Stock prediction by applying hybrid Clustering-GWO-NARX neural network technique

Debashish Das[a,b], Ali Safa Sadiq[a], Seyedali Mirjalili[c]

[a]Faculty of Computer Systems & Software Engineering, University Malaysia Pahang
Kuantan, Pahang Darul Makmur, Malaysia.

[b]Faculty of Computing, Engineering and Technology, Asia Pacific University of Technology and Innovation
TPM, Bukit Jalil, Kuala Lumpur 57000, Malaysia

[c]School of Information and Communication Technology, Griffith University, Nathan Campus, Brisbane QLD 4111, Australia

debashish.das@apu.edu.my, alisafasadiq@ump.edu.my, ali.mirjalili@gmail.com

Abstract: Prediction of stock price is one of the most challenging tasks due to nonlinear nature of the stock data. Though numerous attempts have been made to predict the stock price by applying various techniques, yet the predicted price is not always accurate and even the error rate is high to some extent. Consequently, this paper endeavours to determine an efficient stock prediction strategy by implementing a combinatorial method of Grey Wolf Optimizer (GWO), Clustering and Non Linear Autoregressive Exogenous (NARX) Technique. The study uses stock data from prominent stock market i.e. New York Stock Exchange (NYSE), NASDAQ and emerging stock market i.e. Malaysian Stock Market (Bursa Malaysia), Dhaka Stock Exchange (DSE). It applies K-means clustering algorithm to determine the most promising cluster, then MGWO is used to determine the classification rate and finally the stock price is predicted by applying NARX neural network algorithm. The prediction performance gained through experimentation is compared and assessed to guide the investors in making investment decision. The result through this technique is indeed promising as it has shown almost precise prediction and improved error rate.

Keywords: Stock prediction, Grey Wolf Optimizer, Non Linear Autoregressive Exogenous algorithm, K-means clustering, Error Rate.

1. Introduction

For few decades, stock market is one of the best investment options for investors despite stock price being uncertain and unpredictable. Due to the nonlinear nature of stock, the prediction of stock price is even more complicated task. Moreover, the stock price of a company does not depend on the economic condition of a country. Researchers are continuously striving to find the best prediction technique for stock market as the economy of a country heavily depends on stock market.

The advancement of Information and Communication technology has made the processing of stock data much faster and the price of stock changes quite frequently. Due to this, the stock brokers, big investors like financial institutions, banks and any other investor need to make the stock investment decision within shortest possible time. However, the changes of stock prices are not depend on only one factor rather it is controlled by various factors like positive or negative news related to the company, political turbulences, rate of interest, dollar price, natural disasters and many more [1].

The investment decision to the stock market is a crucial task. The investor needs to be well educated and trained about the investment decision as the price of stock may be increased or decreased next day. Nevertheless, neural network has proved its potential as it can make better prediction than Fundamental and Technical analysis [2]. This paper is an attempt to guide the investors about the stock investment decision.

Data mining is a very important research dimension which is a way of discovering useful patterns from huge collection of data [3]. Large amount of historical data can be managed well through data mining technique as it is a way of converting data into more manageable form [4]. Clustering is a data mining method which can classify the unclassified data into smaller groups that has common properties [5]. Hence, the clustering data mining technique can be used for classifying the stock data.

Neural network is another very important method that consists of a lot of interconnected processors and it can have the capacity to deal with non-linear data like stock data which changes very fast [6]. Neural network is found to be very effective by several researches compared with traditional methods. It can be used for prediction of many different types of data including financial field i.e. prediction of stock and performance [7,8]. Classification is a data analysis method of neural network by which it can predict categorical class labels. Grey Wolf Optimizer (GWO) is a classification algorithm that can be optimized for solving a challenging problem with unknown search spaces [9]. Grey Wolf Optimizer is a process through which the algorithm is optimized for dealing with stock data. A nonlinear autoregressive exogenous model (NARX) is a nonlinear autoregressive time series modeling which can predict the recent value of a time series based on past values of same series.

In this research, we have collected data from renowned stock market like New York Stock Exchange (NYSE), NASDAQ and rising stock market like Dhaka Stock Exchange (DSE). We have employed hybridized technique on the stock data

called Clustering-GWO-NARX, where the Clustering is a process by which the most promising stock is categorized and in this research work we have employed k-means clustering, next the classification algorithm called Grey Wolf Optimizer (GWO) is applied to determine the classification rate, then Nonlinear Autoregressive Exogenous (NARX) algorithm is applied to predict the stock price.

The remaining part of the paper is organized as follows. Section 2 presents brief overview of related research. Section 3 describes the basic concepts of GWO. Section 4 illustrates the architecture of the proposed prediction system, while Section 5 explains the experimental results obtained. The paper is concluded in section 6.

2. Related Research Work

Numerous researches have been directed towards stock prediction based on not only fundamental or technical analysis but also through data mining and neural network based applications. However, each technique has its own pros and cons. Moreover, none of the techniques has proved to be capable in making 100% correct prediction. Due to this reason, the research is on-going for developing an accurate stock prediction system. As data mining and neural network based technique has proved to be one of the most accurate stock prediction technique, we are going to review previous studies on data mining and neural network strategies applied for stock prediction in the field of our research study.

2.1 General stock prediction systems and their limitation

Prediction of stock price is very popular in financial and academic arena. Fundamental method is widely used in analysing the stock market through Time series analysis. Numerous existing forecasting systems are available based on time series data and factors for stock prediction. Xu has attempted to predict weekly stock price by mingling the conventional time series analysis technique with information from the Google trend website and the Yahoo finance website. He had found that there is a correlation between the changes in weekly stock prices and the values of important news/events computed from the Google trend website. He also argued that his result is better than traditional time series analysis however he has also agreed that his study has many drawbacks as well. He has finally suggested that further research can be conducted to improve the stock prediction and more accurate algorithm can be used for such prediction [10]. Statistical analysis is another way of making stock prediction by applying regression, co-linearity, ARIMA and so on. But, main drawback of statistical model is that it can deal with linear data whereas stock data is non-linear in nature. Due to this reason, the statistical model is becoming obsolete.

2.2 Stock prediction using data mining: Issues and Challenges

Extracting knowledge from database is known as Data mining. Data mining is able to extract the appropriate data based on various techniques i.e. Classification, Prediction, Clustering, Association. Classification and Prediction is predictive model, whereas Clustering and Association is descriptive model. Unknown pattern of data can be discovered through data mining [11]. Kannan et al. has emphasized that the amount of data is growing day by day and hence we need to apply an appropriate technique to better manage the huge amount of data. Therefore, data mining is identified as very demanding and one of the best key research areas in data management. So, they have applied data mining technology and performed data analysis to predict whether the closing price of a stock will increase or decrease for a day better than level of significance. However, their algorithm was able to make 50% correct prediction and their algorithm was silent for few situations i.e. when to sell the stock [4]. Hence, better data mining algorithm is needed to be explored to make more accurate prediction.

2.3 Stock prediction systems using neural network and their limitation

Neural network is established based on artificial intelligence and it consists of interconnected processors like human brain. It uses mathematical model to make prediction. The most exciting feature of neural network is its ability to deal with non-linear or dynamic data. Due to this, it is extensively applied in dealing with financial data i.e. forecasting the prices of stock market. Neural network can make good prediction if it is configured well for a particular problem as it can learn based on patterns. However, the main challenge of neural network is optimizing number of neurons for a problem and selection of data [11]. Tomasz Kozdraj has attempted to apply neural network in predicting stock price for Warsaw Stock Exchange. He has considered stock prices of 30 companies and performed necessary research and analysis which provided him average level of mean absolute percentage error. The error value was acceptable for his study however he encountered some disturbances in forecasting fitness. He added that though the neural network is worthwhile in stock prediction, but there is no existence of universal forecasting method. The problem which lowered his result is due to various reasons such as one single outcome of a network to a specific internal decision is very difficult, noisy data, overfitting which affects generalization, choice of input series, network parameter setting etc. He concluded that complex network suffers from various difficulties such as data pre-processing, selection of architecture and parameterization. Finally, he resolved that neural network is having limitation and problems in stock prediction which needs further research [12].

2.4 Use of classification in stock prediction systems and their limitation

Analysing the data to determine the possible course of action is known as classification, for example: stock data is analysed to know whether it is feasible to invest or not. The dynamic and non-linear nature of stock data demands the application of sophisticated algorithm. However, the technological advance has made us available the computational intelligence by providing various tools and algorithms. Classification and various machine learning algorithms which use combination of classifiers are also

gaining researchers' attention in dealing with the stock prediction. Among the various classification methods, some of the researchers use single classification for stock forecasting, whereas others use combination of multiple classifiers to improve the accuracy of prediction [13]. Teixeira et al. have proposed of applying combination of technical analysis and nearest neighbour classification in predicting Brazilian stock market. They have compared their result with buy-and-hold strategy and found the combination of technical analysis and classification produced more profits. Eventually, their proposed method also reduced the risk. However, they have suggested using of better classification algorithm such as Support Vector machines, k-NN etc. to achieve more generalization. Additionally, they have also recommended applying classification technique in Asian market [14]. B. Narayanan et al. have attempted applying two classification methods AdaSVM and AdaNaive and compared their result with widespread classification methods i.e. Support Vector Machine (SVM) and Naïve Bayes which gave them more accuracy with significant reduce of classification error. They have also mentioned that change of technique and increase of input data may provide even more accurate result. Finally, they have agreed that continuous research is essential in developing an acceptable stock forecasting method [15].

2.5 Stock prediction systems using hybrid data mining and neural network and their limitation

Data mining is capable of describing cause and effect relation of information, whereas, much prediction problems can be solved using neural network. However, distinct application of neither data mining nor neural network is as effective as the combination of these two techniques [15]. Prior researches also support that hybridization of these two techniques provide more accurate result with less error rate. Tsai et al. have applied hybridization of artificial neural network (ANN) and decision tree (DT) data mining technique in creating stock price forecasting model for Taiwan stock market which provided them 77% accuracy which was higher than the single use of ANN or DT. However, they recommended performing further research by applying other techniques such as support vector machine, genetic algorithm etc. They have also suggested pre-processing data applying genetic algorithm to achieve better prediction accuracy. Finally, they have added that other industries can be merged with electronic for comparisons [16]. The database of shares was studied and prediction was made accordingly applying data mining and artificial intelligence by Navale et al. They have gained highest probability prediction for share in their study. However, they have emphasized the necessity of improving parameters in accurate prediction performance through data mining and artificial intelligence. They have also recommended for further research on hybridizing artificial intelligence and data mining to reveal the weaknesses of other researches [17].

The advantages and shortcomings of the relevant researches above states that further research is required for revealing the uncertainty of stock market. In search of appropriate stock prediction method, most effective methods like data mining, neural network, classification etc. can be applied for developing accurate prediction system.

3. Basic concepts of Grey Wolf Optimizer (GWO)

An algorithm which is executed iteratively to produce optimum solution through the comparison of various solutions is known as Optimization algorithm. Grey Wolf Optimizer (GWO) is a new meta-heuristic algorithm which is emulated from grey wolves (Canis lupus) and it mimics the leadership hierarchy and hunting mechanism of grey wolves in nature [9]. The leadership hierarchy is simulated applying four types of grey wolves i.e. α, β, δ, ω and hunting is implemented applying three main steps i.e. searching for prey, encircling prey, attacking prey. The GWO algorithm is applicable to challenging problems in unknown search spaces which produces better result than PSO, GSA, DE, EP and ES [9]. In this research, GWO algorithm is analyzed and it is applied as a training algorithm for Multi-layer perceptron such that it suits with stock data. The novelty of the research is that GWO is applied first time for stock prediction. The proposed algorithm improves the wolves attack strategy. It calculates the weights based on wolves fitness function and gives the highest weight to the dominant wolf concurrently to improve the convergence. The empirical results confirm the power of GWO and demonstrate that the algorithm can faster decide the suitable thresholds, provide good classification rate, efficiency and accuracy.

Grey wolves (Canis lupus) belong to Canidae family and they are considered as apex predators, meaning that they are at the top of the food chain. They have a very strict social dominant hierarchy. The algorithm divides the wolves into four types: α, β, δ, and ω, whereas, each type of wolves displays the following social behavior.

The leaders are a male and a female, called alpha. Decision making about hunting, sleeping place, time to wake etc. is made by alpha, that's the reason alpha becomes the leader in the pack and others follow its orders. The best solution of a problem can be determined identifying the location of alpha, as it is the best member in managing the pack.

The second level in the hierarchy of grey wolves is beta. The betas are subordinate wolves that help the alpha in decision-making or other pack activities. The beta maintains the discipline of the pack that enforces the alpha's commands in the pack and hence, it takes the role of the advisor to the alpha.

The third level in the hierarchy of grey wolves is delta and they need to submit to alphas and betas, but they dominate the omega. They are responsible for watching the boundaries of the territory, warning the pack in case of any danger, protecting and guaranteeing the safety of the pack, helping the alphas and betas when hunting prey, and providing food for the pack and caring for the weak, ill, and wounded wolves in the pack.

The last grey wolf in the hierarchy is omega and it may not be an important individual in the pack, but it has been observed that the whole pack face internal fighting and problems in case of losing the omega, which is harmful to the group structure.

Group hunting is another interesting social behavior of grey wolves in addition to the social hierarchy. The main phases of grey wolf hunting are as follows: searching for the prey; tracking, chasing, and approaching the prey; pursuing, encircling, and harassing the prey until it stops moving; attacking toward the prey.

The fittest solution is considered as the alpha (α) in designing mathematical model of the social hierarchy of wolves for designing GWO [9]. Beta (β) and delta (δ) are respectively the second and third fittest solution. Omega (ω) is considered for the rest of the candidate solutions. In the GWO algorithm, the hunting (optimization) is guided by α, β, and δ. The ω wolves follow these three wolves.

3.1 Preliminary of GWO Mathematical Model

The mathematical model of MGWO can be formed as per the equations below:

$$\vec{P} = \left| \vec{N} . \vec{Z}_a(i) - \vec{Z}(i) \right| \tag{1}$$

$$\vec{Z}(i + 1) = \vec{Z}_a(i) - \vec{M}.\vec{P} \tag{2}$$

Where, i denotes the current iteration, $\vec{M}$ and $\vec{N}$ are coefficient vectors, $\vec{Z}_a$ is the position vector of the prey, and $\vec{Z}$ denotes the position vector of a grey wolf.

The vectors $\vec{M}$ and $\vec{N}$ are calculated as follows:

$$\vec{M} = 2\vec{q}.\vec{s}_1 - \vec{q} \tag{3}$$

$$\vec{N} = 2. \vec{s}_2 \tag{4}$$

Where, components of $\vec{q}$ are linearly reduced from 2 to 0 over the number of iterations and and is used for controlling the trade-off between exploitation and exploration. The following equations will be employed for updating the value of variable:

$$\vec{q} = 2 - i\,(2\,/\,X_i) \tag{5}$$

$$\vec{Z}(i + 1) = (\,\vec{Z}_1 + \vec{Z}_2 + \vec{Z}_3\,)\,/\,3 \tag{6}$$

Where, X_i denotes the number of iterations, $\vec{s}_1$ and $\vec{s}_2$ are random vectors between [0, 1] which are employed to find the optimal solution. Appropriate idea about the potential location of prey can be availed by Alpha, Beta and Delta, where they help the Omega to follow the suitable positions. The values of $\vec{Z}_1$, $\vec{Z}_2$ and $\vec{Z}_3$ can be obtained through the equations below:

$$\vec{Z}_1 = \vec{Z}_\alpha - \vec{M}_1.\vec{P}_\alpha \tag{7}$$

$$\vec{Z}_2 = \vec{Z}_\beta - \vec{M}_2.\vec{P}_\beta \tag{8}$$

$$\vec{Z}_3 = \vec{Z}_\delta - \vec{M}_3.\vec{P}_\delta \tag{9}$$

In iteration i, the best 3 solutions are respectively, $\vec{Z}_1$, $\vec{Z}_2$ and $\vec{Z}_3$. Where, the values of $\vec{P}_\alpha$, $\vec{P}_\beta$ and $\vec{P}_\delta$ are as below:

$$\vec{P}_\alpha = \left| \vec{N}_1.\vec{Z}_\alpha - \vec{Z} \right| \tag{10}$$

$$\vec{P}_\beta = \left| \vec{N}_2.\vec{Z}_\beta - \vec{Z} \right| \tag{11}$$

$$\vec{P}_\delta = \left| \vec{N}_3.\vec{Z}_\delta - \vec{Z} \right| \tag{12}$$

In this study, GWO is modified to suit it with stock data and train the MLP using GWO for achieving a precise neural network and for removing the drawback of the back-propagation algorithm. Here, two important steps are included for this method that comprises: firstly, train the neural network using GWO to determine the optimal initial weights and secondly, test the results of the proposed GWO approach. This approach can improve the efficiency of the back-propagation to seek global optima in the search space. For the proposed GWO approach, the weights are achieved as a vector of variables.

For this approach, Root Mean Square Error (RMSE) is the cost function which can determine the error between the actual value and predicted value which can be expressed as below:

$$RMSE = \sqrt{\sum_{i=1}^{n}(a_i - p_i)^2} \tag{13}$$

Where, n is the number of observations, a_i is the number of actual values and p_i is the number of predicted values from neural network. The lower RMSE value is expected for determining acceptable prediction and making the model acceptable.

3.2 Pseudocode of GWO

The steps used in MGWO are as follows:

Step 1: Initialize the grey wolf population, P_i (i = 1,2,3,…..n)

Step 2: Initialize q, M and N

Step 3: Calcualte the fitness of each search agent, where, $Z_\alpha, Z_\beta, Z_\delta$ are the best, second best and
       third best search agent consecutively

Step 4: While i<MaxIterations
       update the position of current search agent
    for each search agent by equation 6
        Update q, M and N
        Calculate the fitness of all search agents
        Update Z_α, Z_β, and Z_δ
        i = i + 1
       End While

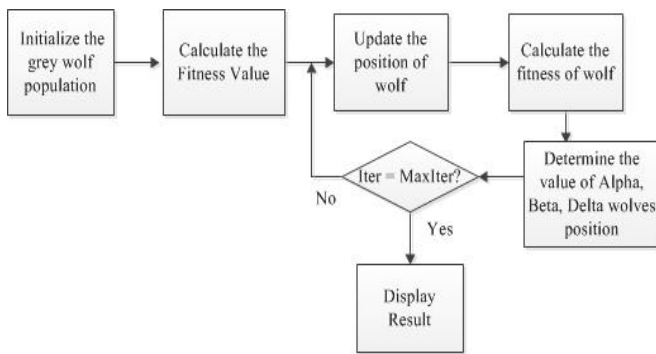Step 5: Return Z_α

Steps of MGWO algorithm are as follows:

Figure 1. Steps of GWO Algorithm

4. Architecture of the proposed System

Literature has supported the application of data mining, classification and neural network technique for the prediction of non-linear stock market. In this study, we have collected data from New York Stock Exchange (NYSE) and NASDAQ through Yahoo Finance and from Dhaka Stock Exchange (DSE) through DSE Library/Website. We have collected about 6 years of stock data from 2011-2016. Kumar et al. have used five input parameters in predicting stock price applying Multi-Layer Perceptron with Back propagation neural network learning algorithm [18]. Many other prior researches also support us to consider the following parameter in stock prediction:

Open price, Close Price, Low Price, High Price, and Volume

We have pre-processed the data and then applied k-means clustering data mining technique to observe the growth of share price for each company. Consequently, k-means clustering helps us to identify the best company in terms of volume transacted, fluctuation of high price and low price of the shares. Classification algorithm GWO is applied on the selected company at this stage to determine the classification rate and decide whether the company can be selected for prediction or not. Then, we apply the optimized neural network Nonlinear Autoregressive Exogenous (NARX) is applied to predict the stock price, whether buy, sell or hold the shares. Finally, the prediction is evaluated by the cost function as explained in equation (13). We have also performed a comparative study of the prediction between prominent stock market and emerging stock market. The proposed system includes the following architecture:
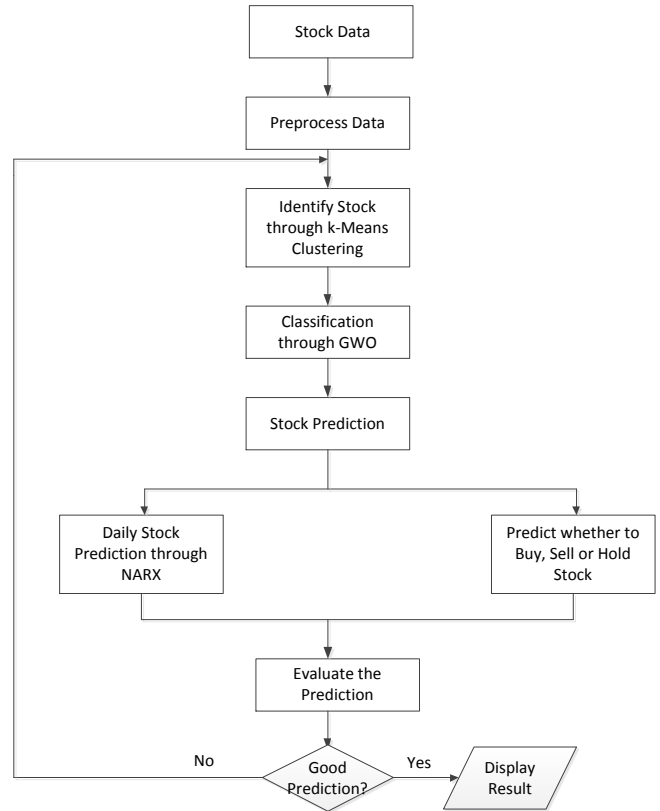

Figure 2. Architecture of the proposed prediction system

5. Experimental Results

The experiment for this work has been conducted in several stages. Firstly, the collected stock data is pre-processed and then clustering is applied to select the share to perform prediction. The raw stock data is pre-processed as per the Table 1 and Table 2 below to use it for clustering:

Table 1. Pre-processed stock data from NASDAQ for K-means clustering.

Company Name	Trade Date	Low Price	High Price	Open Price	Closing Price	Volume of Trades
Microsoft	2014-1-2	37.1	37.4	37.35	37.16	30632200
Hudson Tech.	2014-1-2	3.59	3.74	3.70	3.69	192700
AMD	2014-1-2	3.84	3.98	3.85	3.95	2054800
NVIDIA Corp.	2014-1-2	15.72	15.98	15.92	15.86	6502300
Spartan Motors	2014-1-2	6.50	6.79	6.69	6.71	66400

Table 2. Pre-processed stock data from DSE for K-means clustering.

Company Name	Trade Date	Low Price	High Price	Average Price	Closing Price	Volume of Trades
ACI	2013-02-12	140	147	145.07	145.04	8800
BXPHARMA	2013-02-12	56.70	65	62.49	62.51	568223
GLAXOSMITH	2013-02-12	547	547.10	547.01	547.01	218805
RECKITBEN	2013-02-12	725.00	725.00	725.0	725.00	500
RENATA	2013-02-12	788.20	798.00	793.71	793.85	4350
SQURPHARMA	2013-02-12	170.00	183.80	182.89	182.94	224,634

5.1 Clustering Result

K-means clustering places the data in two groups as we have selected two clusters to categorize the data based on number of volumes being traded. Clustering categorizes in a way that one cluster includes the pattern of data with highly growing, whereas another cluster contains the pattern with slowly growing. The following Table 3 lists the overall clustering data, though we will be using the highly growing data for classification.

Table 3. k-means Clustering Process

Attribute	Full Data (17092)	Cluster #0 (12091)	Cluster #1 (5001)
Company Name	ACI	ACI	SQURPHARMA
LOWPRC	1309.27	280.95	3795.47
HIPRC	1348.31	293.25	3899.15
AVGPRC	1330.63	288.11	3851.13
CLSPRC	1330.63	287.91	3851.63
TRDVOL	110681.94	147207.60	22373.25

The close observation of data pattern and also considering the growth, it has been determined that ACI has high growth and hence it has been selected for classification through GWO.

## 5.2 Classification Result through GWO

The Table 4 demonstrates the sample dataset used for classification, where 1 is used for Hold, 2 for Buy and 3 for Sell.

Table 4. Dataset used for Classification Process

Instrument	Low Price	High Price	Average Price	Closing Price	Volume	Decision
ACI, DSE	499	510.90	507.03	507.07	62446	1(Hold)
	490	510.09	504.86	504.76	61170	2(Buy)
	478	504.00	500.83	500.71	65774	2
	490	502.80	499.15	499.16	54554	1
	490	508.00	502.03	502.23	75680	1
	501	505.00	502.69	502.71	38720	1
	496.10	505.00	499.16	499.17	52370	1
	490.50	499.80	494.08	493.96	52000	2
	480	500.00	491.54	491.51	48281	1
	470	495.00	491.04	491.13	28930	2
	480	500.00	497.24	497.19	29702	3(Sell)

The classification through GWO produces the result indicated in Figure 3 and the classification rate through the algorithm is 100, which indicates that the proposed stock data is good for prediction [19,20].

fx>> main

fobj =

@MLP_ACI

ans =

classification rate = 100
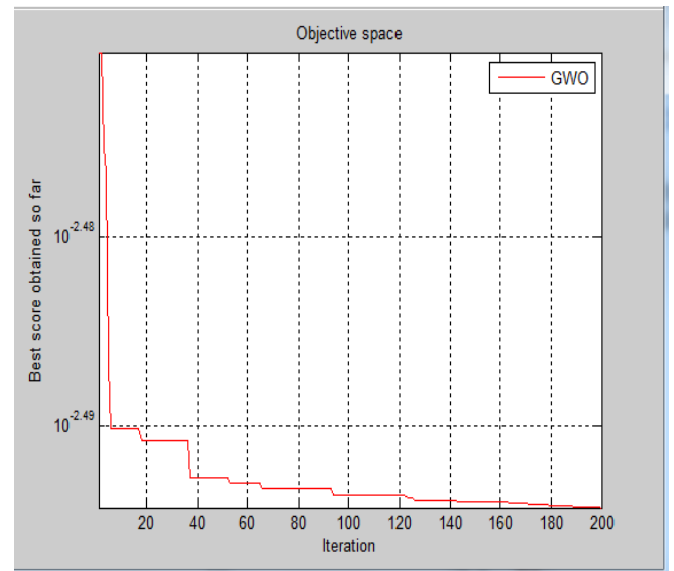
Figure 3. Result through MGWO

5.3 Result through Non Linear Autoregressive Exogenous (NARX) neural network model

The six years of daily stock price for companies from NYSE, Nasdaq, Bursa Malaysia and DSE has been provided to predict the High Price through Non Linear Autoregressive Exogenous neural network model which categorizes the data into Training, Validation and Test Set as indicated in Table 5. Where, Training Set consists 70% of total data, Validation Set contains 15% of data and Test Set holds 15% of data. The neural network model indicated in Figure 4 contains 10 Hidden Neurons and Number of Delays (d) is 2 in this architecture.

Table 5. Categorization of Data Sets

Dataset Type	Amount
Training Set	70% of the target timesteps
Validation	15% of the target timesteps
Testing	15% of the target timesteps
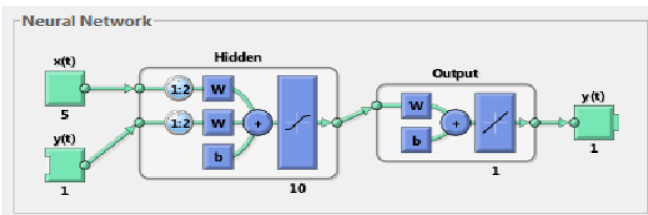
Figure 4. Architecture of the neural network Model

The equation 14 is applied to perform the stock prediction using Non-Linear Autoregressive Exogenous neural network.

$$P(k) = f(A(d-1),.....,A(d-k),P(k-1), ...,P(d-k)) \quad (14)$$

Where, $P(k)$ is the predicted Series, k is the time interval, d is the past values, $P(k)$ is the actual price of the stock and $A(k)$ is the target price to be predicted. The Table 6 lists Actual Value Vs Predicted value of stock price through Non Linear Autoregressive neural network algorithm for different stock market for a day, where the stock value for NYSE and Nasdaq in USD, stock value for Bursa Malaysia in Ringgit Malaysia (RM) and stock value for DSE in Bangladesh Taka (BDT).

Table 6. Actual Price Vs Predicted Price for various stock markets worldwide

Company Name	Stock Market	Date	Actual	Predicted	Forecasting Error (%)
Microsoft	Nasdaq, USA	2016-12-01	60.15	59.71	0.72
IBM	NYSE,USA	2016-12-01	162.2	160.966	0.76
Digi	Bursa Malaysia	2016-12-01	4.96	4.986	0.536
ACI	DSE, Bangladesh	2016-12-01	414	415.37	0.33

5.4 Performance Evaluation

The performance of the predicted price is evaluated through the equation 13 for various instruments of different stock markets produced the Root Mean Square Error (RMSE) value as indicated in Table 7 and the forecasting error related to the prediction is indicated in Figure 5. The RMSE value close to 0 indicates no error and prediction is completely acceptable. The forecasting error remains lower and reasonable through the proposed system. Hence, it can be concluded that the proposed hybridized system can be applied for stock prediction that can also reduce the error.

Table 7. RMSE value for instrument of various stock markets

Company Name	Stock Market	RMSE
Microsoft	Nasdaq,USA	0.7624
IBM	NYSE,USA	0.0199
Digi	Bursa Malaysia	0.0645
ACI	DSE, Bangladesh	0.2249
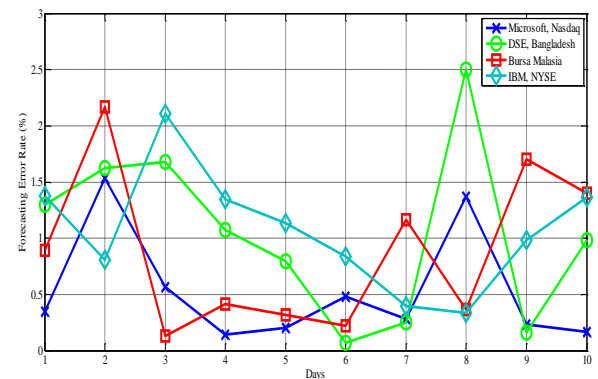
Figure 5. Comparison of Forecasting Error for various stock markets

The performance of a neural network can be also evaluated through generalization capability [21]. The proposed prediction system can perform well for various stock markets and hence it has good generalization capacity.

6. Conclusion

This research attempts to construct a hybridized model for stock prediction and compare the performance of prediction between various stock markets. The empirical results through the hybridized model are satisfactory and the error rate also remained reasonable. Investors look for making good profit through stock investment and hence the proposed system can assist the investors in stock investment decision.

For some instrument the forecasting error remained high which is due to the moving of stock trend through various other factors. Our future plan is to work with the effect of various factors in stock price movement and selection of

parameters. We would also take into consideration the influence of company news either positive or negative in stock price movement. Stock indices are another important component of stock prediction which can also be predicted.

7. Acknowledgement

References

[1] Bonde Ganesh, Khaled Rasheed, *Proceedings of the International Conference on Genetic and Evolutionary Methods (GEM); Athens: 1-6. Athens: The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp)*, 2012.

[2] Coupelon O., "Neural network modeling for stock movement prediction," *Neural Network Modeling for Stock Movement Prediction*, 2007.

[3] Jukic, N., Vrbsky, S., nestorov, S., *Database Systems: Introduction to Database and Data Wearhouses: Pearson*, 2014.

[4] K. Senthamarai Kannan, P. Sailapathi Sekar, M.Mohamed Sathik, P. Arumugam, *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS); Hong Kong: 1-5, Financial Stock Market Forecast using Data Mining Techniques*, 2010.

[5] Armano G. Marchesi M., Murru A., "A hybrid genetic-neural architecture for stock indexes forecasting," *International journal of Information Science,* Vol. 170. pp, 3-33, 2005.

[6] Dase R.K. , Pawar D.D., "Application of Artificial Neural Network for stock market predictions: A review of literature," *International Journal of Machine Intelligence,* ISSN: 0975–2927. Volume 2, Issue 2, pp. 14-17, 2010.

[7] Debashish Das, Ali Safa Sadiq, Noraziah Binti Ahmad, "An Efficient Time Series Analysis for Pharmaceutical Sector Stock Prediction by Applying Hybridization of Data Mining and Neural Network Technique," *Indian Journal of Science and Technology,* Vol. 9(21), pp. 1-7, June 2016.

[8] Debashish Das, Ali Safa Sadiq, Noraziah Binti Ahmad, Jaime Lloret, "Stock Market Prediction with Big Data Through Hybridization of Data Mining and Optimized Neural Network Techniques," *J. of Mult.-Valued Logic & Soft Computing,* In Press, Vol. 0, pp. 1–25, 2017.

[9] Seyedali Mirjalili, Seyed Mohammad Mirjalili, Andrew Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software,* Vol. 69, pp. 46–61, 2014.

[10] Selene Yue Xu, "Stock Price Forecasting Using Information from Yahoo Finance and Google Trend," *Econ Honor Thesis. UC Berkeley*, 2014.

[11] Amit B. Suthar, Hiral R. Patel, Satyen M. Parikh, "A Comparative Study on Financial Stock Market Prediction Models," *The International Journal of Engineering And Science (IJES),* Volume 1, Issue 2, pp. 188-191, 2012.

[12] Tomasz Kozdraj, "USING ARTIFICIAL NEURAL NETWORKS TO PREDICT STOCK PRICES," *ACTA UNIVERSITATIS LODZIENSIS. FOLIA OECONOMICA* 225, pp. 295-308, 2009.

[13] Ballings, M., Van den Poel, D., Hespeels, N., Gryp, R., "Evaluating multiple classifiers for stock price direction prediction," *Expert Systems with Applications*, Vol. 42, pp. 7046-7056, 2015.

[14] Lamartine Almeida Teixeira, Adriano Lorena Inácio de Oliveira, "A method for automatic stock trading combining technical analysis and nearest neighbor classification," *Expert Systems with Applications,* Issue 37, pp. 6885–6890, 2009.

[15] B. Narayanan, M. Govindarajan, "Prediction of Stock Market using Ensemble Model," *International Journal of Computer Applications,* Volume 128, No.1, pp. 18-21, 2015.

[16] Tsai, C.-F., Wang, S.-P., "Stock Price Forecasting by Hybrid Machine Learning Techniques," *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS)*, Hong Kong, 2009.

[17] G. S. Navale, Nishant Dudhwala, Kunal Jadhav, Pawan Gabda, Brij Kishor Vihangam, "Prediction of Stock Market using Data Mining and Artificial Intelligence," *International Journal of Computer Applications,* Volume 134, No.12, pp. 9-11, 2016.

[18] K.K.Suresh Kumar, N.M.Elango, "Performance analysis of stock prediction using ANN," *Global journal of Computer Science and Technology*, Volume 12, issue 1, 2012.

[19] Seyedali Mirjalili,, Ali Safa Sadiq, "Magnetic optimization algorithm for training multi layer perceptron," IEEE 3rd international conference on Communication software and networks (ICCSN). 27-29 May. Xi'an, China. pp: 42-46, 2011.

[20] Seyedali Mirjalili, Andrew Lewis, Ali Safa Sadiq, "Autonomous particles groups for particle swarm optimization," *Arabian Journal for Science and Engineering.* Vol. 39. No. 6. Pp: 4683-4697, 2014.

[21] Kaastra, I., Boyd, M.S, "Designing a Neural Network for Forecasting Financial and Economic Time Series," *Neurocomputing*, 10(3), 215-236, 1996.