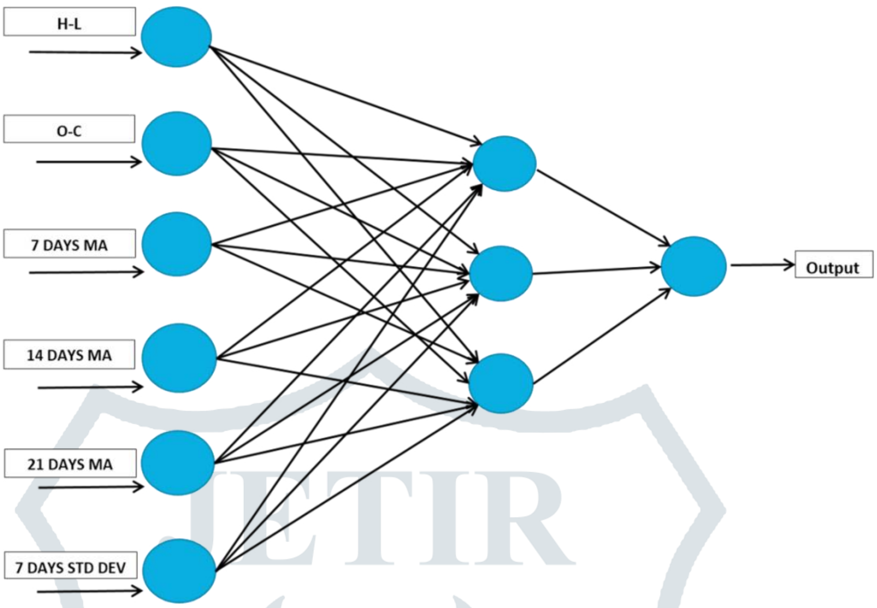

Fig. 1. Detailed architecture of Artificial Neural Network (ANN) for stock price prediction

Random Forest

Random Forest (RF) is an ensemble machine learning technique. It is capable of performing both regression and classification tasks. The idea is to combine multiple decision trees in order to determine the final output instead of relying on individual decision trees which in order reduce the variance in the model. In this work, new created variables are provided for the training of each decision tree which in turn determines the decision at the nodes of the tree. The noise in stock market data is usually quite high because of its huge size and can cause the trees to grow in a completely different manner as compared to the expected growth. It aims at minimizing forecasting error by treating the stock market analysis as a classification problem and based on training variables predicted the next day closing price of the stock for a particular company.

1. Results and Conclusions

To evaluate the effectiveness of the models, a comparison is made between the two techniques on five different sector companies namely, JP Morgan, Nike, Johnson and Johnson, Goldman Sachs and Pfizer using both ANN and RF models. Predicted closing prices are subjected to Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and Mean Bias Error (MBE) for finding the final minimalized errors in the predicted price.

RMSE is computed using eq. 1.

$$\text{RSME}=\sqrt[2]{\textstyle\sum_{i=1}^{n}(Oi-Fi)2/n} \qquad (1)$$

where 'O_i' refers to the original closing price, 'F_i' refers to the predicted closing price and 'n' refers to the total window size. MAPE has also been used to evaluate the performance of the model and is computed using eq. 2.

$$\text{MAPE}=\frac{1}{n}\textstyle\sum_{i=1}^{n}((Oi-Fi)/Oi)*100 \qquad (2)$$

where 'O_i' refers to the original closing price, 'F_i' refers to the predicted closing price and 'n' refers to the total window size. MBE has also been used to evaluate the performance of the model and is computed using eq. 3.

$$\text{MAPE}=\frac{1}{n}\sum_{i=1}^{n}((Oi-Fi)/Oi) \tag{3}$$

where 'O_i' refers to the original closing price, 'F_i' refers to the predicted closing price and 'n' refers to the total window size. Fig. 2 represents graphs showing original closing price of stock with respect to predicted closing price of stock of five different companies using ANN. Fig. 3 represents graphs showing original closing price of stock vs predicted closing price of stocks using RF. Comparative analysis of the RMSE, MAPE and MBE values obtained using ANN and RF model is shown in Table 2, it can be observed that ANN shows better prediction results for stock prices.

where 'O_i' refers to the original closing price, 'F_i' refers to the predicted closing price and 'n' refers to the total window size. Fig. 2 represents graphs showing original closing price of stock with respect to predicted closing price of stock of five different companies using ANN. Fig. 3 represents graphs showing original closing price of stock vs predicted closing price of stocks using RF. Comparative analysis of the RMSE, MAPE and MBE values obtained using ANN and RF model is shown in Table 2, it can be observed that ANN shows better prediction results for stock prices.

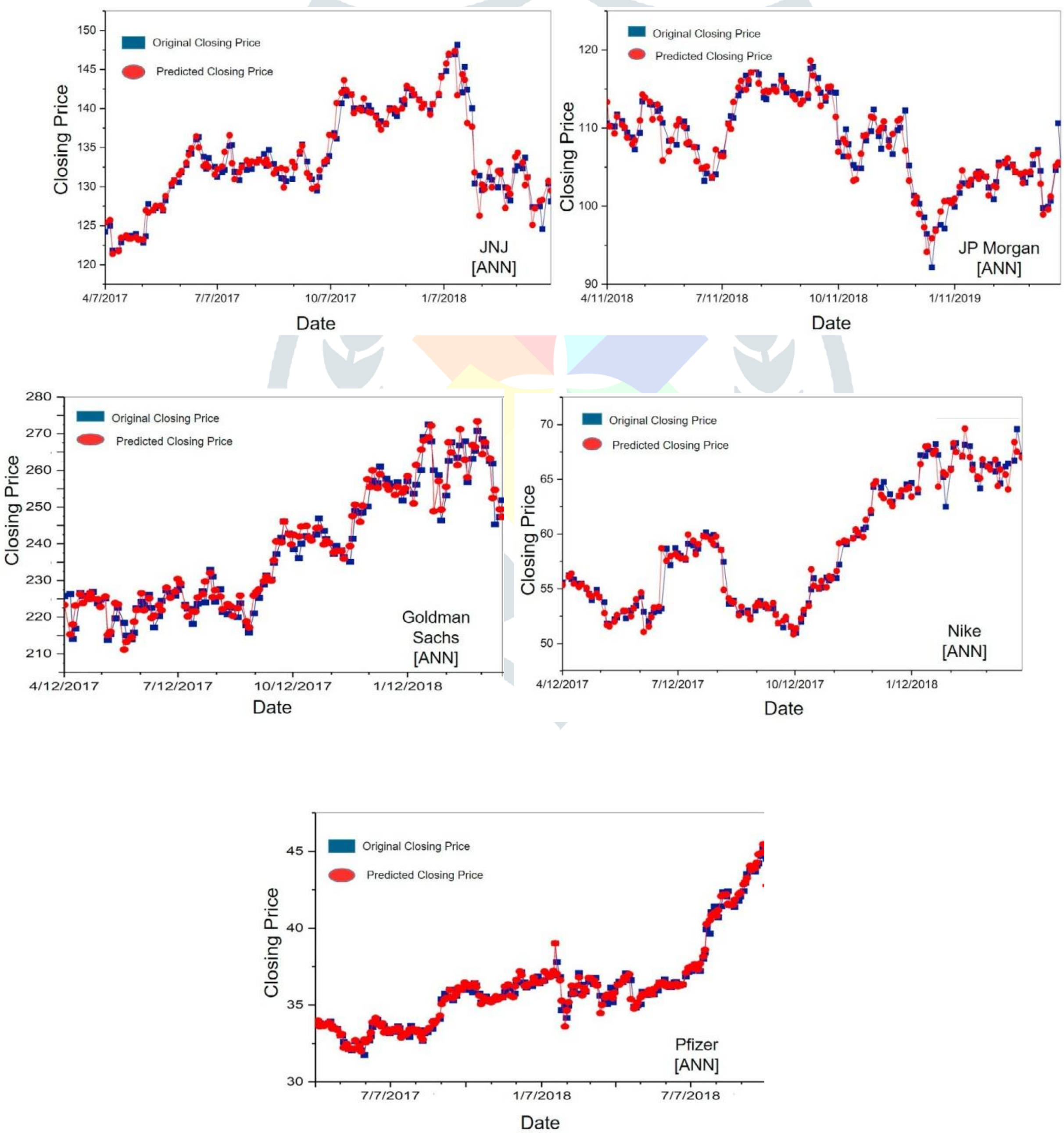

Fig. 2. Predicted v/s original (expected) closing stock price using ANN.

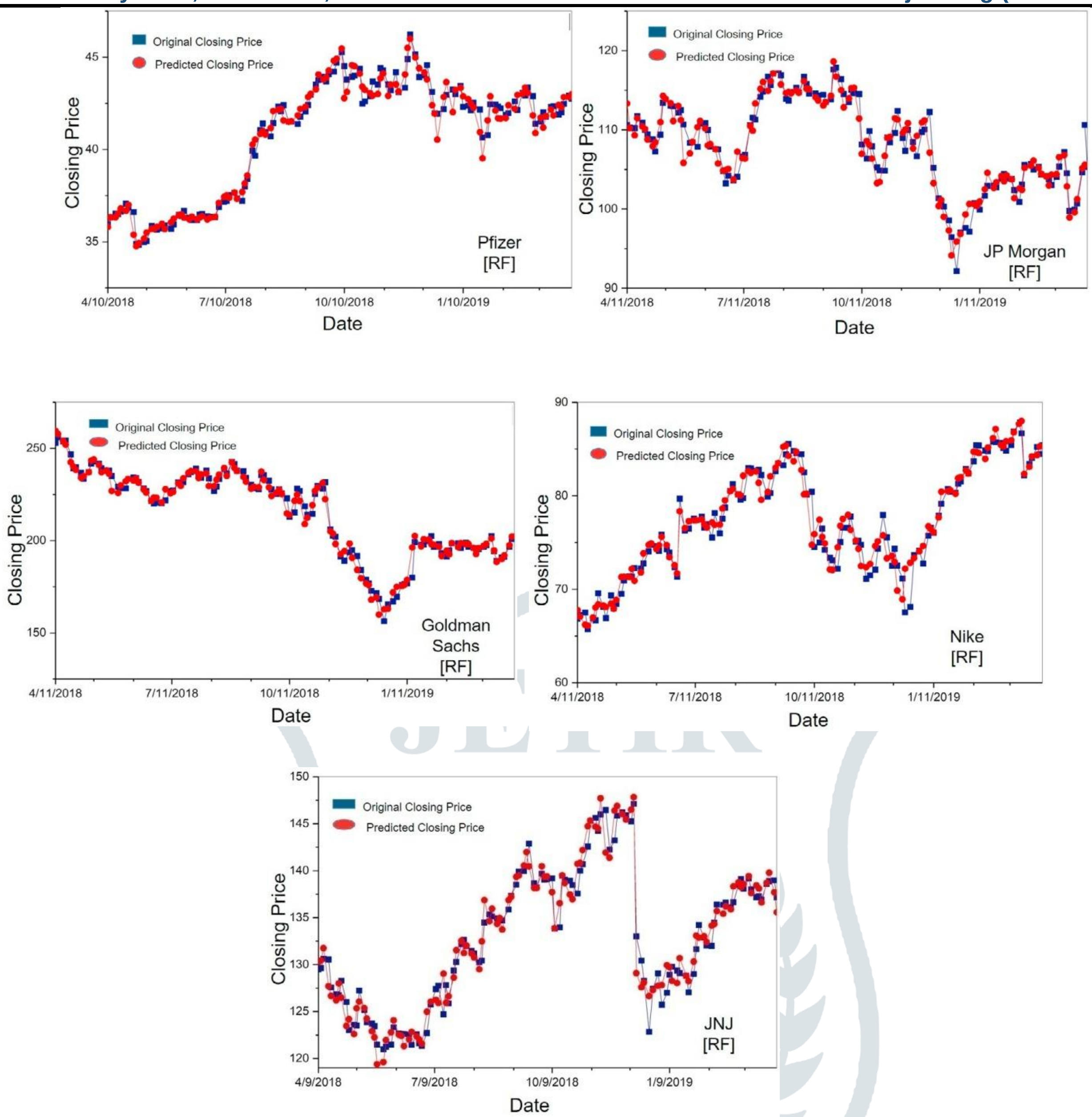

Fig. 3. Predicted v/s original (expected) closing stock price using RF.

Company	*ANN*			*RF*		
	RSME	*MAPE*	*MBE*	*RSME*	*MAPE*	*MBE*
Nike	1.10	1.07%	-0.0522	1.29	1.14%	-0.0521
Goldman Sachs	3.30	1.09%	0.0762	3.40	1.01%	0.0761
JP Morgan and Co.	1.28	0.89%	-0.0310	1.41	0.93%	-0.0313
Johnson and Johnson	1.54	0.70%	-0.0138	1.53	0.75%	-0.0138
Pfizer Inc.	0.42	0.77%	-0.0156	0.43	0.8%	-0.0155

Table 2. Comparative analysis of RMSE, MAPE and MBE values obtained using ANN and RF models.

The comparative analysis indicates, that for Nike, JP Morgan and Co., Johnson & Johnson and Pfizer Inc.

companies, ANN proves to be a better technique, giving better RMSE and MAPE values, as shown in the Table 2.

3. Conclusion and Future Scope

Predicting stock market returns is a challenging task due to consistently changing stock values which are dependent on multiple parameters which form complex patterns. The historical dataset available on company's website consists of only few features like high, low, open, close, adjacent close value of stock prices, volume of shares traded etc., which are not sufficient enough. To obtain higher accuracy in the predicted price value new variables have been created using the existing variables. ANN is used for predicting the next day closing price of the stock and for a comparative analysis, RF is also implemented. The comparative analysis based on RMSE, MAPE and MBE values clearly indicate that ANN gives better prediction of stock prices as compared to RF. Results show that the best values obtained by ANN model gives RMSE (0.42), MAPE (0.77) and MBE (0.013). For future work, deep learning models could be developed which consider financial news articles along with financial parameters such as a closing price, traded volume, profit and loss statements etc., for possibly better results.

References

[1] Masoud, Najeb MH. (2017) "The impact of stock market performance upon economic growth." International Journal of Economics and Financial Issues ***3 (4)*** *: 788–798.*

[2] Murkute, Amod, and Tanuja Sarode. (2015) "Forecasting market price of stock using artificial neural network." International Journal of Computer Applications ***124 (12)*** *: 11-15.*

[3] Hur, Jung, Manoj Raj, and Yohanes E. Riyanto. (2006) "Finance and trade: A cross-country empirical analysis on the impact of financial development and asset tangibility on international trade." World Development ***34 (10)*** *: 1728-1741.*

[4] Li, Lei, Yabin Wu, Yihang Ou, Qi Li, Yanquan Zhou, and Daoxin Chen. (2017) "Research on machine learning algorithms and feature extraction for time series." IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC): 1-5.

[5] *Seber, George AF and Lee, Alan J. (2012) "Linear regression analysis." John Wiley & Sons* ***329***

[6] Reichek, Nathaniel, and Richard B. Devereux. (1982) "Reliable estimation of peak left ventricular systolic pressure by M-mode echographic- determined end-diastolic relative wall thickness: identification of severe valvular aortic stenosis in adult patients." American heart journal ***103 (2)*** *: 202-209.*

[7] Chong, Terence Tai-Leung, and Wing-Kam Ng. (2008) "Technical analysis and the London stock exchange: testing the MACD and RSI rules using the FT30." Applied Economics Letters ***15 (14)*** *: 1111-1114.*

[8] Zhang, G. Peter. (2003) "Time series forecasting using a hybrid ARIMA and neural network mode." Neurocomputing 50 : 159-175.

[9] Suykens, Johan AK, and Joos Vandewalle. (1999) "Least squares support vector machine classifiers." Neural processing letters ***9 (3)*** *: 293-300.*

[10] Liaw, Andy, and Matthew Wiener. (2002) "Classification and regression by Random Forest." R news ***2 (3)*** *: 18-22.*

[11] Oyeyemi, Elijah O., Lee-Anne McKinnell, and Allon WV Poole. (2007) "Neural network-based prediction techniques for global modeling of M (3000) F2 ionospheric parameter." Advances in Space Research ***39 (5)*** *: 643-650.*

[12] Huang, Wei, Yoshiteru Nakamori, and Shou-Yang Wang. (2005) "Forecasting stock market movement direction with support vector machine."
Computers & Operations Research ***32 (10)*** *: 2513-2522.*

[13] Selvin, Sreelekshmy, R. Vinayakumar, E. A. Gopalakrishnan, Vijay Krishna Menon, and K. P. Soman. (2017) "Stock price prediction us- ing LSTM, RNN and CNN-sliding window mode." International Conference on Advances in Computing, Communications and Informatics (ICACCI): 1643-1647.

[14] Hamzaçebi, Coşkun, Diyar Akay, and Fevzi Kutay. (2009) "Comparison of direct and iterative artificial neural network forecast approaches in multi-periodic time series forecasting." Expert Systems with Applications ***36 (2)*** *: 3839-3844.*

[15] Rout, Ajit Kumar, P. K. Dash, Rajashree Dash, and Ranjeeta Bisoi. (2017) "Forecasting financial time series

using a low complexity recurrent neural network and evolutionary learning approach." Journal of King Saud University-Computer and Information Sciences ***29 (4)*** *: 536-552.*

[16] Yetis, Yunus, Halid Kaplan, and Mo Jamshidi. (2014) "Stock market prediction by using artificial neural network." in2014 World Automation Congress (WAC): 718-722.

[17] Roman, Jovina, and Akhtar Jameel. (1996) "Backpropagation and recurrent neural networks in financial analysis of multiple stock market return." Proceedings of HICSS-29: 29th Hawaii International Conference on System Sciences 2: 454-460.

[18] Mizuno, Hirotaka, Michitaka Kosaka, Hiroshi Yajima, and Norihisa Komoda. (1998) "Application of neural network to technical analysis of stock market prediction." Studies in Informatic and control ***7 (3)*** *: 111-120.*

[19] Kumar, Manish, and M. Thenmozhi. (2006) "Forecasting stock index movement: A comparison of support vector machines and random forest"

In Indian institute of capital markets 9th capital markets conference paper

[20] Mei, Jie, Dawei He, Ronald Harley, Thomas Habetler, and Guannan Qu. (2014) "A random forest method for real-time price forecasting in New York electricity market." IEEE PES General Meeting Conference & Exposition: 1-5.

[21] Herrera, Manuel, Luís Torgo, Joaquín Izquierdo, and Rafael Pérez-García. (2010) "Predictive models for forecasting hourly urban water demand." Journal of hydrology ***387 (1-2)*** *: 141-150.*

[22] Dicle, Mehmet F., and John Levendis. (2011) "Importing financial data." The Stata Journal ***14 (4)*** *: 620-626.*

[23] Yahoo Finance - Business Finance Stock Market News, <https://in.finance.yahoo.com/> [Accessed on August 16,2018]

[24] Zurada, Jacek M. (1992) "Introduction to artificial neural systems." St. Paul: West publishing company 8.