An Optimized Gradient Descent Model for Forecasting Stock Movement Trends

Gaurav Mourya[1], Prof. Preetish Kshirsagar[2]

***Abstract*— Stock Market prediction is a category of time series prediction which extremely challenging due to the dependence of stock prices on several financial, socio-economic and political parameters etc. Moreover, small inaccuracies in stock market price predictions may result in huge losses to firms which use stock market price prediction results for financial analysis and investments. Off late, artificial intelligence and machine learning based techniques are being used widely for stock market prediction due to relatively higher accuracy compared to conventional statistical techniques. The proposed work employs the steepest descent based algorithm along with the data pre-processing using the discrete wavelet transform (DWT) for stock market prediction. It has been shown that the proposed system attains lesser mean square percentage error compared to previously existing technique.**

Keywords—Stock Market Forecasting, Artificial Neural Network (ANN), Back Propagation, Discrete Wavelet Transform (DWT), Mean Absolute Percentage Error (MAPE).

I. INTRODUCTION

Stock Markets have long remained one of the one avenues on the forefront which is crucial for the operations of most of the top most companies [1]. Several decisions pertaining to investments, shares etc. depend on the behavior of the stocks of a company. The stock price values are often leveraged by financial and investment firms for gaining profits and investing [2]. However, the volatile nature of the stock markets make it a risky proposition. Therefore, estimating the future trends in stock prices is somewhat mandatory for firms analyzing stock prices and aiming to gain leverage. This calls for stock market forecasting or stock market prediction. Stock market prediction is basically a time series prediction problem. Mathematically:

$$\boldsymbol{P = f(t, v)} \quad \textbf{(1)}$$

Here,
P represents stock price
f represents a function of
t is the time variable
v are other influencing global variables

The dependence of stock process over time makes it somewhat predictable under similar other conditions of global influencing variables. However, even the slightest of changes can derail the prediction completely [3].

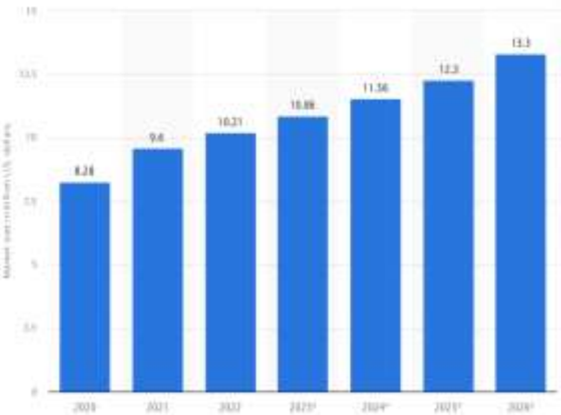

Fig.1 Online Stock Trading Market Business Trends

Statistical techniques are not found to be as accurate as the contemporary artificial intelligence and machine learning based approaches [6]. In this paper, a back propagation based scaled conjugate gradient algorithm is used in conjugation with the discrete wavelet transform (DWT) for forecasting stock market trends. The evaluation of the proposed approach has been done based on the mean absolute percentage error (MAPE). A comparative MAPE analysis has also been done w.r.t. previously existing techniques [4].

II. DEEP LEARNING

Deep learning has evolved as one of the most effective machine learning techniques which has the capability to handle extremely large and complex datasets [5]. It is training neural networks which have multiple hidden layers as compared to the single hidden layer neural network architectures [6]-[7].
The architectural view of a deep neural network is shown in figure 1. In this case, the outputs of each individual hidden layer is fed as the input to the subsequent hidden layer. The weight adaptation however can follow the training rule decided for the neural architecture. There are various configurations of hidden layers which can be the feed forward, recurrent or back propagation etc [8].

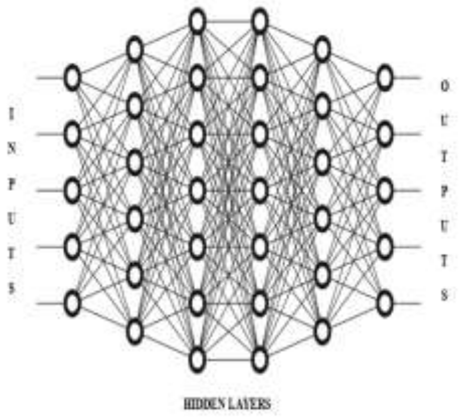

Fig.2 The Deep Neural Network Architecture

The figure above depicts the deep neural network architecture with multiple hidden layers. The output of the neural network however follows the following ANN rule:

$$\boldsymbol{Y = f(\sum_{i=1}^{n} X_i \,.\, W_i \;+\; \theta_i)} \qquad \textbf{(2)}$$

Where,
X are the inputs
Y is the output
W are the weights
ϴ is the bias.
Training of ANN is of major importance before it can be used to predict the outcome of the data inputs.

III. BACK PROPAGATION

Back propagation is one of the most effective ways to implement the deep neural networks with the following conditions [9]:

1) Time series behavior of the data
2) Multi-variate data sets
3) Highly uncorrelated nature of input vectors

The essence of the back propagation based approach is the fact that the errors of each iteration is fed as the input to the next iteration. [10]-[11].

The back propagation based approach can be illustrated graphically in figure 2.

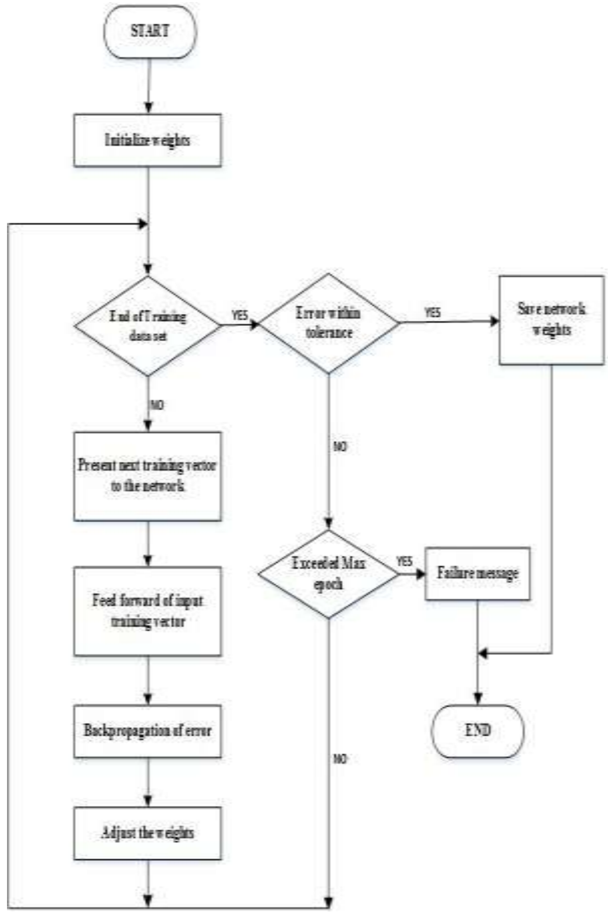

Fig.3 Concept of Back Propagation

The error feedback mechanism generally is well suited to time series problems in which the dependent variable is primarily a function of time along with associated variables. Mathematically,

$$\boldsymbol{Y = f(t, V_1 \ldots . V_n)} \qquad \textbf{(3)}$$

Here,
Y is the dependent variable
f stands for a function of
t is the time metric
V are the associated variables
n is the number of variables

In case of back propagation, the weights of a subsequent iteration doesn't only depend on the conditions of that iteration but also on the weights and errors of the previous iteration mathematically given by [12]:

$$\boldsymbol{W_{k+1} = f(W_k, e_k, V)} \qquad \textbf{(4)}$$

Here,
W_{k+1} are the weights of a subsequent iteration
W_k are the weights of the present iteration
e_k is the present iteration error
V is the set of associated variables

In general, back propagation is able to minimize errors faster than feed forward networks, however at the cost of computational complexity at times. However, the trade off between the computational complexity and the performance can be clearly justified for large, complex and uncorrelated datasets for cloud data sets [13].

IV. GRADIENT DESCENT BASED TRAINING

The gradient descent algorithms (GDAs) generally exhibit:

1) Relatively lesser memory requirement
2) Relatively faster convergence rate

The essence of this approach is the updating of the gradient vector g, in such as way that it reduces the errors with respect to weights in the fastest manner. Mathematically, let the gradient be represented by g and the descent search vector by p, then [14]:

$$p_0 = -g_0 \tag{5}$$

Where,

g_0 denotes the gradient given by $\frac{\partial e}{\partial w}$

The sub-script 0 represents the starting iteration

The negative sign indicates a reduction in the errors w.r.t. weights [15].

The tradeoff between the speed and accuracy is clearly given by the following relations [16]:

$$W_{k+1} = W_k - \alpha g_x\,, \quad \alpha = \frac{1}{\mu} \tag{6}$$

Here,

w_{k+1} is the weight of the next iteration

w_k is the weight of the present iteration

g_x is the gradient vector

μ is the step size for weight adjustment in each iteration.

There are several ways to implement the back propagation technique in the neural networks [17]. One consideration however always remains that of the least time and space complexity so as to reduce the amount of computational cost that is associated with the training algorithm. The essence of the scaled conjugate gradient algorithm is the fact that it has very low space and time complexity making it ideally suited to large data sets to be analyzed in real time applications where the time is a constraint. The training rule for the algorithm is given by [18]:

$$A_0 = -g_0 \tag{7}$$

A is the initial search vector for steepest gradient search

g is the actual gradient

$$w_{k+1} = w_k + \mu_k g_k \tag{8}$$

Here,

w_{k+1} is the weight of the next iteration

w_k is the weight of the present iteration

μ_k is the combination co-efficient

V. THE DISCRETE WAVELET TRANSFORM

The wavelet transform is an effective tool for removal of local disturbances. Stock prices show extremely random behavior and local disturbances. Hence conventional Fourier methods do not render good results for highly fluctuating data sets. Mathematically, the wavelet transform can be given as [19]

$$Z\,(S,P) = \int_{-\infty}^{\infty} z(\mathrm{t})\;((S,P,t))dt \tag{9}$$

Here,

S denotes the scaling operation

P denotes the shifting operation

t denotes the time variable

Z is the image in transform domain

z is the image in the spatial domain

The major advantage of the wavelet transform is the fact that it is capable of handling fluctuating natured data and also local disturbances. The DWT can be defined as [20]:

$$\mathrm{W\Phi}\ (\mathrm{Jo, k}) = \frac{1}{\sqrt{M}} \sum_n S(n).\Phi(n)_{jo\prime k} \tag{10}$$

The entire methodology can be understood using the system flowchart depicted below.

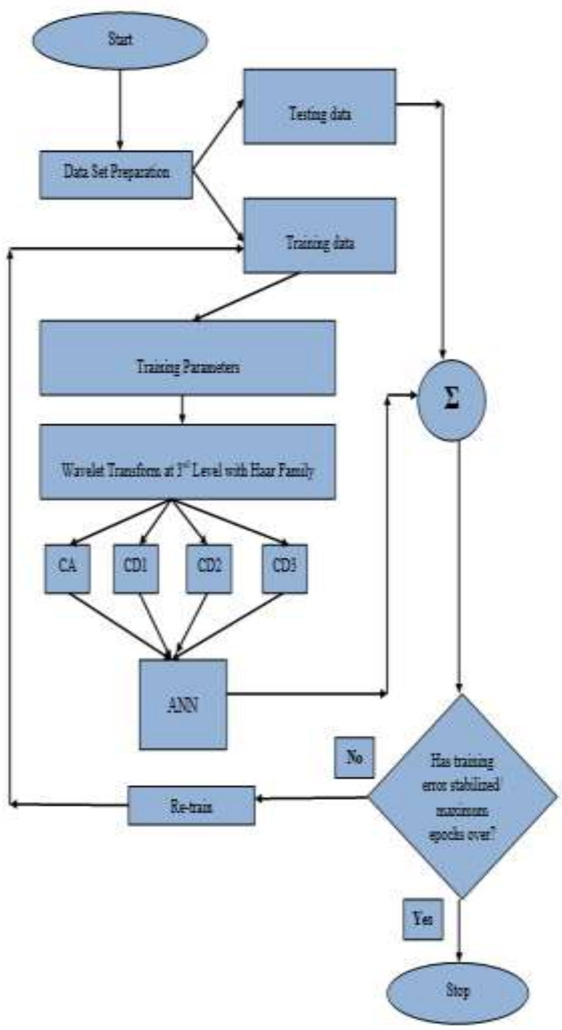

Fig.4 Flowchart of Proposed System

The data is divided in the ration of 70:30 for training and testing data set bifurcation.

The final performance metrics computed for system evaluation are:

1) Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{100}{M} \sum_{t=1}^{N} \frac{E - E_t|}{E_t} \qquad (11)$$

Here E_t and $E_t^{\sim}$ stand for the predicted and actual values respectively.

The number of predicted samples is indicated by M.

2) Regression

The extent of similarity between two variables is given by the regression where the maximum value is 1 and the minimum is 0.

VI. RESULTS

The results have been evaluated based on the following parameters:

1. (MAPE)
2. Regression
3. MSE w.r.t. the number of epochs

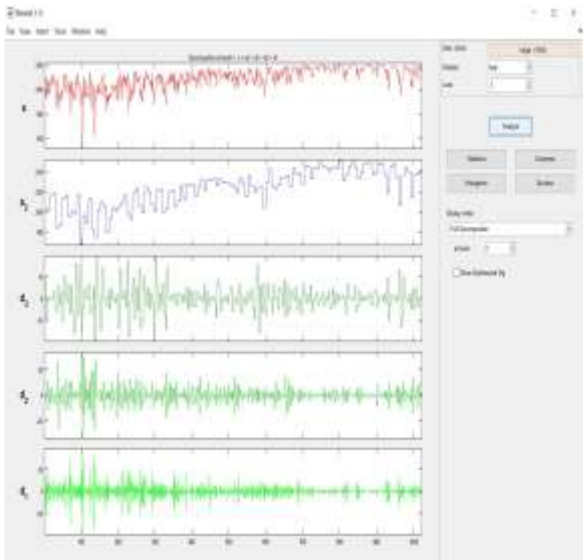

Fig.5 Decomposition of Data at level 3 using Haar Wavelets

The figure above depicts the Haar wavelet decomposition of the stock data at level 3.

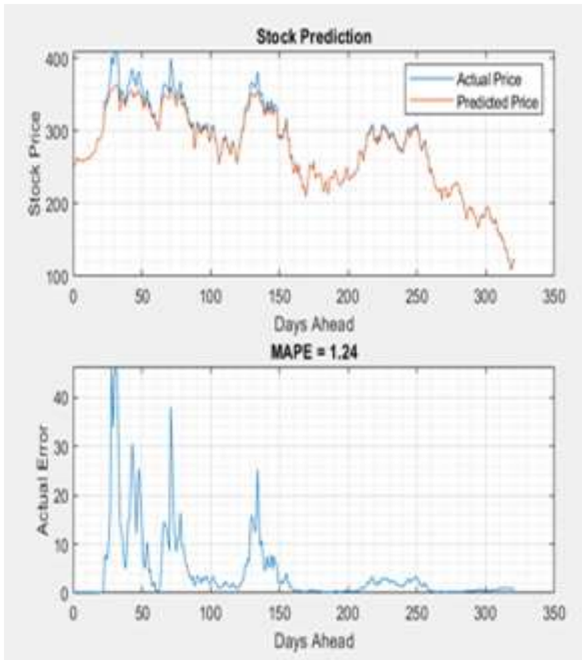

Fig.6 Predicted and Actual Stock Behavior

The figure above depicts the predicted and actual stock behavior.

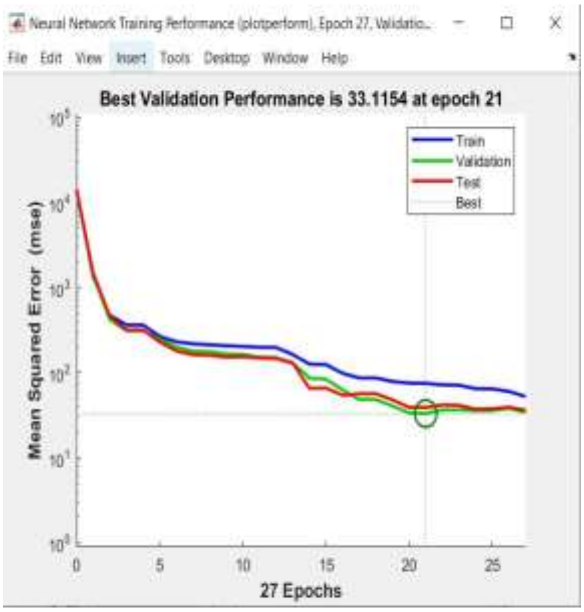

Fig.7 Variation of MSE with respect to epochs

The figure above depicts the variation of the mse w.r.t. the epochs

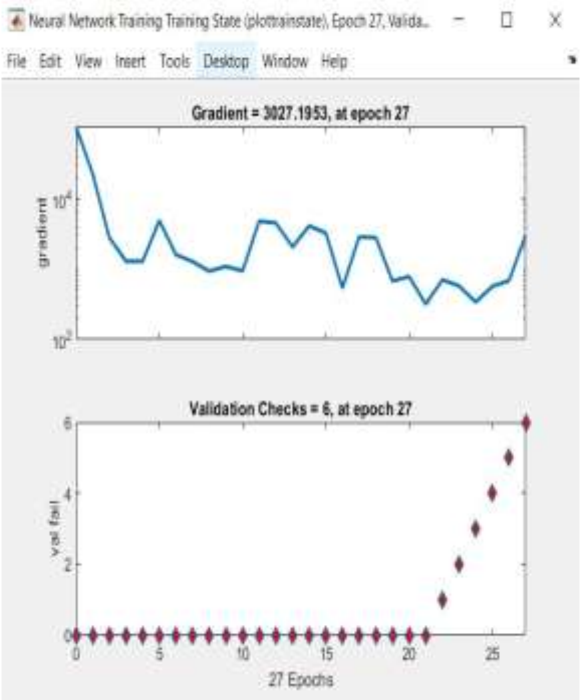

Fig.8 Training Parameters

The figure above depicts the training states as a function of iterations.

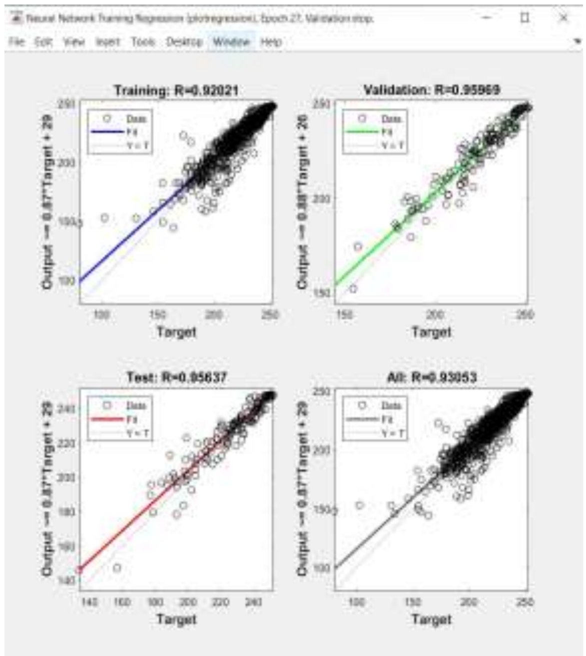

Fig.9 Regression Analysis

The figure above depicts the regression analysis of the proposed system

From the previous discussions, it can be observed that the proposed system obtaines an MAPE of 1.24% which is much higher than the previous work's forecasting accuracy for Tesla stocks.

CONCLUSION

It can be concluded from previous discussions that stock market prediction is a category of time series prediction with high sensitivity and dependence on external factors. Hence it is often challenging to attain high levels of accuracy in prediction. In the proposed approach a back propagation based deep learning model is proposed with a 7-10-10-10-10-10-1 configuration. The adaptive gradient descent algorithm (GDA) is used to train the neural network. Data pre-processing is done using the discrete wavelet transform. It has been shown that the proposed work attains a mean absolute percentage error of 1.24% compared to a mean absolute percentage error of previous work [1]. Thus the proposed system is able to achieve higher accuracy and relatively low number of iterations. Also the proposed work outperforms previously existing systems in terms of the accuracy for the benchmark datasets used.

REFERENCES

1. Subakkar, S Graceline Jasmine, L Jani Anbarasi, J Ganesh, CM Yuktha, "An Analysis on Tesla's Stock Price Forecasting Using ARIMA Model", Proceedings of the International Conference on Cognitive and Intelligent Computing, Springer, 2023, pp 83–89
2. S Kim, S Ku, W Chang, JW Song, "Predicting the Direction of US Stock Prices Using Effective Transfer Entropy and Machine Learning Techniques", IEEE Access 2022, Vol-8, pp: 111660 – 111682. DOI: 10.1109/ACCESS.2020.3002174
3. S Bouktif, A Fiaz, M Awad, Amir Mosavi, "Augmented Textual Features-Based Stock Market Prediction", IEEE Access 2021, Volume-8, PP: 40269 – 40282. DOI: 10.1109/ACCESS.2020.2976725
4. X Li, P Wu, W Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong", Information Processing & Management, Elsevier 2020. Volume 57, Issue 5, pp: 1-19. https://doi.org/10.1016/j.ipm.2020.102212
5. Gaurang Bansal; Vikas Hasija; Vinay Chamola; Neeraj Kumar; Mohsen Guizani, "Smart Stock Exchange Market: A Secure Predictive Decentralized Model", 2019 IEEE Global Communications Conference (GLOBECOM), IEEE 2019 pp. 1-6. DOI: 10.1109/GLOBECOM38437.2019.9013787
6. Jithin Eapen; Doina Bein; Abhishek Verma, "Novel Deep Learning Model with CNN and Bi-Directional LSTM for Improved Stock Market Index Prediction", 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), IEEE 2019 pp. 0264-0270. DOI: 10.1109/CCWC.2019.8666592
7. Min Wen; Ping Li; Lingfei Zhang; Yan Chen, "Stock Market Trend Prediction Using High-Order Information of Time Series", IEEE Access 2019, Volume 7, pp : 28299 – 28308. DOI: 10.1109/ACCESS.2019.2901842
8. Y Guo, S Han, C Shen, Y Li, X Yin, Y Bai, "An adaptive SVR for high-frequency stock price forecasting", Volume-6, IEEE Access 2018, pp: 11397 – 11404. DOI: 10.1109/ACCESS.2018.2806180
9. MS Raimundo, J Okamoto, "SVR-wavelet adaptive model for forecasting financial time series", 2018 International Conference on Information and Computer Technologies (ICICT), IEEE 2018, pp. 111-114. DOI: 10.1109/INFOCT.2018.8356851
10. Y Baek, HY Kim, "ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module" Journal of Expert System and Applications, Elsevier 2018, Volule-113, pp: 457-480. https://doi.org/10.1016/j.eswa.2018.07.019
11. S Selvin, R Vinayakumar, E. A Gopalakrishnan ; Vijay Krishna Menon; K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model", 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE 2017, pp. 1643-1647.
12. Z Zhao, R Rao, S Tu, J Shi, "Time-weighted LSTM model with redefined labeling for stock trend prediction", 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), pp. 1210-1217. DOI: 10.1109/ICTAI.2017.00184
13. DMQ Nelson, ACM Pereira, Renato A. de Oliveira , "Stock market's price movement prediction with LSTM neural networks", 2017 International Joint Conference on Neural Networks (IJCNN), IEEE 2017, pp. 1419-1426 DOI: 10.1109/IJCNN.2017.7966019
14. M Billah, S Waheed, A Hanifa, "Stock market prediction using an improved training algorithm of neural network", 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), IEEE 2016, pp. 1-4, doi: 10.1109/ICECTE.2016.7879611
15. Yanhui Guo, Siming Han, Chuanhe Shen, Ying Li, Xijie Yin, and Yu Bai, "An Adaptive SVR for High-Frequency Stock Price Forecasting", 2169-3536 2018 IEEE, VOLUME 6, 2018
16. Xi Zhang, Siyu Qu, Jieyun Huang, Binxing Fang, and Philip Yu, "Stock Market Prediction via Multi-Source Multiple Instance Learning", 2169-3536 2018 IEEE, VOLUME 6, 2018
17. Bruno Miranda Henrique, Vinicius Amorim Sobreiro, Herbert Kimura, "Stock price prediction using support vector regression on daily and up to the minute prices", The Journal of Finance and Data Science 4 (2018) 183-201
18. Wen Fenghua, Xiao Jihong, He Zhifang, Gong Xu, "Stock Price Prediction Based on SSA and SVM", Procedia Computer Science 31 (2014) 625 – 631.
19. HJ Sadaei, R Enayatifar, MH Lee, M Mahmud, "A hybrid model based on differential fuzzy logic relationships and imperialist competitive algorithm for stock market forecasting", Journal of Applied Soft Computing, Elsevier 2016, Volume 40, pp: 132-149.