Stock Market Prediction using RFR, DTR & SVR

Ravikant[1] Suman Kumar Swarnkar[2] L. P. Bhaiya[3]
[1,2,3]Bharti College of Engineering & Technology, Durg, Chhattisgarh, India

Abstract— Stock market or equity market have a profound impact in today's economy. An increase or fall within the share worth has a very important role in deciding the investor's gain. the prevailing foretelling strategies build use of each regression (AR, MA, ARIMA) and non-linear algorithms (ARCH, GARCH, Neural Networks), but they concentrate on predicting the indicator movement or worth foretelling for one company victimization the daily price. The planned methodology could be a model freelance approach. Here we have a tendency to don't seem to be fitting the info to a particular model, rather we have a tendency to be distinguishing the latent dynamics existing within the information victimization deep learning architectures. During this work we have a tendency to use 3 totally different deep learning architectures for the value prediction of firms and compares their performance. We have a tendency to be applying a window approach for predicting future values on a brief term basis. The performance of the models were quantified victimization proportion error.

Key words: Stock Market, RFR, DTR, SVR

I. Introduction

Forecasting are often outlined because the prediction of some future event or events by analyzing the historical information. It spans several areas as well as business and trade, economics, biology and finance. Prediction issues are often classified as

- Short term prediction (prediction for few seconds, minutes, days, weeks or months)
- Medium term prediction (prediction for one to a pair of years)
- long run prediction (prediction on the far side 2years)

Many of the prediction issues involve the analysis of your time. A statistic information are often outlined as a temporal arrangement of observations for a specific variable. In our case the variable is stock value. It will either be univariate or variable. Uni-variate information includes data regarding only 1 explicit stock whereas multivariate information includes stock costs of quite one company for varied instances of your time. Analysis of your time series information helps in distinguishing patterns, trends and periods or cycles existing within the information. Within the case of exchange, associate early data of the optimistic or pessimistic mode helps in investment cash showing wisdom. Conjointly the analysis of patterns helps in distinguishing the simplest playing firms for a specified amount. This makes statistical analysis and prediction a crucial space of analysis. The present strategies for stock value prediction are often classified as folows[1]

- elementary Analysis
- Technical Analysis
- statistic prediction

Fundamental analysis may be a form of investment analysis wherever the share price of an organization is calculable by analyzing its sales, earnings, profits and different economic factors. This methodology is most fitted to long run prediction. Technical analysis uses the historical value of stocks for distinguishing the longer-term value. Moving average may be a normally used algorithmic program for technical analysis. It are often thought of because the un-weighted mean of past n information points. This methodology is appropriate for brief term predictions. The third methodology is that the analysis of your time series information. It involves primarily 2 categories of algorithms, they are

- Linear Models
- Non Linear Models

The different linear models square measure AR, ARMA, ARIMA and its variations [2] [3] [4]. These models uses some predefined equations to suit a mathematical model to a uni-variate statistic. The most disadvantage of those models is that, they are doing not ac-count for the latent dynamics existing within the information. Since they contemplate solely univariate statistic, the bury dependencies among the assorted stocks aren't known by these models. Conjointly the model known for one series won't acceptable the opposite. Because of these reasons, it's insufferable to spot the patterns or dynamics gift within the information as an entire.

Non-linear models involve strategies like ARCH, GARCH, [3] TAR, Deep learning algo-rithms [5].In [6] associate analysis on the bury dependency between stock value and stock volume for twenty nine elite firms listed in great fifty has been done. The poposed work focuses on the applying of deep learning algorithms for stock value prediction [7] [8]. Deep neural networks are often thought of as non-linear operate approximators that square measure capable of mapping non-linear functions. Supported the sort of applica-tion, varied styles of deep neural network architectures square measure used. These embrace multi-layer perceptrons (MLP),Recursive Neural Networks (RNN),Long Short Term Memory(LSTM), CNN(Convolutional Neural Network) etc [9].They have been ap-plied in varied areas like image process, tongue process, statistical analysis etc.

Deep learning algorithms square measure capable of distinguishing hidden patterns and underly-ing dynamics within the information through a self-learning method.

In the case of exchange, the information generated is big and is very non-linear. To model such quite energizing information we'd like models that may analyze the hidden patterns and underlying dynamics. Deep learning algorithms square measure capable of distinguishing and exploiting the interactions and patterns existing in a very information through a self-learning method. Not like different algorithms, deep learning models will effectively model these form of information and may provide a smart prediction by analyzing the interactions and hidden patterns at intervals the information. In [5], we will see the applying of various deep learning models for variable statistical analysis. The primary commit to model a monetary statistic employing a neural network model was introduced in [10]. This work created an effort to model a neural network model for cryptography the non-linear regularities in plus value movements for IBM. but the scope of the work was restricted, however it helped in establishing evidences against EMH [11].

Researches within the space of economic statistical analysis victimization NN models used different input variables for predicting the stock come. In some works, information from one statistic were used as input [10], [8]. Sure works thought of the inclusion of heterogeneous market data and economic science variables. In [12], a com-bination of economic statistical analysis and IP are introduced. In [13] and [7], deep learning architectures are used for the modeling of variable financial statistic. In [14], a NN model victimization technical analysis variables are enforced for the prediction of Shanghai exchange. The work compared the performance of 2 learning algorithms and 2 weight format strategies. The results sh6wn that potency of back propagation are often inflated by conjugate gradient learning with multiple regression toward the mean weight format.

In 1996, [15] used back propagation and RNN models for the prediction of index for 5 totally different stock markets. In [16], application of your time delay, continual and probabilistic neural network models were introduced for daily stock prediction. In [17], application of machine learning algorithms like PSO and LS-SVM are used for the prediction of S&P five hundred exchange. Implementation of genetic algorithms in conjunction with neural network models were introduced in [18]. The work combined the applying of each genetic algorithmic program and artificial neural network for prediction. during this work weights for NN was obtained from the genetic algorithmic program. However, the prediction accuracy for this model was low. Application of rippling rework for prediction was introduced in [19]. The work used riffle remodel to explain short term options available trends. With the introduction of LSTM [20], the analysis of your time dependent information become additional economical. These style of networks have the aptitude of holding past information. They need been employed in stock value prediction by [8], [7].

The projected methodology focuses on predicting stock value for NSE (National Stock Exchange) listed firms. The approach we've got adopted may be a window approach with data overlap. The information set contains minute wise data of NSE listed firms. Here we have a tendency to try to get a generalized model for the aim of prediction which might use minute wise information as input. This sort of modeling have applications in algorithmic commerce wherever high frequency commerce happens.

The paper is structured as follows section [II] explains the projected methodology. Results and discussions is found in Section [III] and Section [IV] includes the conclusion.

II. METHODOLOGY

The data set consists of minute wise stock worth for 1721 NSE listed firms for the amount of Gregorian calendar month 2014 to June 2015. It includes informations like day stamp, time stamp, dealings id, stock worth and volume of stock sold-out in every minute. For this work we've got hand-picked 2 completely different sectors, IT sector and company sector. 2 companies from IT sector and one company from company sector were taken for the study. These firms were known by the assistance of NIFTYIT index and NIFTY-Pharma index. Information for these 3 firms were extracted from the on the market data and was subjected to preprocessing to get the stock worth.

The work relies on a window approach for a brief term future prediction. The window size was mounted to be a hundred minutes with AN overlap of ninety minute's information and prediction was created for ten minutes in future. the most effective window length was known by shrewd the error for numerous window sizes. The train information consists of stock worth of Infosys for the amount July-01-2014 to October-14-2014 and take a look at information consists of stock worth for Infosys, TCS and CIPLA for the amount of Octo-ber-16-2014 to November-28-2014.

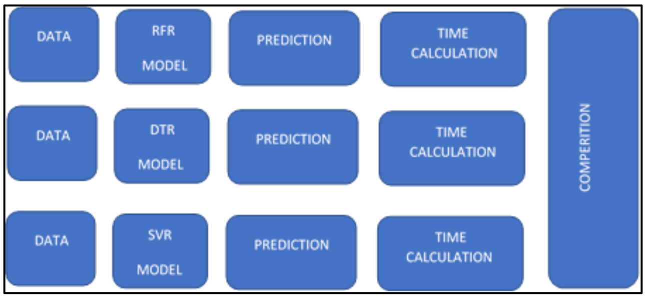

Fig. 1: Proposed Model Block Diagram

The data varies with in a very vary of 2000 to four hundred0 for Infosys and TCS and for Cipla it's 400 to 700. To unify the info vary, it had been subjected to standardization and was mapped to a variety of zero to one. This normalized information was given to the network for training. All the models were trained for one thousand epochs by variable the layer size for fine standardization. If the loss (mean square error) for this epoch is a smaller amount than the worth obtained in previous epoch, the load matrices for that epoch is keep. once the coaching method every of those models were tested and also the model with least RMSE (Root Mean square Error) is taken because the final model for prediction.

We have used 3 completely different deep learning architectures, RNN, LSTM and CNN for this work. RNN could be a category of neural network wherever connections between the computational units kind a directed circle. in contrast to feed forward networks, RNN will use their internal memory to method impulsive sequence of inputs. every of the computing unit in AN RNN incorporates a time variable real valued activation and modifiable weight. RNNs area unit created by applying identical set of weights recursively over a graph-like structure. Several of the RNNs use (1) to outline the values of their hidden units

$$h^t = f(h^{t-1}, x^t ; \theta) \quad (1)$$

In the case of RNN, the learned model perpetually has an equivalent input size, as a result of it's laid out in terms of transition from one state to a different. Conjointly the design uses an equivalent transition perform with an equivalent parameters at when step. LSTM could be a special quite RNN, introduced in 1997 by Hochreiter and Schmidhuber [20]. Within the case of LSTM design, the same old hidden layers area unit replaced with LSTM cells. The cells area unit composed of assorted gates which will management the input flow. AN LSTM cell consists of input gate, cell state, forget gate, and output gate. It conjointly consists of sigmoid layer, tanh layer and purpose wise multiplication operation.

The cell state is updated supported the outputs kind the gates. Mathematically we are able to represent it mistreatment the subsequent equations.

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \quad (2)$$

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \quad (3)$$

$$c_t = \tanh(W_c.[h_{t-1}, x_t] + b_c) \quad (4)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t * \tanh(c_t) \quad (6)$$

Where xt: input vector, ht: output vector, ct: cell state vector, ft: forget gate vector, it: input gate vector, ot: output gate vector and W, b area unit the parameter matrix and vector. Convolutional neural networks or CNNs, area unit a specialised quite neural net-work for process information that contains a acknowledged, grid-like topology. This embody time-series information, which might be thought of as a 1D and image information, which might be thought of as a 2nd grid of pixels. The network employs a computation known as convolution and thus referred to as convolutional neural network. it's a specialised quite linear operation. Convolutional networks use convolution rather than general matrix multi-plication in a minimum of one in every of their layers. The motivation behind mistreatment these 3 models is to spot whether or not there's any long run dependency existing within the given information. this may be known from the performance of the models. RNN and LSTM architectures area unit capable of distinctive long run dependencies and uses them for future prediction. but CNN architectures in the main focuses on the given input sequence and doesn't use any previous history or data throughout the educational method. The motivation behind testing the models with information from alternative firms is to visualize for interdependencies among the businesses and to grasp the market dynamics.

The train information was normalized. check information was conjointly subjected to an equivalent normalization. once getting the expected output, denormalization was applied and per-centage error was calculated mistreatment the accessible true labels. The error share was calculated mistreatment (7)

$$ep = \frac{abs\left[X^i_{real} - X^i_{predicted}\right]}{X^i_{real}} \times 100 \quad (1)$$

Where ep is that the error share, X^i_{real} is that the i^{th} real price and X^i foreseen is that the i^{th} foreseen price. Error share offers the magnitude of error gift within the output.

III. RESULTS & DISCUSSION

The experiment was in serious trouble 3 totally different regression models. The utmost price of error proportion obtained for every model is given in Table [I]. From the table it's clear that CNN is giving a lot of correct results than the opposite 2 models. This is often because of the rationale that CNN doesn't rely upon any previous data for prediction. It uses solely this window for prediction. This allows the model to under-stand the high-powered changes and patterns occurring within the current window. But within the case of RNN and LSTM, it uses data from previous lags to predict the long run instances. Since exchange may be an extremely phase space, the patterns and dynamics existing with within the system won't forever be an equivalent. This cause teaching issues to LSTM and RNN design and thence the models fails to capture the high-powered changes accurately.

For comparison we've got used ARIMA, that may be a linear model used for forecasting. The error proportion obtained for the 3 firms ar as follows

	RFR	DTR	SVR
TIME	0.021	0.0009	0.007
PREDICTED VALUE	20.31	20.35	23.55

Table 1: Time & Predicted Value of RFR, DTR & SVR

	RBF	LINEAR	POLYNOMIAL
TIME	0.007	0.018	59.45
PREDICTED VALUE	23.55	15.17	11.54

Table 2: Time & Predicted Value of rbf, Linear & Polynomial

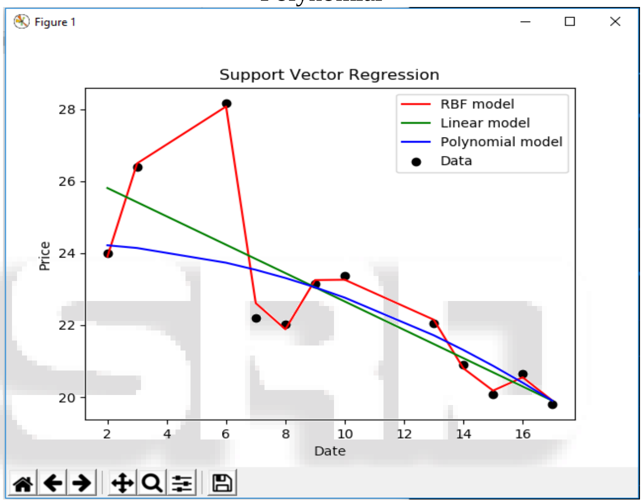

Fig. 2: Support Vector Regression

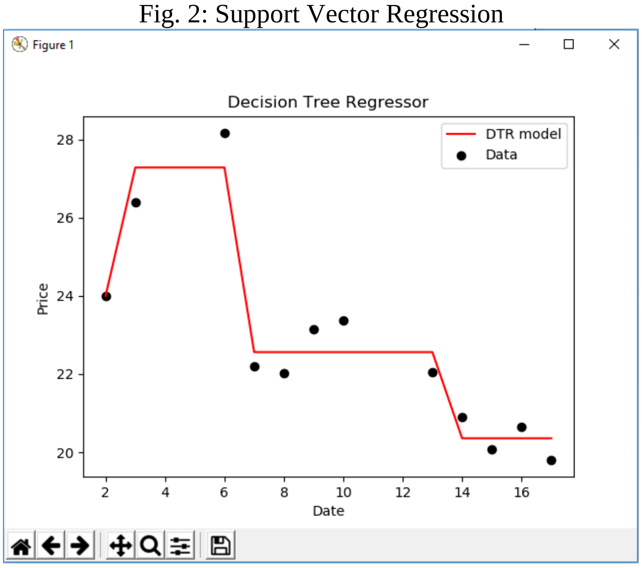

Fig. 3: Decision Tree Regressor

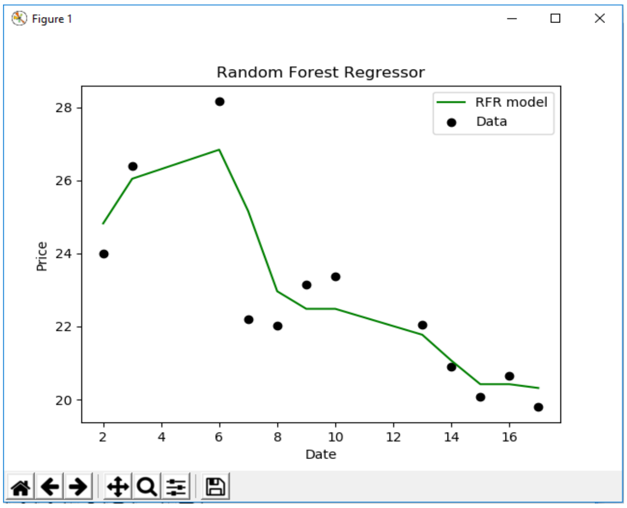

Fig. 4: Random Forest Regressor

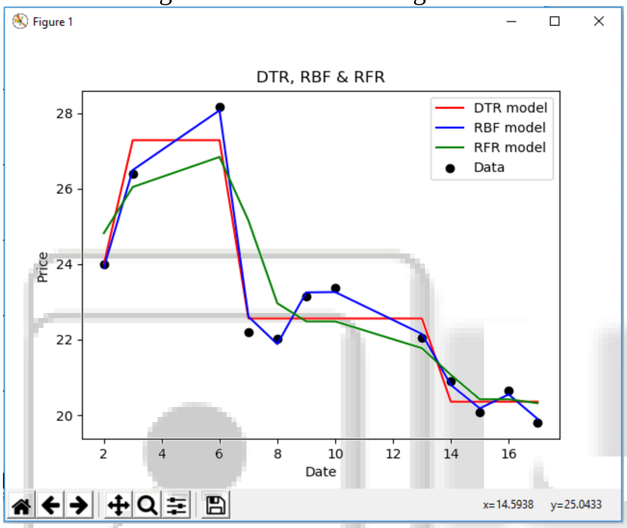

Fig. 5: DTR,RBF & RFR

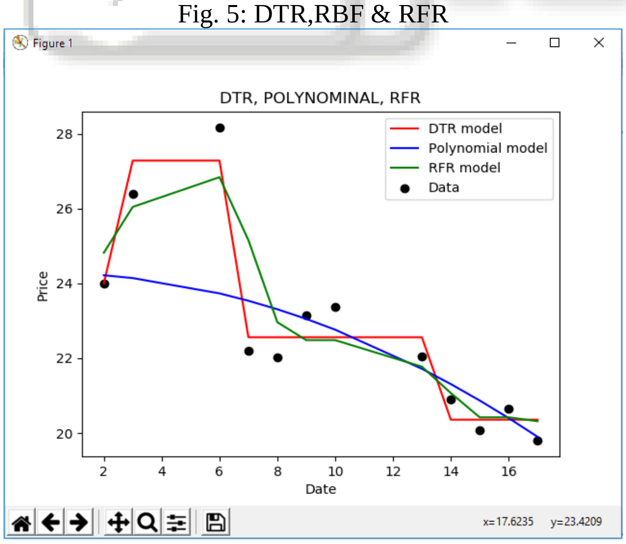

Fig. 6: DTR, POLYNOMINAL, RFR

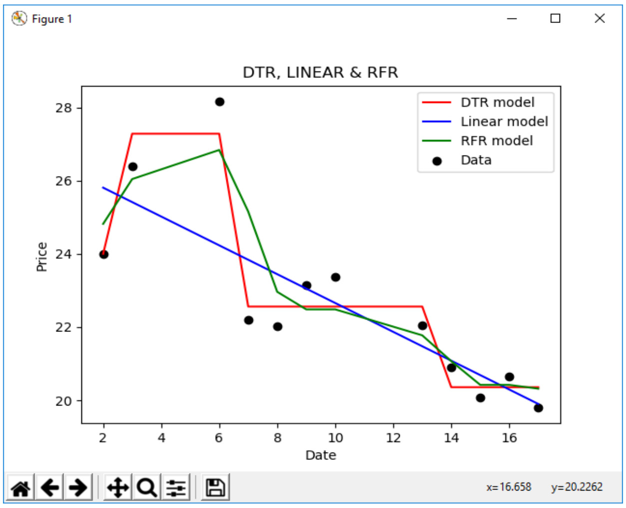

Fig. 7: DTR, LINEAR & RFR

From Fig(2) and Fig(4) it will be discovered that each SVR and RFR fails to cap-ture the trends and dynamics towards the top (between the period 9000 to 11000) ie; there's a amendment within the behavior of the stock pattern for that point window in comparison to the previous windows. just in case of DTR it's evident from the Fig(3), that is capable of capturing the changes within the trend for the period 9000 to 11000

In the case of SVR Fig(2) LINEAR and POLYNOMIAL don't seem to be characteristic the pattern within the starting of the window. there's a amendment within the trend followed by RBF throughout that amount. This makes the predictions more correct.

IV. Conclusion

We propose a regression based formalization for stock price prediction. It is seen that, regression architectures are capable of capturing dynamics and are able to make predictions. We trained the model using the data of stock and was able to predict stock price of stock. This shows that, the proposed system is capable of identifying some inter relation with in the data. Also, it is evident from the results that, SVR, RFR & DTR model is capable of identifying the changes in trends. For the proposed methodology DTR is identified as the best model. It uses the information given at a particular instant for prediction. Even though the other two models are used in many other time dependent data analysis, it is not out performing the DTR model in this case. This is due to the sudden changes that occurs in stock markets. The changes occuring in the stock market may not always be in a regular pattern or may not always follow the same cycle. Based on the companies and the sectors, the existence of the trends and the period of their existence will differ. The analysis of these type of trends and cycles will give more profit for the investors.To analyze such information we must use networks like DTR as they rely on the current information.

References

[1] Devadoss and T. A. A. Ligori, “Forecasting of stock prices using multi-layer perceptron,” Int J Comput Algorithm, vol. 2, pp. 440–449, 2013.

[2] J. G. De Gooijer and R. J. Hyndman, “25 years of time series forecasting,” International journal of forecasting, vol. 22, no. 3, pp. 443–473, 2006.
[3] K. Menon, N. C. Vasireddy, S. A. Jami, V. T. N. Pedamallu, V. Sureshkumar, and K. Soman, “Bulk price forecasting using spark over nse data set,” in International Conference on Data Mining and Big Data. Springer, 2016, pp. 137–146.
[4] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time series analysis: forecasting and control. John Wiley & Sons, 2015.
[5] G. Batres-Estrada, “Deep learning for multivariate financial time series,” ser. Technical Report, Stockholm, May 2015.
[6] P. Abinaya, V. S. Kumar, P. Balasubramanian, and V. K. Menon, “Measuring stock price and trading volume causality among nifty50 stocks: The toda yamamoto method,” in Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on. IEEE, 2016, pp. 1886–1890.
[7] J. Heaton, N. Polson, and J. Witte, “Deep learning in finance,” arXiv preprint arXiv: 1602.06561, 2016.
[8] H. Jia, “Investigation into the effectiveness of long short term memory networks for stock price prediction,” arXiv preprint arXiv: 1603.07893, 2016.
[9] Y. Bengio, I. J. Goodfellow, and A. Courville, “Deep learning,” Nature, vol. 521, pp. 436–444, 2015.
[10] H. White, Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns, ser. Discussion paper- Department of Economics University of California San Diego. Department of Economics, University of California, 1988.
[11] B. G. Malkiel, “Efficient market hypothesis,” The New Palgrave: Finance. Norton, New York, pp. 127–134, 1989.
[12] X. Ding, Y. Zhang, T. Liu, and J. Duan, “Deep learning for event-driven stock prediction.” in IJCAI, 2015, pp. 2327–2333.
[13] J. Roman and A. Jameel, “Backpropagation and recurrent neural networks in financial analysis of multiple stock market returns,” in System Sciences, 1996. Proceedings of the Twenty-Ninth Hawaii International Conference on,, vol. 2. IEEE, 1996, pp. 454–460.
[14] M.-C. Chan, C.-C. Wong, and C.-C. Lam, “Financial time series forecasting by neural network using conjugate gradient learning algorithm and multiple linear regression weight initialization,” in Computing in Economics and Finance, vol. 61, 2000.
[15] J. Roman and A. Jameel, “Backpropagation and recurrent neural networks in financial analysis of multiple stock market returns,” in System Sciences, 1996. Proceedings of the Twenty-Ninth Hawaii International Conference on, vol. 2. IEEE, 1996, pp. 454–460.
[16] E. W. Saad, D. V. Prokhorov, and D. C. Wunsch, “Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks,” IEEE Transactions on neural networks, vol. 9, no. 6, pp. 1456–1470, 1998.
[17] O. Hegazy, O. S. Soliman, and M. A. Salam, “A machine learning model for stock market prediction,” arXiv preprint arXiv: 1402.7351, 2014.
[18] K.-j. Kim and I. Han, “Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index,” Expert systems with Applications, vol. 19, no. 2, pp. 125–132, 2000.
[19] Y. Kishikawa and S. Tokinaga, “Prediction of stock trends by using the wavelet transform and the multi-stage fuzzy inference system optimized by the ga,” IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol. 83, no. 2, pp. 357–366, 2000.
[20] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.

