Journal of Novel Applied Sciences
Available online at www.jnasci.org
©2014 JNAS Journal-2014-3-S2/1602-1611
ISSN 2322-5149 ©2014 JNAS

Predicting the Daily Efficiency of Tehran Stock Share Price by Using of Artificial Neural Networks, Cascade Forward

Roohollah Bahmani Zargari[1*], Mansur Amini Lari[2] and Mohsen Bashiri[3]

1- MA of Computer Software Engineering of Azad University, Broujerd Branch, Iran
2- Assistant Professor of Research and Science University, Fars Branch, Iran
3- Assistant Professor of Sanati Sharif University, Tehran Branch, Iran

***Corresponding author*:** Roohollah Bahmani Zargari

ABSTRACT: The ever increasing development of tendency to be involved in regional and universal financial markets, such as stock exchange which is mainly online today, and also the high speed of trade and fluctuations of this market have made trade strategies to be of a significant importance. Generally, by predicting the situation of market in the future and invested can determine its position, situation and the rate of his investments, in a way that the efficiency of his/her assets reaches to its maximum. In this research, we have tried to offer the best prediction of the movement trend in the next day share price by using technical analysis (stock exchange diagram analysis) and cascade forward, learning by three layer supervised learning, artificial neural networks and Levenberg-Marquardt learning algorithm, weight learning function and delta rule, and then, by having such information, conduct a successful and profitable trade in the stock exchange.

***Keywords*:** prediction of share price, technical analysis, cascade forward neural network.

INTRODUCTION

The main issue regarding a safe investment is having a correct analysis of markets future. The art of capital management is the art of finding and obtaining the highest profit from the market. Predicting the future trend of price in order to determine the most suitable time for selling or per chasing shares, is the main goal of investors in the stock exchange. The two thinking school which prevails the share market literature, are fundamental analysis and technical analysis. Fundamentalists believe that stock share have an inert value and market work forces make sure that the price of each share covers this value in the long term. On the other side, technicality believe that past prices of each share (and price changes); size and the volume of share trade tend to follow the same pattern. As a result, systematic analysis can bring about extra ordinary profits in the short term generally supply and request forces, determine share prices. Perhaps, technical analysis is the only technic for an investor to use form. Fundamental analysis requires all fundamental factors due to the high volume of information and time that it needs obtaining. These factors are costly and timely for small stock holders. On the other side, these small stock holders (investors) have no access to secret information, a low competitive margin is provided for fundamental analysis. Economists believe that technical analysis has assisted market behavior studies through predicting prices by using computer. Both of technical and fundamental analysis methods have always been trying to solve a unit problem that is predicting the movement direction of prices. however, with respect to the above points, if a trade man is having to select one of these two methods, logic says that in Iran stock exchange market, success is more probable by technical analysis, because technical analysis investigates the effect of fundamental factors and as fundamental effect on the price is always considered, so fundamental analysis is definitely inessential. there is such possibility that selling and per chasing in a financial market are done by technical knowledge but it is almost un probable that a person conduct a success full trade (buying and selling) on the basis of fundamental issues without technical evaluation of the market. Technical analysis is the process of historical evaluation of share prices to determine future probable prices. This is done by a comparison of the expected trend of current price with its historical movement in order to predict a logical result. As a principle of technical analysis, we must remember that most of financial market events stem form past incidences of market in every movement or decision of the market, there is a huge amount of past data and these upward or downward movements produce important information for the next movements.

2. A Review of Past Researches

Technical analysis emerged with Charles down and William Hamilton's perspectives and articles (1900-1902). A review of literature on this topic manifests those numerous researchers who worked on this field. Before the systematic constitution of this method of analysis, Brown and Jennings (1989) showed the value of technical analysis associated whit signals and prices.

In his research, Sweeny (1988) concluded that depending on the level of transactions expenses, filtering rules (filtering rules and mobile mean rules are the two transactional rules of technical analysis), lead to a more or less profitable result.

Lakonishok and lebaron (1992) used mobile mean rules and concluded that these rules will also make profitable results.

Lerich and Thomas (1993) and kho (1996) also investigated the mobile mean strategy and concluded that the mentioned strategies are helpful.

Ratner and lill (1993), in some of Asian and Latin American countries, reached to this result that the use of technical analysis methods will lead to profitable results. Menig, Matno and Goro (2010) concluded that mobile mean rules are of more prediction ability and they are able to get more efficiency. Mang, Manzure and Chew (2012) proved the better function of mobile mean method and partial power index comparing with selling and maintenance method in Singapore stock exchange. In his research, amiri (1995),concluded that technical analysis method can be implemented in Tehran stock exchange market to analyses the shares . Khanloo (1996) reached to this result that various methods of technical analysis which were used in the world financial markets, are also applicable to some extent in Iran stock market.

The results of Gholamzade and Norush's research (2000) showed that the process of making annual profit by those Iranian firms which were under investigation was mobile mean method. In predicting shares profitability, Mehrani and Karami (2008) used historical information (both financial and non-financial) to distinguish successful firms from unsuccessful ones. In their research, Sadeghisharif and Sultan Zareii (2011) concluded that technical analysis methods are helpful and profitable for analyzers and investors of Tehran stock exchange market.

MATERIALS AND METHODS

3.1. Designing the Suggested Neural Network

To make a neural network, a designer or manufacture must identify the structure of a network which has a more efficient output regarding the issue of prediction apart from selecting a set of input variables. Changing the architecture of a network, even without changing input and output variables, can disturb the previous predictions totally. In order to find the best architecture, the network manufacturer must proceed by test-error method to reach the most favorable result. In this section, each component of the neural network used in this article, is described.

3.1.1. Selection of Input Variables

In this research, after studding past researches and trend rule, those variables which had a significant role in determining the price, were identified and the model was designed by varying these parameters. In the stock market, lots of data are declared daily as shares characters. The parameters which we consider as input variables in this model, includes five important data in time period of one day: 1. the open price of share; 2. the close price of price of share; 3. the low share price; 4. the high share price; 5. close gold price. Since, for modeling with artificial neural networks, test-error data are required, the information of a 2 year period (2011-2013) about total index of Tehran stock exchange market was considered suitable for this research. This information which was extracted daily and online from formal site of Tehran stock exchange market (www.irbourse.com), constituted the basic data base of this research.

3.1.2. Data Classification

The aim of classification is to know how the situation of the considered share for selling or per chasing, will be in future. In this research, using trend rule (up trend, down trend, sideways trend) and circumstances of previous days occurred in the market, were considered as a standard for data classification. By referring to past data and evaluating the market atmosphere and above cited cases, an investor decides to calibrate share and per chase it in the incoming days. Thus, generally, we consider the standard of data classification in a way that if in the previous day, the close daily price of each share per candlestick is higher than the close price of each share, it means that we have an uptrend and it is the best time for purchasing the probable share (class 1),and if the contrary of this situation occurs, that is in the previous day, the close daily price of each share per candlestick is higher than the close price of each share, we have a down trend and this is the best time to sell the probable share (class 2), and if none of these situations happened, the market is in a steady state or “sideways trend”(class 3). According to the following formula (1), we did data classification:

```
Begin
If (Close (0) >Open (0))      Then
                Buy
      Else if (Close (0) < Open (0))   Then
                Sell
Else
           No Buying and Selling
End
```

Formula 1: data classification

In order to classify data, we have divided them in to 3 classes based on their characteristics (table 1).

Table 1. Data Classification

Class characteristics	Class number
The best time for buying share	1
No buying and selling	2
The best time for selling share	3

Up to this section, a set of 5 variables as the network input and a classification of 1to3 are determined for the output.

3.1.2.1. Class Encoding

In order to prevent from saturating activation functions or “death of some neurons”, we encode the considered output (which has 3 classes) in a binary way. That is, for each class the related output is one and the rest are zero (table 2).

Table 3. Binary Class Encoding

	Output 3	Output 2	Output 1
Class 1	0	0	1
Class 2	0	1	0
Class 3	1	0	0

In fact, we insert the binary equivalent instead of the number of the class. Binary encoding is ideal as above and in practice, instead of zero, a near value like 0.1 and instead of one, a near value like 0.9 is used (table 3). Thus, we will have:

Table 3. Class Encoding

	Output 3	Output 2	Output 1
Class 1	0.1	0.1	0.9
Class 2	0.1	0.9	0.1
Class 3	0.9	0.1	0.1

3.1.3. Data Structure

The first step to train (teach) a neural network is to offer a set of paradigms with which the network can learn. To do so, 661 days of stock exchange data were used (table 4). The implemented data included 661 pairs, classes and characteristics as follows:

Table 4. Characteristics and Structure of Network Data

The number of train data	530
The number of test data	131
The number of characteristics	5
The number of classes	3

The goal of this section is to train the neural network by using training data in a way that it can distinguish the order of stock exchange classification and classify test data in a successful manner. In sum, the range of data changes and input – output information can be observed in table (5).

Table 5. The Input - Output Of Neural Network and the Range of Their Changes

Network input	Changes range
Open price of share	14503 – 88015.5
High price of share	14529.4 - 88199
Low price of share	145009 – 87015.5
Close price of share	14500.9 – 88114.7
Close gold price	1237.72 – 1201.9
Network output	Changes range
Movement trend	1 - 3

3.1.3.1. Data Normalization

In order to obtain a reasonable and ideal respond from the model, it is necessary that before the initiation of network training, inputs and particularly output be limited to a certain extent by using statistical methods. The aim of this correction is to reduce network modeling error. This operation is called data standardization or normalization. To do so, the following relation was used for data normalization (Haykin, 1999).

$$A_I = 0.1 + 0.8\,(\frac{A_i - A_{min}}{A_{max} - A_{min}})$$

In which:

$A_{I:}$ is the normalized value of data, $A_{i:}$ is the target value of data, $A_{Min:}$ the minimum data related to the considered parameter, $A_{Max:}$ the maximum data related to the considered parameters. By using the above relation, the data is normalized between (0.1 – 0.9) and they are prepared to enter the modeling software.

3.1.4. Determining the Number of Layers and Network Neurons

Various models and architectures were examined for determining the suitable topology of neural network. By changing the number of hidden layers, the main prediction model was selected. Finally, the optimum number of layers was calculated as four layers (one input layer, two hidden layers, and one output layer) and the number of neurons was considered as (5-6-1) which we will explain them in a summarized way.

3.1.4.1. Input Layer

Selecting the number of input of input parameters is of significant importance, because every input pattern includes much important information about network's structure and complexity and the interrelation of data. Most of research use trial- error method to calculate the number of input nodes. In this research, the number of input nodes is considered precisely equivalent to the size of network input which is five nodes.

3.1.4.2. Hidden Layer

Hidden layers and nodes play a significant role in the operation or function of neural network. The hidden nodes in the hidden layer allow the network to identify data Characteristics and by which establish the complex nonlinear transcription between input and output variables. In the theory of neural network, the optimal precise can be obtained for approximating functions by using sufficient number of hidden nodes in the hidden layers. In this study, by using trial-error method it was concluded that (according to tables 6) that if we consider the number of hidden layers as two layers with 6 nodes, the network will bring about the best result.

Table 6. The Function of the Number of Neurons in the Network's Hidden Layer

The number of neurons in the hidden layer	Root Mean Square Error(RMSE)
2	0.6253
3	0.5987
4	0.4251
5	0.5463
6	0.1261
7	0.2939
8	0.1865
9	0.1810
10	0.4533

3.1.4.3. Output Layer

Compared to the output parameters which we defined for the network, the considered network has one output per minute. Up to this part, we can observe an overall representation of the structure of optimized neural network to predict the movement trend of the share under investigation (figure 1).

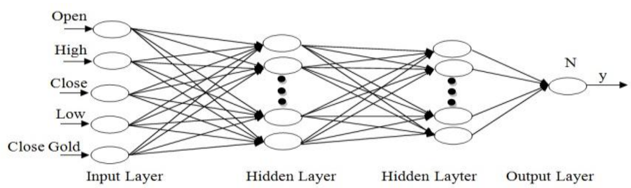

Figure 1. An Overall Representation of the Structure of optimized Neural Network

3.1.5. Selecting the Type of Neural Network

After developing the data base related to training, and defining the suitable number of layers and neurons, now it is time to select the type of network. To do so, in this research, we used the cascade forward network with multi-layered supervisor for the first time (multi layered networks are far more efficient and stronger than unit layered networks to solve problems).

3.1.5.1. The Structure of Nodes in the Cascade forward Network

Generally, the structure of nodes in the suggested network is injected in a way that nodes are placed in alternative layers and their relation is unilateral. When an input pattern is imposed on the network, the first layer calculates its output value and gives it to the next layer. The next layer receives these values and the nodes of previous layers as the input and calculates them and transfer its output to the next layer, in this way, signal transfer is performed up to the last layer (figure 2).thus, the conditions of network structure can be obtained:

1. The neurons of each layer are only attached to the neurons of the next layer.
2. Each neuron is totally attached to the next layer.
3. The operation is dispersed forwardly. All of neurons except input layer are commutative and each neuron can have an independent activation function.
4. Each neuron can have an independent bias.

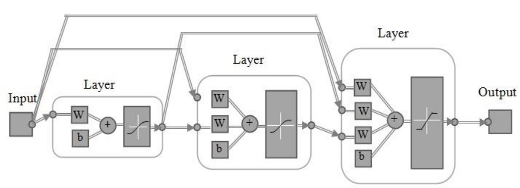

Figure 2. The Structure of Nodes in the Cascade Forward Network

3.1.6. Network Training Function (Algorithm)

Learning rules of neural network are used to train the network a special task. To do so, a procedure which corrects network biases and weighs, is implemented, this is called "training algorithm or function". It is suggested that for forward neural networks, researchers use several error feedback algorithms such as Levenberg-Marquardt or

SGFB. According to the implemented data in this research, the Levenberg–Marquardt algorithm has the best results of all algorithms and this is used in all models which we will explain them in sum.

3.1.7. Network Learning Algorithm

To achieve the ideal response by changing bias and weight variables, "the learning rule" In neural network is used. In this study, we used the "delta rule".

3.1.7.1. Delta Rule

When the problems are not separable linearly, a set of methods are used to solve these problems, one of these methods, is "delta rule". The basic and fundamental theory of this rule is the use of gradient descent algorithm to search in the atmosphere (space) of "possible weights" hypothesis. This is a basic method for all kinds of learning algorithms which searches in to all of continue hypothesis spaces. At each step, weights and biases are changed based on the following delta rule (the negative sign shows that the movement is toward gradient reduction):

$wi = wi + \Delta wi$
Where $\Delta wi = - \eta E'(W)/wi$

In the following, the learning algorithm by using delta rule is summarized.
1. First, it dedicates a stochastic value to the weights.
2. It continues the following steps until it reaches to stop condition (in this study, when $t \sim o$).
It changes the W_i value to make the error very small:
$W_i = W_i + \Delta W_i$
It changes the weight ΔW_i as follows:
$\Delta W_i = \Delta W_i + \eta (t – o) x_i$
It changes the value of b_i as follows:
$b_i = b_i + \Delta b_i$
It changes Δb_i as follows:
$\Delta b_i = \eta (t – o) x_i$
Where;
t: target output
o: output generated
η: constant called the learning rate (e.g., 0.1)
Now, we illustrate the summary of neural network specifications and parameters (in this research) in the table (7).

Table 7. A Summary of Parameters and Specifications of Neural Network

parameters	explanation
The structure of neural network	Cascade forward
Type of neural network	Feed forward
Learning (training) algorithm	Levenberg – marquardt algorithm
Bias and weight learning function	Delta ruler (stochastic gradient)
Error function	Mean square error (MSE) and root mean square error (RMSE)
Number of hidden layers	2
Number of nodes in the hidden layer	6
Validation check	6
Hidden layer transfer function	tan sigmoid
Output layer transfer function	Symmetric saturating function
Number of training data	530
Number of examination data	131

3.1.8. The body of Suggested Neural Network

Generally, if we want to describe the body of neural network designed in this paper, we can show it in the following figure (3).

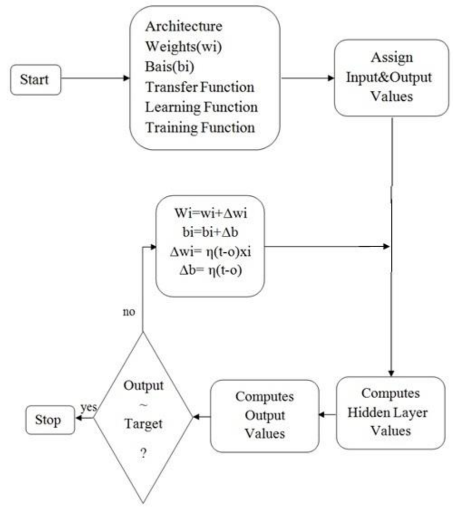

Figure 3. The body of suggested neural network

4. Evaluating the Strength of Neural Network

In every neural network, the aim of training the network is to determine and change neurons and biases weights coefficients in a way that the expected output (the target output) of the network is near to the current output (the output resulted from imposing inputs on the network and receiving its output) as much as possible. In other words, the error between output value and the network output value would be minimum.

4.1. Investigating the Efficiency of Neural Network in Data Classification

For doing so, we formulated a square matrix at the size of class numbers, which it's "ij" array shows that how many of the "i" class data is placed in the "j" class. The best state occurs when the matrix is diagonal, since it shows that no data of the "i" class is located in the "j" class and all data are placed in their own class. The more diagonal this matrix, the more efficient the network is as the number of our classes were 3, the output is generated as follows:

$$AM = \begin{pmatrix} 42 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 88 \end{pmatrix}$$

As it is clear, all of data are classified in their own classes. Finally, the efficien work is demonstrated by a number called correct classification rate. This number is equal to the total proportion of correct classification to the total number of classifications. This is calculated according to the following relation:

CCR = Sum (Diag (Am))/Sum (Sum (Am))*100

As it is guessed from the above matrix, the value of "CCR" is equal to 100, because this matrix is totally diagonal. That is, our trained neural network has been able to classify the test data (that has never seen before) correctly and perfectly.

4.2. Investigating the Errors of Test and Training Data

The thing which is important for a good training is having low errors for prediction data (in training and test) compared to target data. To investigate the precision rate of error in data, the following relation can be used:

Error Train=Normal(Y Train Sim-Y Train) / Train Num

Where:

"Error Train" is the training error rate, "Y Train Sim" is the output resulted from imposing inputs on the network, "Y Train" is the target output and "Train Num" is the number of training data.

For the above training, the response of the error is as follows it is an ideal value.

Error Train=0.0032

Of course, a thing which is more important for a sound training is having low errors for test data (those which the network has not seen before). For this standard of test data error, the following relation is also used:

Error Test =Normal(Y Test Sim-Y Test) /Test Num

Where:

“Error Test” is the rate of test error, “Y Test Sim” is the output resulted from imposing inputs on the network, “Y Test” is the target output and “Test Num” is the number of test data.

For the above training, the response of this error is ideal and it is as follows:

Error Test =0.0420

As it is observed in the figure (4), the errors resulted from prediction, are very small and negligible. It shows the high strength of the designed neural in prediction.

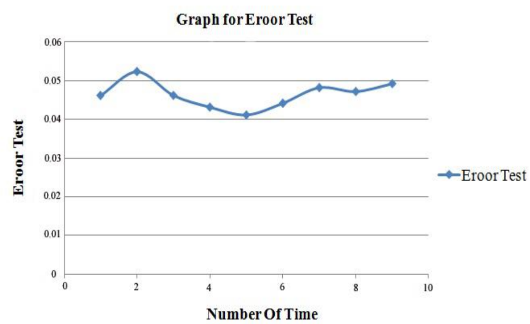

Figure 4. The graphical diagram of test data errors

4.3. Investigating the Structure of Neural Network

In the table (8), the structure of the best neural network and the mean square error function, correlation coefficient (R), the precision of network for training data and test of each network are offered. According to this table, the architecture of “5-3-3-1” (with 5 input parameters, 2 hidden layers which have 3 nodes in each layer) gave the best result. It should be noted that for every structure of the network, the process of network training is performed several times and the values of mean square error function and the correlation coefficient were obtained by “averaging” method.

Table 8. The structure of the best developed neural networks

Network Structure	Mean Square Error (MSE)	Correlation Coefficient (R)	Network Precision For Training Data	Network Precision For Test Data
5-1-1-1	0.0349	0.9651	0.0081	0.1159
5-2-1-1	0.0356	0.9644	0.0082	0.0757
5-1-2-1	0.0341	0.9659	0.0080	0.1311
5-2-2-1	0.0970	0.9030	0.0135	0.0430
5-2-3-1	0.0211	0.9789	0.0063	0.1230
5-3-2-1	0.0279	0.9721	0.0073	0.0587
5-3-3-1	0.0154	0.9840	0.0032	0.0420
5-4-3-1	0.0337	0.9663	0.0080	0.0890
5-3-4-1	0.0380	0.9620	0.0085	0.0653
5-4-4-1	0.0319	0.9681	0.0078	0.1082
5-5-4-1	0.0355	0.9645	0.0082	0.0690
5-4-5-1	0.0352	0.9648	0.0082	0.1385
5-5-5-1	0.0241	0.9759	0.0152	0.1233

In this section, the determined parameters of the neural network and the best obtained structure for this network, a comparison of network response with test data outputs, are expressed.

4.4. Investigating the Value of Real Data and Prediction

After training the network, now it is time to test the network for the new data in order to evaluate the precision rate of simulation output with the target value. According to figure (5), we can observe the graph of network output with the obtained target value.

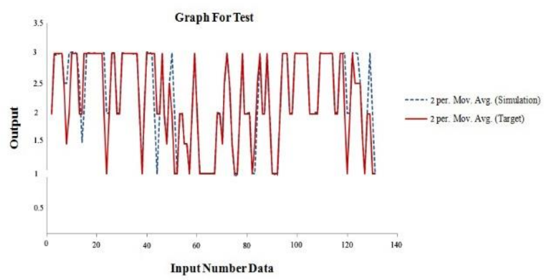

Figure 5. The graph diagram of network output with a target value of test data

4.5. Investigating the Diagram of Changing Mean Square Error

After selecting the configuration of (5-6-1) as the best design, we will discuss the diagram of changing mean square error. As it is observed from the figure (6), this diagram shows the method of training and reduction of training error in each epoch. In the 31 epoch of the network, the network training is stopped to avoid from over filling, as it is observed from this figure, the value of mean square error for training data, is approximately(10^{-3}) (in this epoch) and this value for test data and validation data is respectively (10^{-2}) and (10^{-1}). These results seem appropriate.

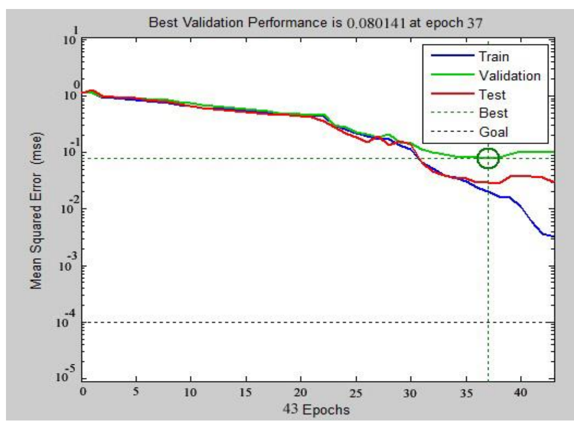

Figure 6. The diagram of changing MSE (MSE variation)

4.6. Investigating the Strength of (R^2) Standard Resulted from the Neural Network

The standard of (R^2) measures the variation rate of "Y" (which is described by regression model). This is used as an index for final investigation of regression model. The value of (R^2) is always between 1 and 0. The closer this values to one, the more ideal the final investigation (for regression model). As it is shown in the figure (7), in the time of operation, the value of "R" for training data is "R=0.99838", for test data "R=0.98529", for validation data (the data that the network has not seen before) "R=0.94721" and eventually for all of network data, we have "R=0.9856", as seen, this value is desirable.

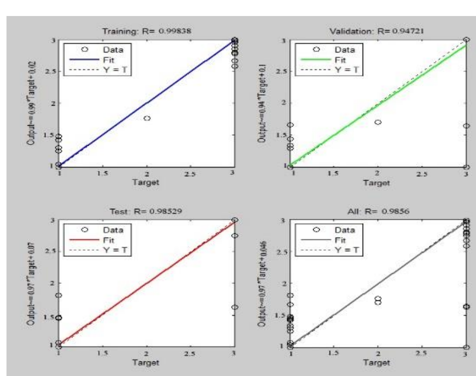

Figure 7. The diagram of investigating the strength of standard (R^2)

RESULTS AND DISCUSSION

Generally, the results of this research can be expressed as follows:

- The first finding of this study understands the complexity and procedure of price variations in Tehran stock market.
- By technical analysis, Tehran stock market can achieve an acceptable prediction.
- Implementation results showed that the trained neural network could classify stock market data with a high precision (correct classification=100).
- The time behavior of stock market (based on a daily extent) is not a stochastic process, instead, it is a non-stochastic process which can be predicted to considerably. Although this time behavior is a non-stochastic process, but it has many complications and it needs a network with many hidden neurons and layers. While increasing the number of hidden neurons and layer leads to higher precision of the network, however, this increase is limited and above that limit, the efficiency of network will be decreased.
- Among the previously designed neural networks, the neural network with the feed forward cascade architecture and five input parameters, two hidden layer and 6 hidden node (with a 5-6-1 architecture), transfer function of tan-sigmoid in hidden layers and symmetric saturating function in the output layer, Levenberg-Marquardt training algorithm and weight learning function, delta bias with the validation check of 6, epoch=37, MSE=0.0154, R=0.9840 and RMSE=0.1261, is the best network modeled for predicting the total index of stock market.

At present, the best method to predict the price in Tehran stock market by investors and those who expect a profitable transaction (trade) in a short term period, are focusing on variations of past prices in the market.

REFERENCES

Charalambous C and Kaurou F. 2010. “Comparative Analysis of Artificial Neural Network Models: Application in Bankruptcy Prediction”.

Massah M. 2005. “technical analysis (from A to Z)”, the first publication, chalesh publication institute, Tehran,Iran.

Mohammadi S. 2004. “Technical Analysis in Tehran Stock Market”, financial research, no 17, pp 97-129.

Morfi J, Farahani Fard K and LangrudiQasemian R. 2005. “technical analysis in the capital market”, the second publication, chalesh publication institute, Tehran,Iran.

Moosa Imad A. 2000. "Exchange Rate Forecasting: Techniques and Applications", MacMillan Business, London, Great Britain.

Pruden Henru O and Douglas R. 2005. "Catastrophe Theory and Technical Analysis Applied to a Call Tech Experiment on Irrational Exuberance", Managerial Finance, Vol.31, No.5, p 38.

Sadeqisharif SJ and Zareemasood S. 2011. “the benefit of technical analysis methods in Tehran market”, the magazine of accounting and auditing investigations, no. 49.

Safarnavade M. 2001. “predicting the price of shares in Tehran stock market”, a MS thesis, Imam Sadeq university.

Sehgal S and Anurag G. 2002. "Abnormal Return Using Technical Analysis: The Indian Experience", Finance India, No.16, pg81.

Subrata Kumar M. 2002. "Profiting Technical Analysis in Indian Stock Market", Finance India, No.1, pg 109-120.

Wong WK, Manzur M and Chew BK. 2012. "How Rewarding is Technical Analysis? Evidence from Singapore Stock Market", Departmental Working Paper, No. 0216, Dept. of Economics, National University of Singapore.