Neural Networks to Predict Stock Market Price

Qeethara K. Al-Shayea, Member, IAENG

Abstract—The prediction of a stock market price has been influenced by a set of the highly nonlinear financial and non-financial indicators may serve as a warning system for investors. In this research, the predicting of the future close price of Dow Jones Index Stocks was conducted using artificial neural networks. Feed forward neural network was used to predict next day closing in Dow Jones stock market. Nonlinear Autoregressive Exogenous (NARX) model is implemented by using feed forward neural network. To optimize the stock market price prediction, the performance of NARX model was examined and compared with different training algorithms. Although these algorithms had adequate results in predicting the NARX model with training, validation and testing. The Bayesian regularization has had the best performance in testing compared with levenberg-marquardt algorithm and the scaled conjugate gradient algorithm. While, the levenberg-marquardt algorithm has had less epochs in the best training performance of the network than the other algorithms. The performance of this model is found as a dominant model for stock market prediction.

Index Terms—Feed Forward Neural Network, Neural Networks, Nonlinear Autoregressive Exogenous, Stock Market Price.

I. INTRODUCTION

STOCK market price prediction is a very complex and challenging research area where different methods have been developed to predict stock price movement in the market [1]. Recently, artificial neural networks have been used as an auxiliary tool to predict stock price time series [2]. The novelties of neural networks lie in their ability to model nonlinear relations without a priori assumptions [3]. Neural networks have been useful to predict volatile financial variables, which are quite difficult to guess with classical statistical and econometric methods such as exchange rate and stock market [4]. Feed-forward neural networks are widely and successfully used models for classification, forecasting and problem solving [5]. Artificial neural network is a field of research which has enjoyed a rapid expansion in the literature for their ability to model nonlinear and multivariate processes without priori assumptions [3]. However, little attention has been paid to predict stock market price using neural-based nonlinear autoregressive exogenous model. Further, most research has been used stock pattern recognition model by matching template with many fixed weights assigned by researchers [6]. Thus, the aim of this research is to predict the future close price of the stock using promising classes of autoregressive network model. Dow Jones Index data set is used in this study because it represents 30 of the largest companies in the United States. The importance of Dow Jones comes from that the fluctuations of Dow Jones have some anticipating effect on other markets like German one [7]. Dow Jones Index stock is used as a data set to evaluate the performance of the proposed neural network. We use three algorithms (Levenberg-Marquardt algorithm, Bayesian Regularization, and the scaled conjugate gradient algorithm) with NARX model that allows predicting the future close price of the stock. From the results, it seems that using these algorithms are improved the accuracy of the model after each training stage. We show that the application of the NARX model enhances the performance of the proposed neural network methodology and improves the ability of it to predict the stock market price.

The rest of the paper is structured as follows. In section 2, we present the literature review and the architecture for the neural networks and NARX model. Section 3 describes the data and methodology used in this paper and the results of the empirical analysis. Finally, section 4 concludes the paper.

II. LITERATURE REVIEW

A. Artificial Neural Networks

In the last decades, many researchers have attempted to develop and enhance artificial neural networks methods to support the prediction in stock market segments. Cavalcante et. al. [8] presented a review of the recent computational intelligence approaches designed to solve financial market problems. The paper was focused on a specific financial market application or on machine learning algorithms fields. Mizuno et. al. [9] suggested a neural network model for technical analysis of stock market, and its application to a buying and selling timing prediction system for stock index. In their paper, a learning method is proposed for improving prediction accuracy of other categories, controlling the numbers of learning samples by using information about the importance of each category. Diaconescu [10] proved that an architectural approach of RNN with embedded memory, “Nonlinear Autoregressive model process with exogenous input” (NARX), showing promising qualities for dynamic system applications. He noticed that the performances of the NARX model are verified for several types of chaotic or fractal time series applied as input for neural network, in relation with the number of neurons, the training algorithms and the dimensions of his embedded memory. Wang [11] adopted artificial neural network (ANN) and two varieties of time series neural network to forecast the stock index of Chinese market. He mentioned that the NARX will be improved when more external data imported. A combination

Qeethara K. Al-Shayea is with Department of Management Information Systems – Al Zaytoonah University of Jordan. Amman-Jordan, P. O Box 130 Amman 11733 (e-mail: dqeethara@zuj.edu.jo).

ISBN: 978-988-14047-5-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)
of data pre-processing methods involves genetic algorithm and Levenberg–Marquardt algorithm for learning feed forward neural networks for stock market prediction was proposed by Asadi et. al. [12]. Laboissierea, Fernandesb and Lage [13] was proposed an artificial neural network to forecasts the range of closing prices of three Brazilian power distribution companies’ stocks. Araújo [14] and Araújo [15] submitted two studies for Brazilian stock market prediction. Oztekin, Kizilaslan, Freund, and Iseri [16] was developed a methodology involving three models to predict daily stock price movements. These models are adaptive neuro-fuzzy inference systems, artificial neural networks, and support vector machines. Zahedi and Roumaghetti [17] assessed stock market prediction using artificial neural networks in Tehran Stock Exchange. Ou and Wang [18] applied ten data mining techniques to predict price movement of Hang Seng index of Hong Kong stock market. The approaches include Linear discriminant analysis (LDA), Quadratic discriminant analysis (QDA), K-nearest neighbor classification, Naïve Bayes based on kernel estimation, Logit model, Tree based classification, neural network, Bayesian classification with Gaussian process, Support vector machine (SVM) and Least squares support vector machine (LS-SVM). Results showed that the SVM and LS-SVM generated superior predictive performance than the other models. Patel et. al. [19] presented four prediction models to predict the direction of movement of stock and stock price index for Indian stock markets. These models are Artificial Neural Network (ANN), Support Vector Machine (SVM), random forest and naïve-Bayes with two approaches for input to these models. Experimental results with continuous-valued data shows that naïve-Bayes (Gaussian process) model exhibits least performance while random forest exhibits highest performance. Oliveira, Nobre, and Zárate [20] proposed a neural network model for predictions of stocks closing prices future behaviour negotiated in BM&FBOVESPA in the short term, using the economic and financial theory, combining technical analysis, fundamental analysis and analysis of time series, to predict price behaviour, addressing the percentage of correct predictions of price series direction (POCID or Prediction of Change in Direction). As a case study, they used Petrobras stock PETR4, traded in BM&FBOVESPA. Experimental results show that he best performance was achieved with a window size of 3, which the POCID index of correct direction predictions was 93.62% for the test set and 87.50% for a validation set. Göçkena et. al. [21] presented a hybrid model based on a heuristic optimization methodology and artificial neural network to improve stock market forecasting performance in terms of statistical and financial terms. They applied the new model to predict stock price on Turkish stock market. They found that this model is a dominant model for stock market forecasting. Chiang et. al. [22] proposed an adaptive intelligent stock trading decision support system that utilizes particle swarm optimization and an artificial neural network to predict a stock index’s future movement direction. Hafezia, Shahrabib, and Hadavandi [23] presented bat-neural network multi-agent system as an intelligent model. After test the model to predict DAX stock price, results show that the proposed model works better than other models. Bisoï & Dash [24] proposed a simple IIR filter based dynamic neural network and an innovative optimized adaptive unscented Kalman filter for forecasting stock price indices of four different Indian stocks. Four different learning strategies were used to adjust the weights of the dynamic neural information system. These strategies include gradient calculation, unscented Kalman filter (UKF), differential evolution (DE), and a hybrid technique (DEUKF) by alternately executing the DE and UKF for a few generations. Experimental results show that significant accuracy is achieved using the hybrid DEUKF algorithm than the others that include only DE, UKF, and gradient descent technique in chronological order. Qiu, Song, and Akagi [25] applied artificial neural network for the prediction of Japanese stock market. This paper presents nonlinear autoregressive exogenous models for future close price of Dow Jones stock index prediction. The proposed model used three different training algorithms which are called Levenberg-Marquardt algorithm, Bayesian regularization and Scaled conjugate gradient algorithm. According to the performance of this model, the three different training algorithms based neural network model is found as a dominant model for stock market prediction.

B. Neural Networks Architectures

An ANN is a computational model that attempts to account for the parallel nature of the human brain. An ANN is a network of highly interconnecting processing elements (neurons) operating in parallel. Biological nervous systems inspire these elements. As in nature, the connections between elements largely determine the network function. A subgroup of processing element is called a layer in the network. The first layer is the input layer and the last layer is the output layer. Between the input and output layer, there may be additional layer(s) of units, called hidden layer(s). Fig. 1 represents the typical neural network. You can train a neural network to perform a function by adjusting the values of the connections (weights) between elements [5].

All the specific dynamic networks discussed so far have either been focused networks, with the dynamics only at the input layer, or feedforward networks. The nonlinear autoregressive network with exogenous inputs (NARX) is a recurrent dynamic network, with feedback connections enclosing several layers of the network. The NARX model is based on the linear ARX model, which is commonly used in time-series modeling.
The defining equation for the NARX model is

\[ y(t) = f(y(t-1), y(t-2), ..., y(t-n_y), u(t-1), u(t-2), ..., u(t-n_u)) \]  

where the next value of the dependent output signal \( y(t) \) is regressed on previous values of the output signal and previous values of an independent (exogenous) input signal. You can implement the NARX model by using a feedforward neural network to approximate the function \( f \). A diagram of the resulting network is shown below, where a two-layer feedforward network is used for the approximation. This implementation also allows for a vector ARX model, where the input and output can be multidimensional [26]. Fig. 2 illustrates the NARX model.

![NARX Model](image)

**Fig. 2. The NARX Model.**

There are many applications for the NARX network. It can be used as a predictor, to predict the next value of the input signal. It can also be used for nonlinear filtering, in which the target output is a noise-free version of the input signal. The use of the NARX network is shown in another important application, the modeling of nonlinear dynamic systems [26].

Some important qualities about NARX networks with gradient-descending learning gradient algorithm have been reported: (1) learning is more effective in NARX networks than in other neural network (the gradient descent is better in NARX) and (2) these networks converge much faster and generalize better than other networks [10] [27] [28].

Fig. 3 illustrate the proposed NARX model presented in this paper to predict the next day close price in Dow Jones Index stock.

![Proposed NARX Model](image)

**Fig. 3. The proposed NARX.**

Marquardt [29] indicated that the Levenberg-Marquardt (LM) algorithm is least-squares estimation of nonlinear parameters. It has been used to upgrade artificial neural network weights [17]. The Levenberg-Marquardt algorithm was designed to approach second-order training speed without having to compute the Hessian matrix. When the performance function has the form of a sum of squares (as is typical in training feedforward networks), then the Hessian matrix can be approximated as:

\[ H = J^T J \]

and the gradient can be computed as:

\[ g = J^T e \]

where \( J \) is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and \( e \) is a vector of network errors. The Jacobian matrix can be computed through a standard backpropagation technique that is much less complex than computing the Hessian matrix.

The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

\[ X_{k+1} = X_k - [J^T J + \mu I]^{-1} J^T e \]

When the scalar \( \mu \) is zero, this is just Newton's method, using the approximate Hessian matrix. When \( \mu \) is large, this becomes gradient descent with a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift toward Newton's method as quickly as possible. Thus, \( \mu \) is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function is always reduced at each iteration of the algorithm.

This algorithm appears to be the fastest method for training moderate-sized feedforward neural networks (up to several hundred weights). It also has an efficient implementation in MATLAB software, because the solution of the matrix equation is a built-in function, so its attributes become even more pronounced in a MATLAB environment [26].

Hagan and Menhaj [30] found after tested that the Marquardt algorithm on several function approximation problems is much more efficient than conjugate gradient algorithm and a variable learning rate algorithm when the network contains no more than a few hundred weights. Wilamowski [31] and Hagan & Menhaj [30] revealed that the LM optimization technique is more powerful than the conventional gradient descent techniques.

Bayesian Regularization is a training algorithm that updates the weight and bias values according to Levenberg-Marquardt optimization. It minimizes a combination of squared errors and weights, and then determines the correct combination to produce a network that generalizes well [26].
The scaled conjugate gradient algorithm is a supervised algorithm with feed forward backpropagation neural network. The scaled conjugate gradient algorithm (SCG), developed by Moller [32] was designed to avoid the time-consuming line search. This algorithm combines the model-trust region approach (used in the Levenberg-Marquardt algorithm described above in this study), with the conjugate gradient approach.

III. METHODOLOGY

A. Data Selection and Normalization

The data are obtained from the Yahoo/finance website. It is daily data for Dow Jones Index stock. The parameters used as an indicator are the price of the stock at the beginning of the day (open), the highest price of the stock during the day (high), the lowest price of the stock during the day (low), the price of the stock at the end of the day (close) and the number of shares of stock that traded hands in the same day (volume). The closing price of the stock in the following day is used as output. At the same time, another financial indicator is used in this study, but it is found that there is no change in predicted next day close value. The dataset covers the time from January 2010 up to September 2016. The dataset is 1679 records. 1000 records are used in training the network while 679 records are used in testing the network.

B. Results and Performance Evaluation

A two-layer feed forward neural network with sigmoid hidden neurons and linear output neurons was created using dynamic time series neural network toolbox from MATLAB R2015a as shown in Fig 1. The network was trained and simulated with three algorithms.

The first algorithm implemented is Levenberg-Marquardt algorithm. The dataset used to train the network is 1000 sample while 679 sample is used to test the network. Fig. 4 shows the proposed network performance of the L-M algorithm. The best validation performance is equal to 0.0010071 at epoch 1000 and the training is stopped after 1000 epochs. Fig. 5 shows the training state value.

Fig. 6 shows regression plots display the network outputs with respect to targets for training, validation, and test sets. For a perfect fit, the data should fall along a 45-degree line, where the network outputs are equal to the targets. For this problem, the fit is reasonably good for all data sets, with R values in each case of 1. Fig. 7 illustrates the time series response which is displayed the inputs, targets, and errors versus time. It also indicates which time points were selected for training, testing and validation.

In Bayesian Regularization algorithm, the proposed network performance is shown in Fig. 8 while Fig. 9 shows the training state value. The best validation performance is equal to 0.44312 at epoch 1000 and the training is stopped after 1000 epochs. The dataset used to train the network is 1000 sample while 679 sample is used to test the network.

Fig. 10 shows regression plots display the network outputs with respect to targets for training, validation, and test sets. For a perfect fit, the data should fall along a 45-degree line, where the network outputs are equal to the targets. For this problem, the fit is reasonably good for all data sets, with R values in each case of 1. The time series response is illustrated in Fig. 11 which is displayed the inputs, targets and errors versus time. It also indicates which time points were selected for training, testing and validation.

With SCG algorithm, the proposed network performance is shown in Fig. 12 while Fig. 13 shows the training state value. The best validation performance is equal to 4719.0433 at epoch 23 and the training is stopped after 29 epochs. The dataset used to train the network is 1000 sample while 679 sample is used to test the network.

Fig. 14 shows regression plots display the network outputs with respect to targets for training, validation, and test sets. For a perfect fit, the data should fall along a 45-degree line, where the network outputs are equal to the targets. For this problem, the fit is reasonably good for all data sets, with R values in each case is approximately of 0.99. The time series response is illustrated in Fig. 15 which is displayed the inputs, targets and errors versus time. It also indicates which time points were selected for training, testing and validation.
Fig. 6. The regression plots.

Fig. 7. The time series response.

Fig. 8. The proposed network performance for Bayesian regularization.

Fig. 9. The training state value.

Fig. 10. The regression plots.

Fig. 11. The time series response.
Table 1 shows the results after training the network while table 2 shows the results after simulating the network with the three algorithms we used it in the proposed model.

**Table 1. Results after training the network**

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Phase</th>
<th>MSE (Mean Square Error)</th>
<th>R (Regression)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-M ALGORITHM</td>
<td>Training</td>
<td>8.45569e-5</td>
<td>9.99999e-1</td>
</tr>
<tr>
<td></td>
<td>Validation</td>
<td>1.41385e-4</td>
<td>9.99999e-1</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>2.94345e-4</td>
<td>9.99999e-1</td>
</tr>
<tr>
<td>Bayesian regularization algorithm</td>
<td>Training</td>
<td>4.43115e-1</td>
<td>9.99999e-1</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>4.56249e-1</td>
<td>9.99999e-1</td>
</tr>
<tr>
<td>SCG ALGORITHM</td>
<td>Training</td>
<td>6649.78824e-0</td>
<td>9.99999e-1</td>
</tr>
<tr>
<td></td>
<td>Validation</td>
<td>4719.04325e-0</td>
<td>9.999940e-1</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>6052.19057e-0</td>
<td>9.98571e-1</td>
</tr>
</tbody>
</table>

**Table 2. Results after testing the network**

<table>
<thead>
<tr>
<th></th>
<th>MSE (Mean Square Error)</th>
<th>R (Regression)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-M algorithm</td>
<td>71.99126e-0</td>
<td>9.999993e-1</td>
</tr>
<tr>
<td>Bayesian</td>
<td>249.86358e-0</td>
<td>9.99997e-1</td>
</tr>
<tr>
<td>regularization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>algorithm</td>
<td>1049125.49139e-0</td>
<td>9.37816e-1</td>
</tr>
</tbody>
</table>

**IV. CONCLUSION**

The aim of this research is to predict the future close price of Dow Jones index stocks using nonlinear autoregressive network with exogenous input model. The proposed model has been examined and compared with three different algorithms; namely Levenberg-Marquardt algorithm, Bayesian regularization and Scaled conjugate gradient algorithm. The proposed model is a combination of neural network and NARX technique. The capabilities of the proposed neural network model are tested by...
predicting stock price indices of the Dow Jones stock between January 2010 to September 2016. The advantage of NARX model is fast training and learning abilities and a high degree of prediction results. The results show that the proposed neural network model is coping with the fluctuation of the Dow Jones stock market prices and yields good predictive performance.

REFERENCES


Assoc. Prof. Dr. Qeethara Kadhim Abdulrahman Al-Shayea has received Ph. D. in Computer Science, Computer Science Department, University of Technology, Iraq, 2005. She received her M.Sc. degree in Computer Science, Computer Science Department from University of Technology, Iraq, 2000. She has received her High Diploma degree in information Security from Computer Science Department, University of Technology, Iraq, 1997. She has received B. Sc. Degree in Computer Science Department from University of Technology, Iraq, 1992. She joined in September (2001-2006), Computer Science Department, University of Technology, Iraq as Assistant Professor. She joined in September 2006, Department of Management Information Systems Faculty of Economics & Administrative Sciences Al-Zaytoonah University of Jordan as Associate Professor. She is the head of the Management Information Systems Department in Al Zaytoonah University of Jordan. She is interested in Artificial intelligent, image processing, computer vision, coding theory and information security.