

Forecasting the Saudi Arabia stock prices based on artificial neural networks model

S. O. Olatunji^{1,*}, Mohammad Saad Al-Ahmadi², Moustafa Elshafei², Yaser Ahmed Fallatah²

¹Computer Science Department, Adekunle Ajasin University, Akungba Akoko, Nigeria

²King Fahd Univ. of Pet. & Mineral, (KFUPM), Dhahran, Saudi Arabia

Email address:

oluolatumji.aadam@gmail.com(S. O. Olatunji), alahmadi@kfupm.edu.sa(M. S. Al-Ahmadi), Shafei@MIT.edu(M. Elshafei), fallatah@kfupm.edu.sa(Y. A. Fallatah)

To cite this article:

S. O. Olatunji, Mohammad Saad Al-Ahmadi, Moustafa Elshafei, Yaser Ahmed Fallatah. Forecasting the Saudi Arabia Stock Prices Based on Artificial Neural Networks Model. *International Journal of Intelligent Information Systems*. Vol. 2, No. 5, 2013, pp. 77-86.

doi: 10.11648/j.ijis.20130205.12

Abstract: In this paper, we have proposed artificial neural network for the prediction of Saudi stock market. The proposed predictions model, with its high degree of accuracy, could be used as investment advisor for the investors and traders in the Saudi stock market. The proposed model is based mainly on Saudi Stock market historical data covering a large span of time. Achieving reasonable accuracy rate of predication models will surely facilitate an increased confidence in the investment in the Saudi stock market. We have only used the closing price of the stock as the stock variable considered for input to the system. The number of windows gap to determine the numbers of previous days to be used in predicting the next day closing price data has been choosing based on experimental simulation carried out to determine the best possible value. Our results indicated that the proposed ANN model predicts the next day closing price stock market value with a very low RMSE down to 1.8174, very low MAD down to 18.2835, very low MAPE of down to 1.6476 and very high correlation coefficient of up to 99.9% for the test set, which is an indication that the model adequately mimics the trend of the market in its prediction. This performance is really encouraging and thus the proposed system will impact positively on the analysis and prediction of Saudi stock market in general.

Keywords: Stock Markets, Stock Prices, Prediction Models, Forecasting, Artificial Neural Networks, Saudi Arabia

1. Introduction

Several works have been presented on the use of soft computing technique for the forecasting of stock prices. But the case of Saudi stock market forecasting has not been explored and thus it is a potential/ideal research target to explore the forecasting of the ever growing Saudi stock market using artificial intelligence and soft computing techniques.

In this work, we have proposed the use of artificial neural network as a predictive solution for forecasting the Saudi Arabia stock market prices. Generally, the standard neural networks have the capacity to learn existing relationships among data. Multilayer perceptron (MLP) is the most widely used neural network with the Back-Propagation Algorithm networks. This type of neural networks is excellent at both prediction and classification and it has been widely used in several fields often with better results.

The proposed model has been tested on three different companies selected as the major determinants of Saudi stock market. These companies include SABIC, Saudi Telecommunication Company (STC) and AlRajhi Bank. Empirical results from simulation have demonstrated the viability of the proposed model in predicting Saudi Arabia stock markets. The results from the empirical work carried indicated that the proposed ANN model predicts the next day closing price stock market value with a very low RMSE down to 1.8174, very low MAD down to 18.2835, very low MAPE down to 1.6476 and very high correlation coefficient of up to 99.9% for the test set, which is an indication that the proposed model adequately mimics the pattern and trend of the market (the actual target closing price values) in its prediction.

This remaining part of the article is organized as follows. Section 2 presents a detailed review of related works. Section 3 contains the proposed model's description and details. Empirical works, data set description and results

discussion are presented in section 4. Section 5 contains conclusion and future work recommendations emanating from this study.

2. Literature Review

Hadavandi, Shavandi et al. (Hadavandi, Shavandi et al. 2010) presents an integrated approach based on genetic fuzzy systems (GFS) and artificial neural networks (ANN) for constructing a stock price forecasting expert system, in order to fulfill the central purpose of achieving best results using minimum required input data and the least complex stock market model. At first, they use stepwise regression analysis (SRA) to determine factors which have most influence on stock prices. At the next stage they divided the available raw data into k clusters by means of self-organizing map (SOM) neural networks. Finally, all clusters were fed into independent GFS models with the ability of rule base extraction and data base tuning. The capability of the proposed approach was evaluated by applying it on stock price data gathered from IT and Airlines sectors and then comparing the outcomes with previous stock price forecasting methods using mean absolute percentage error (MAPE). They reported that the results show that the proposed approach outperforms all previous methods, and thus they concluded that their proposed method can be considered as a suitable tool for stock price forecasting problems.

Kara, Acar Boyacioglu et al (Kara, Acar Boyacioglu et al. 2010) presented the use of artificial neural networks and support vector machines in Predicting direction of stock price index movement for the sample of the Istanbul Stock Exchange. In this study, they attempted to develop two efficient models and compared their performances in predicting the direction of movement in the daily Istanbul Stock Exchange (ISE) National 100 Index. The models are based on artificial neural networks (ANN) and support vector machines (SVM). Ten technical indicators were selected as inputs of the proposed models. Two comprehensive parameter setting experiments for both models were performed to improve their prediction performances. The final experimental results showed that average performance of ANN model (75.74%) was found significantly better than that of SVM model (71.52%). Ten technical indicators were used as input variables. A tangent sigmoid transfer function was selected on the hidden layer. On the other hand, a logistic sigmoid transfer function was used on the output layer. That is, the outputs of the model will vary between 0 and 1. If the output value is smaller than 0.5, then the corresponding case is classified as a decreasing direction; otherwise, it is classified as an increasing direction in movement. The research data used in this study is the direction of daily closing price movement in the ISE National 100 Index. The entire data set covers the period from January 2, 1997 to December 31, 2007. The total number of cases is 2733 trading days. The number of cases with increasing direction is 1440 while the

number of cases with decreasing direction is 1293. That is, 52.7% of the all cases have an increasing direction and 47.3% of the all cases have a decreasing direction. The historical data was obtained from the technical analysis module of Matriks gold 2.4.0., produced by Matriks Information Delivery Services Inc.

Shen, Guo et al. (Shen, Guo et al. 2010) made use of a radial basis function neural network (RBFNN) to train data and forecast the stock indices of the Shanghai Stock Exchange. They also made use of the artificial fish swarm algorithm (AFSA) to optimize RBF. In order to further increase forecasting efficiency, a K-means clustering algorithm is optimized by AFSA in the learning process of RBF. As a means of verifying the usefulness of their proposed algorithm, they compared the forecasting results of RBF optimized by AFSA, genetic algorithms (GA) and particle swarm optimization (PSO), as well as forecasting results of ARIMA, BP and support vector machine (SVM). They concluded that the results that emanated from the experiments indicate that RBF optimized by AFSA is an easy-to-use algorithm with considerable accuracy.

Furthermore, Chih-Fong and Yu-Chieh (Chih-Fong and Yu-Chieh 2010) combined multiple feature selection methods to identify more representative variables for better prediction and to effectively predict stock price for investors. In particular, three well-known feature selection methods, which are Principal Component Analysis (PCA), Genetic Algorithms (GA) and decision trees (CART), are used. The combination methods to filter out unrepresentative variables are based on union, intersection, and multi-intersection strategies. For the prediction model, the back-propagation neural network is developed. It was concluded that their experimental results show that the intersection between PCA and GA and the multi-intersection of PCA, GA, and CART perform the best, which are of 79% and 78.98% accuracy respectively. In addition, these two combined feature selection methods filter out near 80% unrepresentative features from 85 original variables, resulting in 14 and 17 important features respectively. They also opined that these variables are the important factors for stock prediction and can be used for future investment decisions.

One of the recent work in this direction is that presented by Esfahanipour and Aghamiri (Esfahanipour and Aghamiri 2010). In this work, the authors developed Neuro-Fuzzy Inference System adopted on a Takagi–Sugeno–Kang (TSK) type Fuzzy Rule Based System for stock price prediction. The TSK fuzzy model applies the technical index as the input variables and the consequent part is a linear combination of the input variables. Fuzzy C-Mean clustering was also implemented for identifying number of rules while initial membership function of the premise part was approximately defined as Gaussian function and the TSK parameters was tuned by Adaptive Neuro-Fuzzy Inference System (ANFIS). Their proposed model is tested on the Tehran Stock Exchange Indexes (TEPIX) and it was reported that the model perform successful forecasting

using several experimental tests from different sectors while achieving an accuracy of up to 97.8%

Lee (Lee 2009) developed a prediction model based on support vector machine (SVM) with a hybrid feature selection method to predict the trend of stock markets. The proposed hybrid feature selection method, named F-score and Supported Sequential Forward Search (F_SSFS), combines the advantages of filter methods and wrapper methods to select the optimal feature subset from original feature set. The authors evaluated the prediction accuracy of this SVM-based model combined with F_SSFS by comparing its performance with back-propagation neural network (BPNN) along with three commonly used feature selection methods including Information gain, Symmetrical uncertainty, and Correlation-based feature selection via paired t-test. It was reported that SVM outperforms BPN for the problem of stock trend prediction, and that the proposed SVM-based model combined with F_SSFS has the highest level of accuracies and generalization performance in comparison with the other three feature selection methods.

Another interesting work was presented by Boyacioglu and Avci (Boyacioglu and Avci 2010) where the predictability of stock market return using Adaptive Network-Based Fuzzy Inference System (ANFIS) was investigated. The objective of this study was to determine whether an ANFIS algorithm is capable of accurately predicting stock market return. They model and predict the return on stock price index of the Istanbul Stock Exchange (ISE) using ANFIS. Six macroeconomic variables and three indices as input variables were made use of and it was reported that experimental results reveal that the model successfully forecasts the monthly return of ISE National 100 Index with an accuracy rate of 98.3%.

Chang, Liu et al. (Chang, Liu et al. 2009) presented an integrated system, CBDWNN by combining dynamic time windows, case based reasoning (CBR), and neural network for stock trading prediction and it includes three different stages: (1) screening out potential stocks and the important influential factors; (2) using back propagation network (BPN) to predict the buy/sell points (wave peak and wave trough) of stock price and (3) adopting case based dynamic window (CBDW) to further improve the forecasting results from BPN. The system developed in this research is a first attempt in the literature to predict the sell/buy decision points instead of stock price itself. In this study, nine different stocks with different trends, i.e., upward, downward and steady, are studied and one individual stock (AUO) was studied as case example. The empirical results show that the CBDW can assist the BPN to reduce the false alarm of buying or selling decisions. It was also reported that the rates of return for upward, steady, and downward trend stocks are higher than 93.57%, 37.75%, and 46.62%, respectively. These results, according to the authors, are all very promising and better than using CBR or BPN alone.

Another related work was presented by Lean Yu, Huanhuan Chen et al (Lean Yu, Huanhuan Chen et al. 2009)

where the authors proposed an evolving least squares support vectormachine (LSSVM) learning paradigm with a mixed kernel to predict stock market trends. In the proposed learning paradigm, a genetic algorithm (GA) is first used to select input features for LSSVM learning, thereafter another GA is used for parameters optimization of LSSVM. Finally, the evolving LSSVM learning paradigm with best feature subset, optimal parameters, and a mixed kernel is used to predict stock market movement direction in terms of historical data series. They later applied this on three important stock indices, S&P 500Index, Dow Jones Industrial Average (DJIA) Index, and New York Stock Exchange (NYSE) Index. They reported that their experimental results indicated that the proposed evolving LSSVM performed better than other forecasting models considered in the work.

Yet another related work was presented in the work of Guresen, kayakutlu (Guresen, kayakutlu et al.). In this paper, the authors evaluate the effectiveness of neural network models which are known to be dynamic and effective in stock-market predictions. The models analyzed include multi-layer perceptron (MLP), dynamic artificial neural network (DAN2) and the hybrid neural networks which use generalized autoregressive conditional heteroscedasticity (GARCH) to extract new input variables. The comparison for each model is done using Mean Square Error (MSE) and Mean Absolute Deviate (MAD) using based on real exchange daily rate values of NASDAQ Stock Exchange index.

Md. Rafiul (Md. Rafiul 2009) presents a novel combination of the hidden Markov model (HMM) and the fuzzy models for forecasting stock market data. This research follows the pattern of using the historical data to generate a 'one-day-ahead' forecast. This paper uses this approach to identify data patterns by using the HMM and then uses fuzzy logic to obtain a forecast value. The HMM's log-likelihood for each of the input data vectors is used to partition the data space. Each of the divided data spaces is then used to generate a fuzzy rule. The fuzzy model developed from this approach is tested on stock market data drawn from different sectors. It was reported that experimental results clearly show an improved forecasting accuracy compared to other forecasting models such as, ARIMA, artificial neural network (ANN) and another HMM-based forecasting model.

Meysam, Mohsen et al. (Meysam, Mohsen et al. 2011) implemented the design of experiment (DOE) technique to identify the significant parameters in the design of adaptive neuro-fuzzy inference systems (ANFIS) for stock price prediction.

Mabu, Chen et al. (Mabu, Chen et al. 2009) applied RasID-GA to predict stock market prices using the multi-branch feed forward neural networks. The authors predicted the next day's closing stock price with several past closing stock prices. The stock prices of 20 brands for 720 days were used in order to evaluate the generalization ability of the proposed method

Based on the fact that stock trend prediction is regarded as a challenging task and the recent confirmation by researches that a successful feature selection method can improve the prediction accuracy of stock market, Ni, Ni et al (Ni, Ni et al. 2011) hybridizes fractal feature selection method and support vector machine to predict the direction of daily stock price index. Fractal feature selection method is suitable for solving the nonlinear problem and it can exactly spot how many important features to be used. To evaluate the prediction accuracy of the proposed method, its performance was compared with other five commonly used feature selection methods. The reported results show that the fractal feature selection method selects the relatively smaller number of features and at the same time achieves the best average prediction accuracy.

Bajestani and Zare (Bajestani and Zare 2011) presented a new method to forecast TAIEX based on a high-order type 2 fuzzy time series. Extra observations are modeled as type 2 fuzzy sets and fourth-order fuzzy time series in this study, while using extra observations to improve forecasting performance. The reported results indicated that the proposed model outperforms previous studies.

Finally, having presented a brief review of most of the recently published work in this area, we refer the reader to the work of Atsalakis and Valavanis (Atsalakis and Valavanis 2009) for further reviews on the earlier works in this direction. This paper (Atsalakis and Valavanis 2009) presented detailed review of earlier stock market forecasting based on soft computing techniques. Other recent works in this direction can be found in (Lean Yu, Huanhuan Chen et al. 2009; Boyacioglu and Avci 2010; Shen, Guo et al. 2010; Ni, Ni et al. 2011)

3. The Proposed Artificial Neural Network

Artificial neural networks is that technology which initially grew from the full understanding of some ideas and aspects about how biological systems work, especially the human brain.

A multi-layered network consists of numerous neurons, which are arranged into levels. Each level is interconnected with the one above and below it. The first layer receives external inputs and is aptly named the input layer. The last layer provides the classification solution, and is called the output layer. Sandwiched between the input and output layers are any number of hidden layers. It is believed that a three-layered network can accurately classify any non-linear function (Duda, Hart et al. 2001; Castillo, Fontenla-Romero et al. 2002).

3.1. How Neural Networks Operates

Neural systems are typically organized in layers. Layers are made up of a number of artificial neurons, which contain an activation function. Patterns are presented to the network via the input layer, which communicates to one, or

more hidden layers where the actual processing is done through a system of fully or partially weighted connections. The hidden layers are then linked to an output layer where the answer is output. Standard neural networks contain learning rules that modify the weights of the connections according to the input patterns. Generally, the standard neural networks have the capacity to learn, memorize and create relationships amongst data. Multilayer perceptron (MLP) is the most widely used neural network with the back-Propagation Algorithm networks (Tutu, Cukrowska et al. 2005; Abdolreza, Fahime et al. 2009). This type of neural networks is excellent at both prediction and classification. Using this neural network algorithm, both input and corresponding desired output data are given to the calibrating phase. As the network starts training, the input layer receives the input signals, and then it processes the data through the hidden layers until the output layer yielding the resulted outputs. These outputs are then compared with the desired outputs computing the error, which is back propagated through the system causing it to adjust the weights, which control the network.

Once a neural network is trained to a satisfactory level, it may then be used as an analytical tool on other data. To do this, the user no longer specifies any training runs and instead allows the network to work in forward propagation mode only. Inputs are presented to the network through input layer and processed by the middle layers as the training was done. The resulting output is retained and no back propagation occurs.

There are many ways in which the same type of neural network can be named. Generally, artificial neural networks are described as connectionist system (Ali 1994; Duda, Hart et al. 2001; Castillo, Guijarro-Berdiñas et al. 2006), because of the connections between individual processing nodes. They are also called parallel-distributed processing system, which emphasize the way in which the many nodes or neurons in a network operate in parallel. Sometimes they are referred to as adaptive system (Liang and Mendel 2000; Duda, Hart et al. 2001), because the values of these connections can change so that the network is performing more effectively and efficiently. Figure 1 shows the architecture of MLP networks with optional hidden neurons.

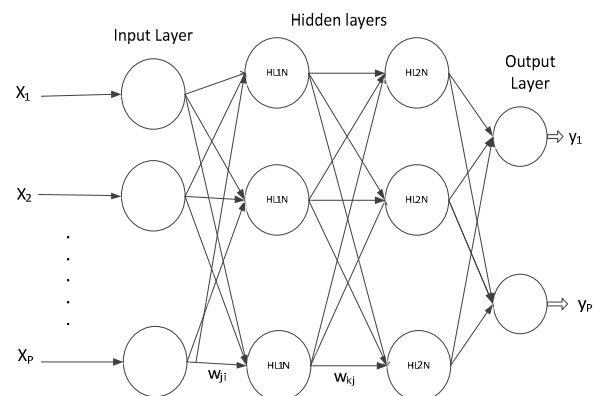


Figure 1: Multilayer Perceptron with two hidden layers

4. Empirical Studies and Discussions

We carried out the experimental studies using Saudi Arabia stock market closing prices. At this stage of the research we made use of the stock prices for three popular company namely Saudi Basic Industries Corporation (SABIC), Al Rajhi Banks, and Saudi telecoms (STC). Each of the dataset was divided into training set and testing set. The division was made using the stratify sampling approach where the data set is randomly divided using specified percentage for the training and testing sets. In the data partitioning phase, 70% was reserved for training the model while the remaining 30% was used to test the built model in order to ascertain the predictive capability of the proposed models.

4.1. Datasets Description

We made use of three different stock market dataset belonging to STC, SABIC and Al Rajhi bank. These are the three most prominent major companies on the stock market whose stock seriously impact on others.

SABIC Stock Dataset: These represent the closing price values for SABIC between 6th January 1993 and 22nd December 2010. The total number of daily closing price recorded and use is 2130.

STC Stock Dataset: These represent the closing price values for Saudi telecommunication company (STC) between 27th January 2003 and 22nd December 2010. The total number of daily closing price recorded is 2130.

Al Rajhi bank Stock Dataset: These represent the closing price values for Al Rajhi bank between 9th January 1993 and 22nd December 2010. The total number of daily closing price recorded is 2130.

4.2. Parameters search Procedures

To carry out the experiment, there is the need to choose the value for the windows delay. That is, the number of previous days to be used to predict the next day value for the stock. Since there is no agreed value for such we have decided to carried out a parameter search oriented simulation to look for the ideal values for this important parameter. To do this an experiment was set up to search for the optimum value between 2 and 50 while running training the model and testing it to determine the performance measure for each value of the parameter being searched for.

Based on the above described experiment we got the optimal delay window (window of past measurements) to be chosen as shown in figures 2 and 3. It must be noted that since the central idea to successful stock market prediction is achieving best results using minimum required input data and the least complex stock market model (Hadavandi, Shavandi et al. 2010), then the lesser the delay window values in the number of days to be used as predictor variables, the better. Fortunately the proposed model has been able to perform optimally at the least possible values of 2 days. This can be clearly visualized in the experimental results of contained in figures 2 and 3.

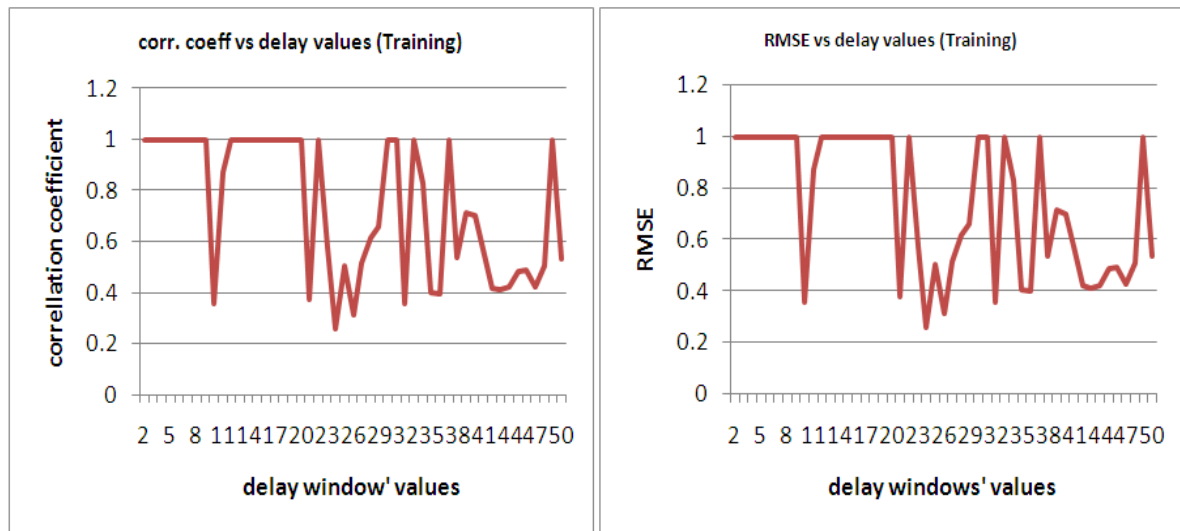


Fig. 2: Results outcome for the training set in determining the optimal values for window of past measurements

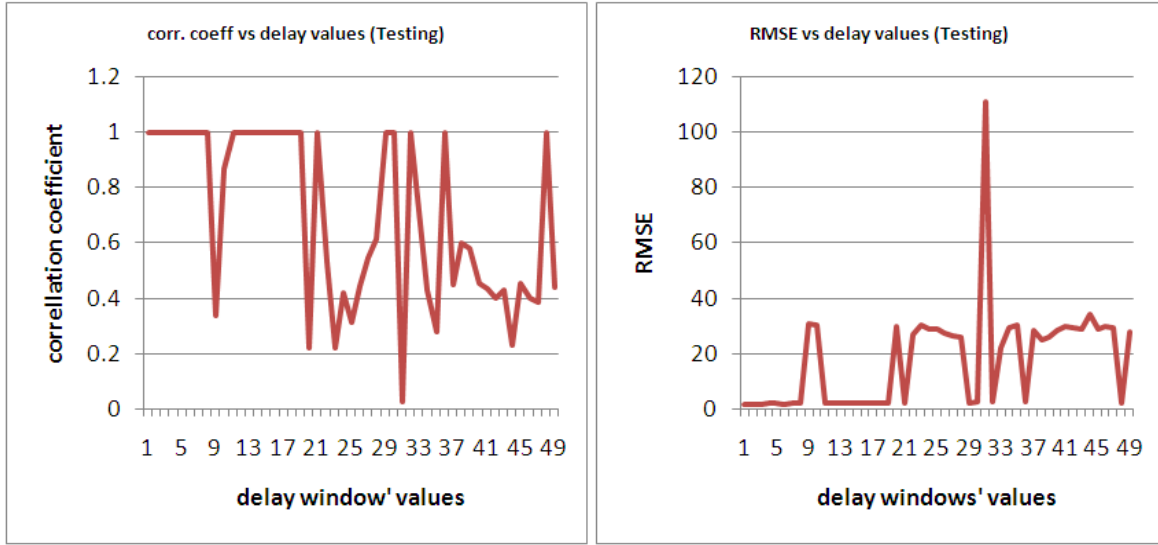


Fig. 3: Results outcome for the testing set in determining the optimal values for window of past measurements

From the result presented, we found out that between 2 and 9 window of past measurements, the model achieve the highest possible accuracy. Therefore we have decided to use the window value of 5 in the succeeding experiment since it provide the best result using one of the least possible number of input variables, That is, we are using the last five days stock prices to forecast the sixth day stock price.

Furthermore, to arrive at a unique and better ANN model, we did another parameter search to find the optimum combination for the architecture of the proposed model. This parameter search procedure is contained in the following sub-section and summarized in algorithm 1.

Algorithm 1: Algorithm for searching optimal parameter for ANN

```

Initialization;  $kx=0, vx=0, jx=0$ 

for  $j = 1 \rightarrow nf$ 
  for  $k = 1 \rightarrow nh$ 
     $accuracy = f(A_j(j))$  {Save the present performance accuracy for the
    current parameters selected}

    if  $accuracy$  is better than  $vx$  then  $vx = accuracy$ 

     $jx = j ; kx = k$  {Saving the pointer or index to the currently identified good
    parameters}
  end
end
end

```

4.3. Searching Optimal Parameters for ANN

To carry out an optimal-parameter search to ensure better model development, the procedures that have been presented in algorithm 1 below are as followed.

To develop this algorithm, we take the following assumption and use of symbols. Let the B represent a set consisting of all the available activation functions, with the

element that are of the form $B_j(k)$, where j is the number representing activation function, k is the number of layers, nf is the total number of activation functions available, and nh is the maximum number of hidden layer assumed. Also $accuracy$ represents performance measure taken, jx represents index for best performing activation function, kx represents index for the best performing number of layer. The algorithm then goes thus:

4.4. Performance Measure Employed

To evaluate the performance of the proposed ANN for Saudi stock price prediction, the most common quality measure in the field of time series forecasting, and they include correlation coefficient (R^2), root mean squared error (RMSE), mean absolute deviation (MAD) and mean absolute percentage error (MAPE). Their mathematical descriptions are briefly given as follows:

Mean absolute percentage error (MAPE) is measure of accuracy in a fitted time series value in statistics, specifically trending. It usually expresses accuracy as a percentage, and is defined by the formula:

$$M = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (14)$$

where A_t is the actual value and F_t is the forecast value.

The difference between A_t and F_t is divided by the actual value A_t again. The absolute value of this calculation is summed for every fitted or forecast point in time and divided again by the number of fitted points n . This makes it a percentage error so one can compare the error of fitted time series that differ in level.

Mean absolute deviation: The mean absolute deviation (MAD), also referred to as the mean deviation, is the mean of the absolute deviations of a set of data about the data's mean. In other words, it is the average distance of the data set from its mean during certain number of time periods.

The equation for MAD is as follows:

$MAD = 1/n \sum(|e_i|)$, where $e_i = F_i - D_i$

The correlation coefficient measures the statistical correlation between the predicted and actual values. This method is unique in model evaluations. The equation representing this measure is given thus:

$$R = \frac{\sum (y_a - y'_a)(y_p - y'_p)}{\sqrt{\sum (y_a - y'_a)^2 \sum (y_p - y'_p)^2}} \quad (15)$$

where y_a and y_p are the actual and predicted values while y'_a and y'_p are the mean of the actual and predicted values.

Root Mean-Squared Error

The mean-squared error is one of the most commonly used measures of success for numeric prediction. This value is computed by taking the average of the squared differences between each predicted value x_i and its corresponding actual value y_i . The root mean-squared error is simply the square root of the mean squared error. The formula is:

$$\sqrt{\frac{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}{n}} \quad (16)$$

where n is the size of the dataset used.

Correlation Coefficient

The correlation coefficient measures the statistical correlation between the predicted and actual values.

$$R^2 = \frac{\sum (y_a - y'_a)(y_p - y'_p)}{\sqrt{\sum (y_a - y'_a)^2 \sum (y_p - y'_p)^2}} \quad (17)$$

where y_a and y_p are the actual and predicted values while y'_a and y'_p are the mean of the actual and predicted values.

4.5. Results and Discussions

Table 1 presents final results from the proposed model using the optimized parameters in the earlier experiments. The results speaks for themselves as the proposed model predict the next day stock closing price with a very low RMSE and high correlation coefficient (R^2) of up to 99.9% for the test set, which is an indication that the model adequately mimics the trend of the market in its prediction. Specifically the proposed model achieve such a high correlation coefficient as high as 99.9% for the testing set for the SABIC stock while achieving 99.8 % for the STC stock price and 99.9% for the Al Rajhi bank stock price estimations. These results indicate the viability of using ANN as a predictive tool for the Saudi stock market.

Moreover, from the table 1, it became very clear that the proposed ANN model for the Saudi Arabia stock market prediction achieved higher accuracy level with as low as 1.8676 MAPE for the STC stock, and 2.4188 for the Alrajhi stock. Other recorded results including RMSE and MAD also indicated very low error values which are signs of the promising and satisfactory performance of the proposed ANN model.

Table 1: Performance measures

	R		RMSE		MAD		MAPE	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
SABIC stock	0.9992	0.9992	2.4061	2.4297	18.2835	18.4483	8.9984	8.3578
STC stock	0.9983	0.9976	1.8414	2.3344	24.1784	25.0995	1.6476	1.8676
Al-Rajhi bank Stock	0.9991	0.9987	1.8174	2.3634	5.1591	5.3966	1.9838	2.4188

The proposed model made use of five previous days values (5 window of past measurements) to arrive at such accurate next day stock value prediction. It is interesting to note that the proposed model met the central idea of a successful stock market prediction, which is to achieve the best possible accuracy using minimum required input data and the least complex stock market model. We have also generated trend plots showing the trends in the predicted values versus the target values, for testing results, as contained in figures 4, 5 and 6, which represent the trend

plots for SABIC, Rajhi-bank and STC respectively. The plots indicate very good close trend between the target and the predicted values for all the three datasets investigated. These can be easily noticed as the target and the predicted values mostly overlap in the figures, which is an indication of very good and close to the target predictive ability of the proposed model for the Saudi Arabia stock prices. Thus the proposed ANN model has proved to be a viable tool in successfully predicting the Saudi stock market where little or no research has been done of recent.

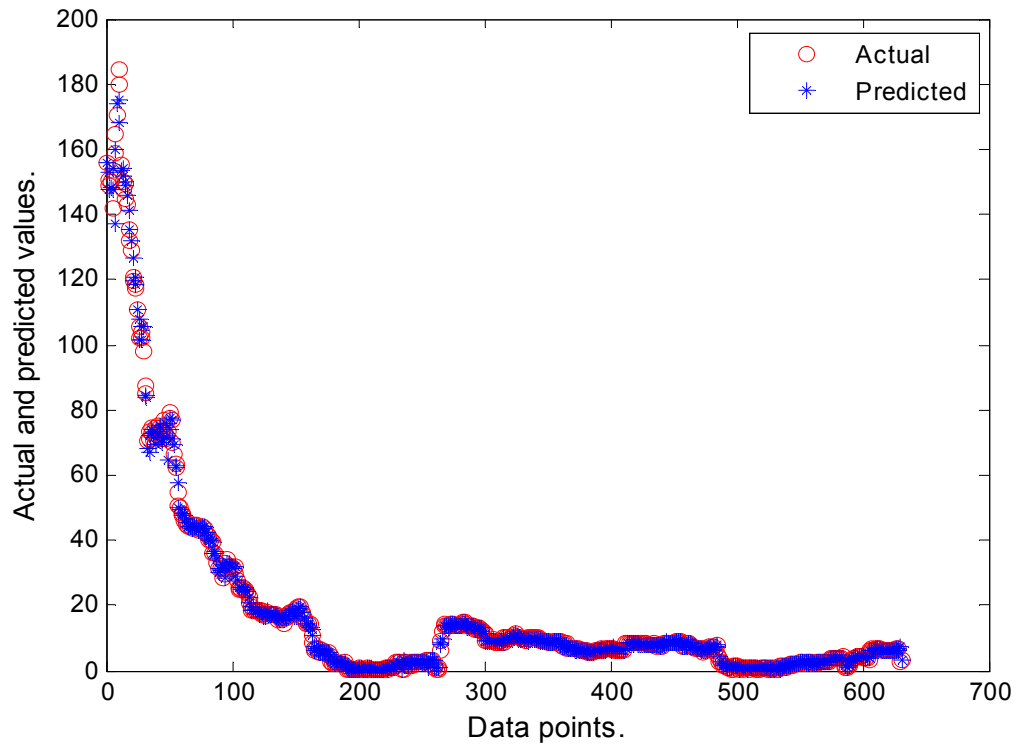


Figure 4: Trend of predicted vs. Target values for SABIC dataset

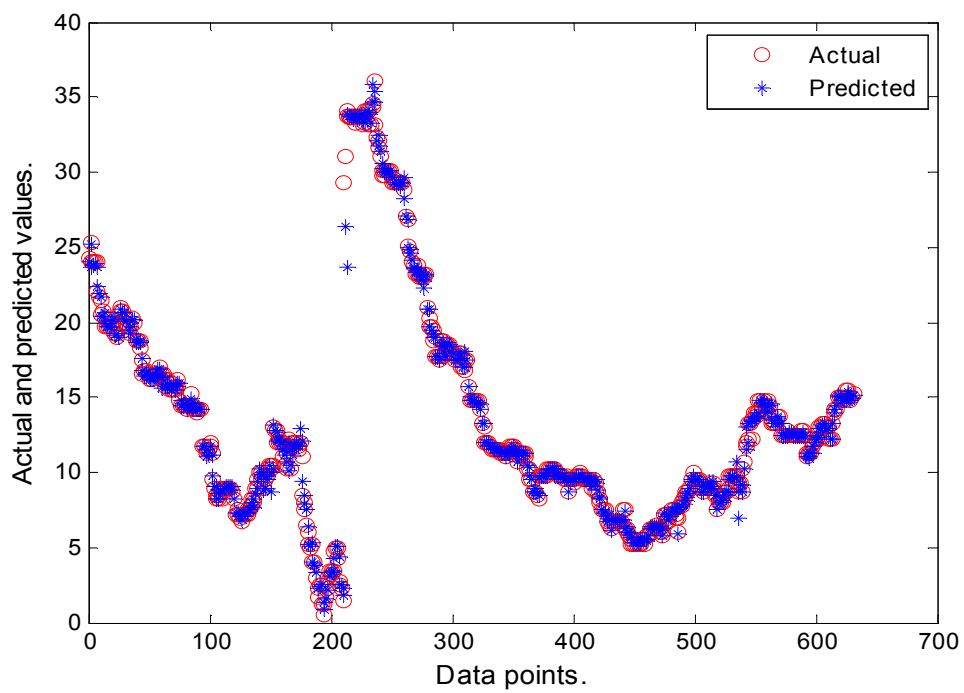


Figure 5: Trend of predicted vs. Target values for the Rajhi Bank dataset

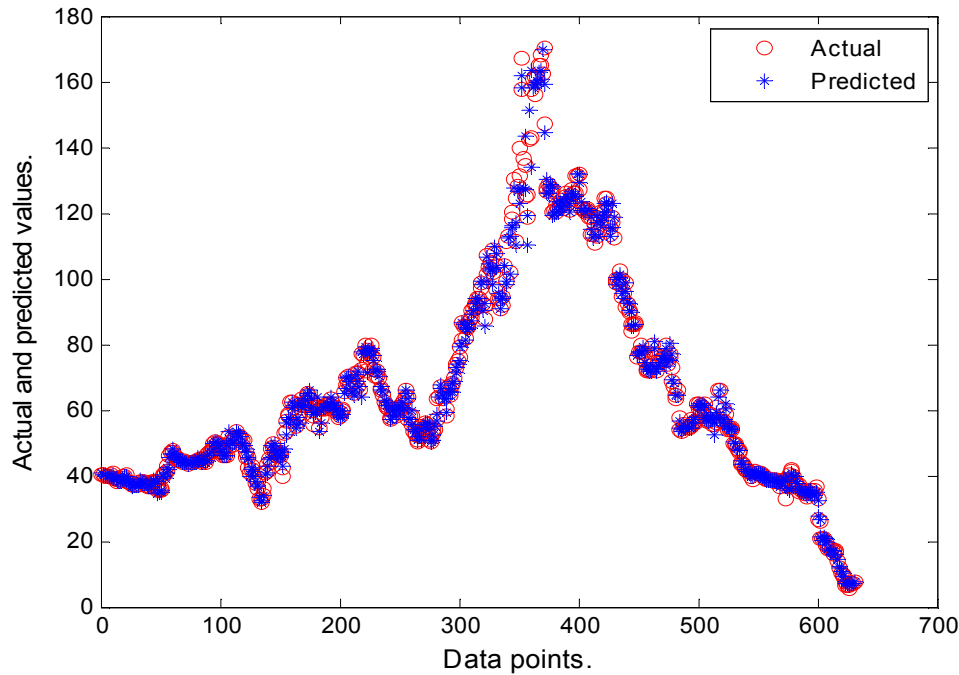


Figure 6: Trend of predicted vs. Target values for the STC dataset

5. Conclusion

We have presented an artificial neural network based model for predicting the Saudi Arabia stock market, which is a stock market that is yet to be explored as far as the use of artificial intelligence and soft computing methods are concerned. The proposed model has been tested on three different company selected as the major determinants of Saudi stock market. These companies include SABIC, Saudi Telecommunication Company (STC) and AlRajhi Bank. We investigated the optimal value for the window of past days to be used in predicting the next stock price. The optimized parameter is then used to build the final model and then tested using part of the data set reserved for such purpose as unseen data for testing the performance of the proposed model on new dataset. Empirical results from the experiments carried out have demonstrated the viability of the proposed model in predicting Saudi Arabia stock market. The results from this work are very accurate and encouraging with up to 99.9% accuracy in terms of correlation coefficient vis-à-vis how close the predicted values were in following the trend of the actual price values. Specifically, the proposed method achieve up to 99.9% for the testing set in the case of the SABIC stock while achieving 99.8 % for the STC stock price and 99.9% for the Al Rajhi bank stock price estimations. These results indicate the viability of using ANN as a predictive tool for the Saudi stock market. The RMSE, MAD and MAPE errors recorded very low values thereby further indicating that the model predicts with high accuracy. Therefore, this model can be implemented in a real-time trading system for stock price prediction during the trading period in order to determine the closing price. We hope to investigate the

capability of other soft computing techniques in the next stage of the research. Also it is part of our planned researches to extend the work to include prediction of opening price, minimum and maximum price and most importantly predicting Saudi stock index will be looked into.

Acknowledgment

The authors would like to acknowledge the support provided by King Abdulaziz City for Science and Technology (KACST) through the Science and Technology Unit at King Fahd University of Petroleum and Minerals (KFUPM) for funding this work through project number 09-INF769-04 as part of the National Science, Technology and Innovation Plan.

References

- [1] Abdolreza, M., P. Fahime, et al. (2009). "A New Approach Based on Artificial Neural Networks for Prediction of High Pressure Vapor-liquid Equilibrium." *Australian Journal of Basic and Applied Sciences* 3(3): 1851-1862.
- [2] Ali, J. K. (1994). *Neural Networks: A New Tool for the Petroleum Industry*. European Petroleum Computer Conference, Aberdeen, U.K.
- [3] Atsalakis, G. S. and K. P. Valavanis (2009). "Surveying stock market forecasting techniques - Part II: Soft computing methods." *Expert Systems with Applications* 36(3, Part 2): 5932-5941.
- [4] Bajestani, N. S. and A. Zare (2011). "Forecasting TAIEX using improved type 2 fuzzy time series." *Expert Systems with Applications* 38(5): 5816-5821.

- [5] Boyacioglu, M. A. and D. Avci (2010). "An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange." *Expert Systems with Applications* 37(12): 7908-7912.
- [6] Castillo, E., B. Guijarro-Berdiñan, et al. (2006). "A Very Fast Learning Method for Neural Networks Based on Sensitivity Analysis." *Journal of Machine Learning Research* 7: 1159-1182.
- [7] Castillo, O., O. Fontenla-Romero, et al. (2002). "A global optimum approach for one-layer neural networks." *Neural Computation* 14(6): 1429-1449.
- [8] Chang, P.-C., C.-H. Liu, et al. (2009). "A neural network with a case based dynamic window for stock trading prediction." *Expert Systems with Applications* 36(3, Part 2): 6889-6898.
- [9] Chih-Fong, T. and H. Yu-Chieh (2010). "Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches." *Decis. Support Syst.* 50(1): 258-269.
- [10] Duda, R. O., P. E. Hart, et al. (2001). *Pattern Classification*. New York, John Wiley and Sons.
- [11] Esfahanipour, A. and W. Aghamiri (2010). "Adapted Neuro-Fuzzy Inference System on indirect approach TSK fuzzy rule base for stock market analysis." *Expert Systems with Applications* 37(7): 4742-4748.
- [12] Guresen, E., G. Kayakutlu, et al. "Using Artificial Neural Network Models in Stock Market Index Prediction." *Expert Systems with Applications* In Press, Accepted Manuscript.
- [13] Hadavandi, E., H. Shavandi, et al. (2010). "Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting." *Knowledge-Based Systems* 23(8): 800-808.
- [14] Kara, Y., M. Acar Boyacioglu, et al. (2010). "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange." *Expert Systems with Applications* In Press, Corrected Proof.
- [15] Lean Yu, Huanhuan Chen, et al. (2009). "Evolving Least Squares Support Vector Machines for Stock Market Trend Mining." *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION* VOL. 13(NO. 1).
- [16] Lee, M.-C. (2009). "Using support vector machine with a hybrid feature selection method to the stock trend prediction." *Expert Systems with Applications* 36(8): 10896-10904.
- [17] Liang, Q. and J. M. Mendel (2000). "Equalization of Non-linear Time-Varying Channels Using Type-2 Fuzzy Adaptive Filters." *IEEE Trans. on Fuzzy Systems* 8: 551-563.
- [18] Mabu, S., Y. Chen, et al. (2009). "Stock price prediction using neural networks with RasID-GA." *IEEJ Transactions on Electrical and Electronic Engineering* 4(3): 392-403.
- [19] Md. Rafiul, H. (2009). "A combination of hidden Markov model and fuzzy model for stock market forecasting." *Neurocomput.* 72(16-18): 3439-3446.
- [20] Meysam, A., G. Mohsen, et al. (2011). "Design and analysis of experiments in ANFIS modeling for stock price prediction." *International Journal of Industrial Engineering Computations* 2: 409-418.
- [21] Ni, L.-P., Z.-W. Ni, et al. (2011). "Stock trend prediction based on fractal feature selection and support vector machine." *Expert Systems with Applications* 38(5): 5569-5576.
- [22] Shen, W., X. Guo, et al. (2010). "Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm." *Knowledge-Based Systems* In Press, Corrected Proof.
- [23] Tutu, H., E. M. Cukrowska, et al. (2005). "Application of artificial neural networks for classification of uranium distribution in the Central Rand goldfield, South Africa." *Environmental Modeling and Assessment* 10: 143-152.