فصلنامه بورس اوراق بهادار شمارۀ ۱۴ تابستان ۹۰ سال چهارم ــ تاریخ دریافت: ۸۹/۰۴/۰۵ ــ تاریخ پذیرش: ۹۰/۰۷/۱۲ ــ ص ۲۳۷- ۲۶۱

Short-term Prediction of Tehran Stock Exchange Price Index (TEPIX): Using Artificial Neural Network (ANN)

Mohsen Dastgir[1]
Mohammad Hadi Enghiad[2]

Abstract

The main objective of this study is to find out whether an Artificial Neural Network (ANN) will be useful to predict stock market price, which is highly non-linear and uncertain. Specifically, this study will focus on forecasting TSE Price Index (TEPIX) as the most significant index of Iran Stock Market.

Many data have been used as inputs to the network. These data are observations of 2000 days for a period of 9 years from 02/29/2000 to 12/03/2008. Data are divided into two categories; fundamental and technical data. The fundamental data used here are principal economic values like Dollar/Rials Exchange Rate, Gold price and Oil price. The technical data used are technical indices such as Moving Average (MA), Moving Average Convergence/Divergence (MACD), Relative Strength Index (RSI), Rate of Change (ROC), Momentum (MOM) and daily trading volume of stocks.

The selected data are divided into training set and test set, in order to be entered into the network and the remaining 10% was used as the testing set. Training set consists 90% of data. This classification uses 3 different approaches to assemble the training and test data, including random, deterministic and consecutive selection.

Here, a feed-forward neural network (FFNN) with the most suitable algorithm for finance (i.e. Back Propagation algorithm) was used for the prediction. Predictions were made for the next day of TEPIX with a 3-4-1 topology and 1500 epochs. The performance of the ANN was evaluated by MSE. Finally, the results showed that ANN could properly recognize the relationships between fundamental and technical data and TEPIX, so that the prediction of the next day was quite possible.

Keywords: Short-term Prediction, Forecasting, Tehran Price Index (TEPIX), Artificial Neural Network (ANN)

JEL classification: G17

1. Professor of Accounting and Finance- Shahid Chamran University
2. PhD. Candidate.

1. Introduction

According to the complexity of stock market as a result of various daily stocks trading by different investors, this market has become uncertain and indefinite. Therefore, in this situation predicting the stock market in order to investigate these uncertain conditions is very important. Using Artificial Neural Network (ANN) in predicting stock market behavior, as a better technique compared to other statistical techniques which are limited when encountering non-linear data set, this research attempts to predict TSE price index behavior.

The most important evaluation criteria of investment performance in stock markets, is stock indices, in particular stock price index. Stock market indices which are included in principal economic variables are recognized as pioneer variables. Pioneer variables are used for tracing the landscape of national economic movement. In fact, investors can evaluate their portfolio through pursuing indices fluctuations. In general, stock market index like thermometer shows capital market conditions and economic status of the country. Declining stock market index demonstrates economic inactivity and rising of the index represents economic growth. In other words, stock market index is a useful scale for current expectations about the future of stocks and reflects the effects of political and economic phenomena with structural and long-term changes in economy (Salim, 2006).

In accordance with the nature of uncertainty dominated in Tehran Stock Exchange and the fact that most capital markets in the world do not contain definite regularity, artificial intelligence is a proper tool for the improvement of predictions in stock market because it is non-linear, dynamic and exempt from any model. So, using ANN for the prediction of stock price index is for its ability to learn non-linear relationships between inputs and outputs based upon data set that is introduced to the network as the training set to be learned and recognized by the network. Then the network generates the results to the test data set (Lucas and Karimi, 2004).

Unlike efficient market Hypothesis (EMH), many researchers and academics claim that stock markets display chaos behavior. Chaos behavior is a non-linear determined process that appears in a casual manner and is not readily understandable (Weiss, 1992). Through the capability of neural networks in learning irregular and non-linear systems, there is the possibility for ANN to identify non-linear relations better than conventional methods.

2. Literature review

The ANN has been used in several areas of finance, such as forecasting bankruptcy and business failure (Tam & Kiang, 1990; Wilson & Sharda, 1994); foreign exchange rate (Refenes & Zaidi, 1995; Yao & Tan, 2001; Nag and Mitra, 2002); stock prices (White, 1988; Yoon & Swales, 1991; Yao, & Poh, 1995; Yao, Tan, & Poh, 1999; Kohara, et al. 1997; Leigh, et al. 2002); Bond (Singleton and Surkan, 1995; Daniels and Kamp, 1999) and other values (Kaastra & Boyd 1995).

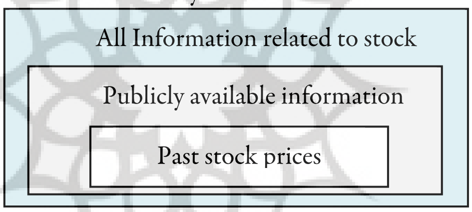
White (1988) did one of the earliest works. He investigated a three-layer feed forward network for extracting non-linear regularities from economic time series, in particular IBM daily stock returns. However his research results were poor. Later researches by Yoon and Swales (1991), Steiner and Wittkemper (1995), Yao and Poh (1995b), and others have proved that the ANN could predict stock price better than statistical techniques.

The issue of market predictability has been discussed a lot by researchers and academics. In finance, a hypothesis has been formulated known as the Efficient Market Hypothesis (EMH), which implies that there is no way to make profit by predicting the market, but so far there has been no consensus on the validity of EMH (Kalyvas, 2001). In general, it states that stock prices fully reflect available information. Every time new information arises, the market corrects itself and absorbs it. Consequently there is no space for prediction (Malkei, 1999).

The most recognized classification system divides information into three groups (Ross, Westerfield, Jaffe, & Roberts, 1999); information on past prices, publicly available information and all other information.

The picture below shows the classification:

Figure1. Classification system of information of a stock



More specifically the EMH has got three forms (Malkeil, 1999).

Other studies indicated that some macroeconomic variables have impact on stock returns and stock market indices Brown and Otsoki (1993) using the APT, investigate relationship between oil price variable and stock returns in Japanese firms and find that this variable affects the stock returns. Jones and Kaul (1996) investigate the influence of oil price on American, Canadian, British and Japanese stocks returns and found that, related to British and Japanese stock market, American and Canadian stock markets have showed less reaction. Clare and Tho mas (1994) proved that oil price can affect the stock returns in British stock market.

Foreign currency rates variable is used for reflection of foreign currency rates risky factor. Obviously this variable is an effective factor in multinational companies

(Jorin, 1990) and companies in natural resources industry (e.g. oil, petrochemical and mineral companies). Brown and Otsoki (1993) using the APT, investigate the foreign currency rates variable for Japanese firms and found that this variable is statistically significant in their model. Morelli (2002) investigates the relationship between foreign currency rates (mark-pound) and British companie's stock market and found that this variable cannot explain changes in company's stocks market. McQueen and Roley (1993) investigate the relationship between inflation rates and companies' stock return.

Some other Iranian studies by Dastgir et al. (2009), Sajadi et al. (2010) and Sadegh-vaziri & Yazdan-panah (2006) showed that macroeconomic variables such as foreign currency rates, inflation rates, oil price, gold price, growth rate and money supply have relation with stock prices and stock market indices.

3. Available Data

Information about the condition of the market is available through the data. Data could be categorized into three major groups (Hellstrom & Holmstrom, 1998):

3.1 Technical Data: price and volume data for each stock in daily trading, such as: closing price, highest price, lowest price and volume traded during the day.

3.2 Fundamental Data: data concerning the activities and financial situation of each company. This data includes the general economy, condition of the industry and condition of the company. Data included are inflation, interest rates, other stock indices prices of related commodities (oil, gold and currencies), P/E (Price/Consensus Earnings), book value per share, etc.

3.3 Derived data: data that is built from original data by transforming and combining technical and fundamental data. Some of them are: Returns which are the relative increases in price compared to the previous point in a time series, Volatility which is the variability in a stock price usually used to estimate investment risk and profit possibilities, Technical Indicators such as Moving Average (MA), Relative Strength Index (RSI), etc.

For predicting future market prices, past and present data are thoroughly studied. It is the fundamental task in all prediction methods. The things that have differentiated different prediction methods in the past are the way they were done and their data, (Iskandar, 2005).

4. Prediction Methods

There are both conventional and non-conventional methods of predicting Market. Some conventional methods are; Technical Analysis, Fundamental Analysis and

Traditional Time Series Prediction Methods. In addition, non-conventional methods are examined, such as Genetic Algorithm (GA) and Artificial Neural Network (ANN).

4.1 Non-conventional Methods

There are several non-conventional methods that have been used to predict the stock market. Some of these methods will be discussed, especially ANN which has a long precedent in market prediction.

4.1.1 Genetic Algorithm (GA)

Genetic Algorithms (GAs) is proposed as a way to perform a randomized global search in a solution space. In this space, a population of candidate solutions, each with an associated fitness value, is evaluated by a fitness function on the basis of their performance. Then, using genetic operations, the best candidates are used to evolve a new population that not only has more of the good solutions but better solutions as well.

4.1.2 Artificial Neural Network (ANN)

Another non-conventional method is an ANN. It is typically used for pattern recognition tasks. An ANN is used for non-linear regression, where the task is to find a smooth approximation between multi-dimensional points. For stock market prediction, input and output data are presented to the network simultaneously. Input can be for any type of data, while the output is usually the price index, returns to the trend.

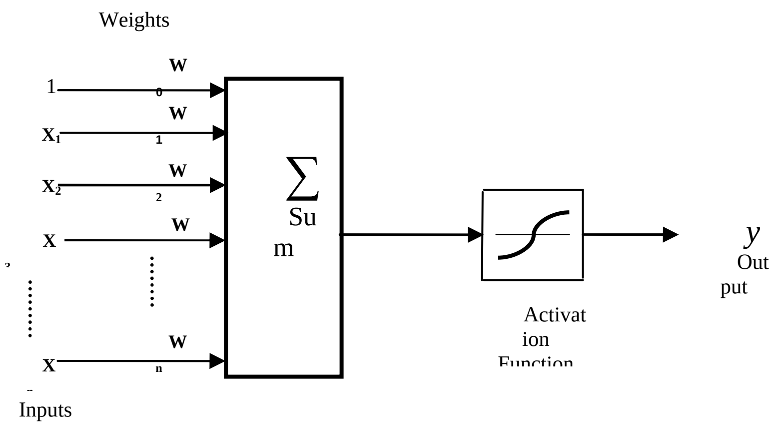
Only ANN will be used here. Further information regarding ANN is discussed thoroughly in the next section.

4.1.2.1 Single Perceptron

Zirilli (1997) proposed the idea that the basic block of an ANN is the perceptron, also known as an artificial neuron. A human neuron operates with 3 basic functions: input, processing, and output. The function of the input components, called dendrites, is to receive impulses from other neurons and to provide inputs to the soma, which is the processor component of the neuron. Next, the soma collects the impulses, sums them and compares the sum by sending signals (outputs) via the axon (Iskandar, 2005).

An artificial neuron simulates the three elements. An ANN also works by receiving inputs and then processing the inputs to generate outputs. While the soma is simulated with a summation element, the axon soma is replicated by an activation function and the dendrites by weighting the inputs. The perceptron sums N weighted inputs and passes the result through a non-linear or activation function. Overall, there are 4 basic types of activation function (Iskandar, 2005).

Figure2. Single Artificial Neuron



Zirilli (1997) proposes the mathematical model for the operation of the artificial neuron with the following equation:

$$y = f\left(\sum_{i=0}^{n} x_i w_i\right) \qquad 2$$

Where, xi are the inputs, yi is the output, wi are the weights and n is the number of inputs.

Each perceptron has the ability to learn or to recognize simple patterns. Its learning algorithm will modify the weights, until the desired output is achieved. The learning process of the perceptron starts with training the perceptron. This is done by presenting it with a set of inputs (x) and recording the output (y). The weights of the perceptron are modified based upon the error between the output and the desired output (d) to reduce the error. The Bias (x0) is set to 1, where the rest of the x provides the actual inputs for the neuron. The following describes the procedure for adapting the weights:

1- Initialize the weights and thresholds: Set wi $(0 \le i \le N)$ to small random variables.

2- Present new inputs and desired outputs: Present new continuous valued inputs xi $(0 \le i \le N)$ along with the desired output d. Note that $\chi_0(\text{bias}) = 1$.

3- Calculate the actual output: based upon equation 2.

4- Adapt weights: where η is the rate at which learning takes place $(0 \le \eta \le 1)$ and d is the desired output.

$$w_i = w_i \chi_o \eta(d - y) \quad 3$$

This procedure is repeated until the error $(d - y)$ reaches an acceptable level.

4.1.2.2 Activation Functions

Activation Function is scalar-to-scalar function that is used to transform the input of the units in the ANN (Sarle, 1997b). One of the activation functions that is used more frequently and can show both linear and non-linear behavior is called Sigmoid (S-shaped) function (Ham & Kostanic, 2001).

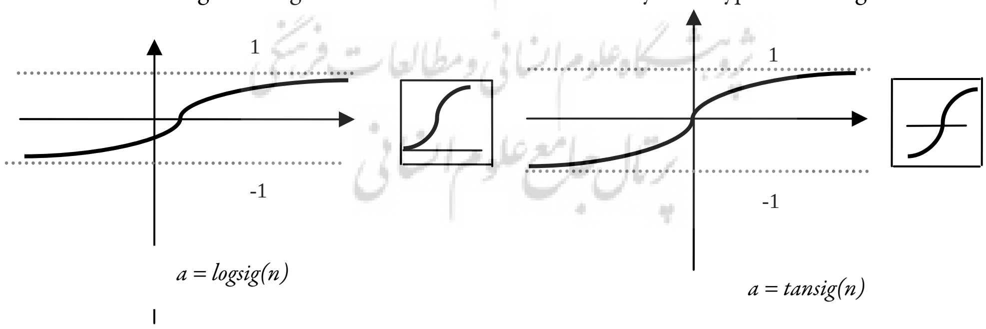
Sigmoid (S-shaped) function: This is the most common type that has been used as an activation function in ANN. Moreover, this function is continuous and differentiable. There are two types of sigmoid function. The first is the binary sigmoid function (logsig) because the output values of this function fall in binary range. The mathematical formula for the logsig function is:

$$y_q = f_{logsig}(v_q) = \frac{1}{1 + e^{-av_q}}$$

Where α is the slope parameter of the binary sigmoid function. The second type of sigmoid function is in bipolar form, which is the hyperbolic tangent sigmoid (tansig). The output of neuron q for the tansig function can be written as:

$$y_q = f_{tansig}(v_q) = \frac{e^{av_q} - e^{-av_q}}{e^{av_q} + e^{-av_q}} = \frac{1 - e^{-2av_q}}{1 + e^{-2av_q}}$$

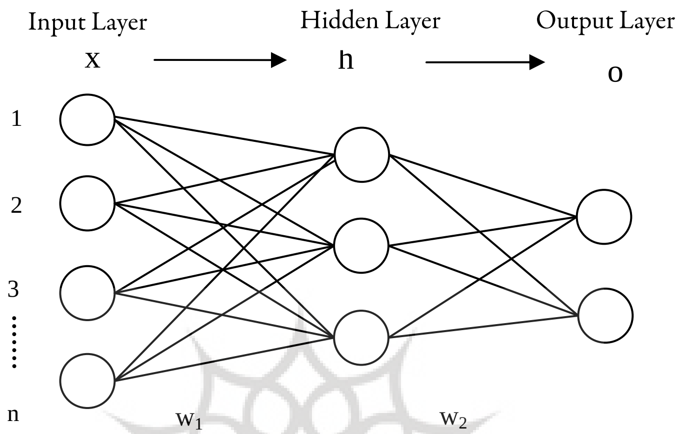
Figure5. Sigmoid Activation Functions: Binary and Hyperbolic tangent



4.1.2.3. Multi Layer Perceptron / Feed-forward Neural Networks

Multilayer perceptrons are built by linking a number of perceptrons together in a network from input to output (Zirilli, 1997). These are commonly referred to as ANN. The neurons in a network are distributed across layers. If the flow of the information through the network is from the input to the output and do not form cycles, it is called a feed-forward neural network (FFNN).

Figue6. Feed-forward Neural Network



From the figure above, it appears that an FFNN has exactly one input layer, one or more hidden layers, and one output layer. The first layer in the FFNN is the input layer, where the nodes take on the value of the input data. The numbers of input neurons are equal to the number of input samples. The last layer is known as the output layer. The number of outputs from the FFNN equals the number of neurons in the output layer. Any layers between the input and output layers are called hidden layers which are a group of nodes/neurons whose inputs and outputs only connect to other nodes, not an input or an output node of the network. The interconnecting lines shown in figure are weighted values. The weights are adjustable parameters and crucial aspects of the net.

The power in ANN models lie in the way their weights are adjusted. The procedures for adjusting the weights based on a specific data set (training set) are referred to as training the network. The ANN can be trained with or without supervision (Sarle, 1997a). The difference for supervised training is the correct results (target values, desired outputs) are known, while they are unknown in the unsupervised case. The objective of supervised learning is to discover the relationship between the two.

A FFNN is trained using the supervised method. Therefore, the target values are given during training so that the ANN can adjust its weights to try to match its outputs to the target values. In the training, the data is paired with the desired response of the network (supervised learning). However, after the training, the ANN is only tested by giving input values. The observed neuron output values in the output layer are compared with the correct target values, their difference being the output errors.

The basic idea of training is that the network will adjust the weights in order to learn the patterns in the training set. It starts by assigning random weights to the connection between each set of neurons in the network. Then the intermediate values in the hidden layer and output are computed. If the output is sufficiently close to the target, the process is terminated.

Using an adjusted network (trained network), it is able to generalize using the unseen data. However, if it cannot, the weights are adjusted and the process is continued until the solution is sufficiently close to the target. Since there are hidden layers in an ANN, the algorithm for a single perceptron fails. A solution to this problem is the back propagation (BP) training algorithm (Zirilli, 1997). Instead of only propagating errors, to correct the weights value, the errors are sent in reverse order through the network. A complete presentation cycle, an iteration to complete the forward and reverse cycles of all the training vectors, is called an epoch. Therefore, the number of epochs equals the number of times the training data are fed to the network. When other variables are held constant, the number of epochs can be used as an indication of the relative amount of training.

The activation function in the hidden layers is needed to introduce nonlinearity, which is the capability to represent nonlinear functions in the network. Without this function, the network is just a plain perceptron, which means that the network is just another linear function. It is the nonlinear attribute of ANN that makes the multilayer networks so powerful.

4.1.2.4. Back propagation Algorithm for FFNN

Here, the MLP or ANN training algorithm is presented. Before ANN can be trained with the back-propagation (BP), the ANN must be fully connected and layered, where each node is connected to every other node in the predicting and succeeding layers. The ANN must all be feed-forward only. As a result, BP is a method for computing the gradient of the error function related to the weights of a feed forward network.

5. Proposed Method

Here, the purpose of the model is to predict Tehran Stock Exchange Price Index (TEPIX) and its changes using FFNN (ANN) trained with Back Propagation. According to Kaastra and Boyd (1996), there are some steps in designing a neural network prediction model.

Table1. Steps in designing an ANN prediction model

Step 1:	Variable selection
Step 2:	Data collection
Step 3:	Assemble Training and Test Data
Step 4:	ANN paradigms
	Number of hidden layers
	Number of hidden neurons
	Number of output neurons
	Transfer functions
Step 5:	Evaluation criteria
Step 6:	ANN training
	Number of hidden layers
Step 7:	Testing Data with new inputs

6. Variable selection

The data used in this research are listed below:

1- Iranian Rial-American Dollar Exchange Rate (Rial vs $US)

2- World Gold Price

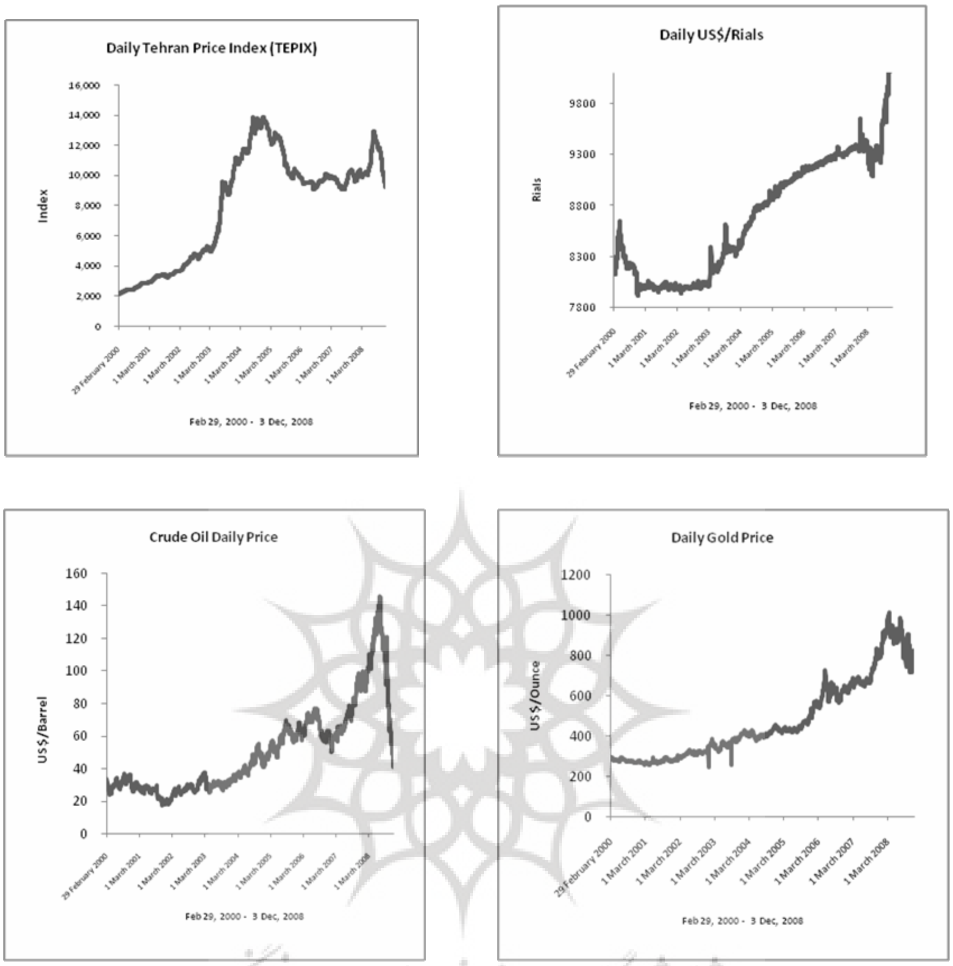
3- World Oil Price

4- Technical Indices (MA, MACD, RSI, ROC, MOM)

5- Daily trading volume of stocks

All data are provided in daily format.

Figure7. Time series plots of commodity prices (Oil and Gold), Dollar Exchange Rate and TEPIX



7. Data collection

TEPIX and all data used for the purpose of prediction consist of 2000 daily observations starting from 29/2/2000 until 03/12/2008.

7.1 Fundamental data structure

As pointed before, fundamental data contain Dollar Exchange Rate, Gold price and oil price.

7.2. Technical data structure

Here is technical data used for the research:

1- Moving Average (MA): is an indicator that shows the average value of a security's price over a period of time. The three time periods used were 5, 15 and 40 days. A simple moving average is defined by the following (Zirilli, 1997):

$$MA_i = \frac{1}{N}\left(\sum_{j=0}^{N-1} C_{i-j}\right) \quad i = 1,\ldots,P$$

N is the length of the moving average and P is the length of the period to the average.

2- Moving Average Convergence/Divergence (MACD): is an oscillator function calculating the difference between two moving averages. The two moving averages are the 26-period exponential and the 12-period moving average. The value of the MACD is calculated by subtracting the two (Murphy, 1999).

3- The Relative Strength Index (RSI): is an indicator that measures a price relative to itself and its past performance. The RSI is defined by the following equation (Murphy, 1999):

$$RSI = 100 - \frac{100}{1+RS}$$

RS = Average price increase in a given period/ Average price decrease in a given period

4- Rate of Change (ROC): calculates the price rate of change from the closing price, within a period difference. The period used is 12. The formula is as follows:

$$\text{Rate of change} = 100(V/V_x)$$

Where V is the latest closing price and Vx is the closing price x days ago (Murphy, 1999).

5- Momentum (MOM): is the difference between two prices (data points) separated by a number of periods. The momentum of a data series was calculated with a time distance of 12 periods. The formula for momentum is:

$$M = V - V_x$$

Where V is the latest closing price and Vx is the closing price x days ago (Murphy, 1999).

8. Assemble data

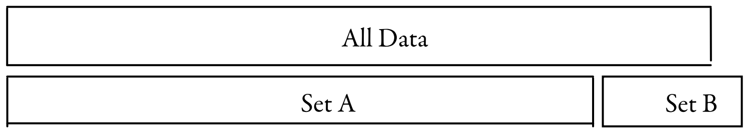
Data selected to be introduced to the network are divided into two subsets. These subsets form two major categories; training data set and test data set.

The training data set is a larger set and is used to define the parameters of the model, while the test data are used to measure their predictive ability. For this study, approximately 10% of the data was used as the testing set and the remaining 90% were used as training data. Furthermore, 3 different approaches were implemented to assemble the training and test data.

8.1. Consecutive selection

During this process, the training and testing sets are consecutively selected from the whole data set. The test data are selected immediately following the training set. The advantage of using this method is that the most recent observations are used. These may be more important than older data. Figure shows the separation of the data. The first 90% of the data was used as the training data and the following 10% as the test data

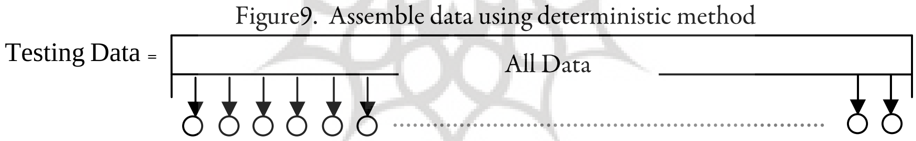
Figure8. Assemble data using consecutive method: Training set (A) and Test set (B)



8.2. Deterministic selection

Every tenth variable was selected to represent the test data. Consequently, approximately 10% of the data was reused for test data. The rest of the data was used for training data, which was 90% of the whole. Using this method, we can test the trained ANN not only by new data but also all data including the older one. Figure shows the deterministic selection method.

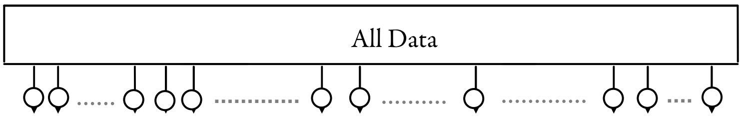
Figure9. Assemble data using deterministic method



8.3. Random selection

Here, the test data are selected as random, while the rest of the data are used as training data. The advantage of selecting the testing data randomly is that the danger of using a testing set characterized by one type of market is largely avoided. 10% of the whole data set was picked at random for the test data set and the rest of the data was used for the training set.

Figure10. Assemble data using random method



9. ANN paradigms

There are a number of parameters considered in building an ANN. As a result, there are unlimited ways to construct an ANN. Since the sole purpose here is to predict future values using historical economic values and data, the training process has a known outcome. For that reason, algorithms designed for supervised learning are ideal. The most suitable algorithm for finance is the Back Propagation (BP) algorithm; it is recognized as a good algorithm for generalization purposes.

The two major ANN parameters to be specified are the architecture and neuron properties. The ANN architecture defines its structure which includes the number of hidden layers and the number of neurons in each layer. On the other hand, neuron properties identify the transfer function that each neuron uses. The number of neurons in input layer is the easiest parameter to select since it is defined automatically according to the number of inputs for the ANN. There are other issues in building an ANN, which are discussed as follows.

9.1. Number of hidden layers

The power of an ANN lies in its hidden layers. The hidden layers can be treated as the core of the BP algorithm. Moreover, they can extract higher-level features and provide the ability for the ANN to generalize. Previous research by Hornik, Stinchcombe, White (cited in Picton, 1994) has shown that only three layers are needed. This paper gave a theoretical proof that the three layer perceptrons with a sigmoid output function are universal approximators, which means that they can be trained to approximate any mapping between the inputs and the outputs. The accuracy of the approximation lies in the number of neurons in the hidden layer. Thus, ANN architectures with more than two hidden layers are eliminated from this study, because they consume more resources in the form of time and storage without delivering results significantly better than or sometimes comparable to those structure with less the hidden layers.

According to Kaastra and Boyd (1996), all the ANN should start with one or at most two hidden layers. The constructive approach to the hidden layer size was to start with a network with only one hidden layer. Then, the ANN was trained until error stabilized. However, during the experimentation, it was found that one hidden later was not enough. If ANN used only one hidden layer, the gradient of the training decreased sharply. This means that the ANN could not generalize and remember the pattern of the data. Thus, additional hidden layer was used for the next training ANN. One of the keys to the constructive approach is to decide when to stop adding hidden layers. The training performed well for small number of neurons in both hidden layers. Consequently, it was decided to use only two hidden layers for this study.

9.2. Number of hidden and output neurons

There is no specified rule in selecting the optimum number of hidden neurons. Therefore, as with the number of hidden layers, the number of neurons in hidden layers was determined by experimentation. The experiments were also done constructively. Starting with a small number of hidden neurons in each hidden layers for the first training, training continues by adding neuron(s) in hidden layers. The experimentations continued by using a higher number of hidden neurons too since it is impossible to decide which number will be the optimum until the experiment is completed. In the end, there were a number selected. While it might be not the optimum number, it provides the 'best' model so far. However, a smaller number of neurons avoid a longer training time.

Deciding the number of output neurons is more straightforward than that for the hidden layer. It is recommended to always use one output neuron. Kaastra and Boyd in 1996 noted that the ANN with multiple outputs, especially where the multiple outputs have a wide range would produce lower quality results compared to an ANN with a single output. Here, the output neuron produced the prediction.

9.3. Transfer Functions

Cybenko stated (cited in Kaastra & Boyd, 1996) that any function could be approximated to an arbitrary accuracy by a network with three layers of neurons (2 hidden layers), specifically for a network that has a linear transfer function in its output layers and squashing transfer function in the hidden layers. Consequently, ANN in this study used sigmoid transfer function for all neurons in the two hidden layers and linear transfer function for the output layer. The hyperbolic binary sigmoid (tansig) was selected and not the binary sigmoid (logsig) because the data had been scaled in the range [-1, 1]. Thu s representational abilities of a tansig function can also show both linear and non-linear behavior. Since almost every financial data is not linear, this feature is appropriate.

10. Evaluation Criteria

To evaluate the performance of the trained ANN, we use MSE. This is the average of the squared values of the prediction errors or squared values of the deviations from the target, i.e. (Zhang, Patuwo and Hu, 1998):

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{x} - x)^2$$

Here $\hat{x}$ is the prediction from the ANN and x is the actual data. MSE eliminates the positive/negative problem by squaring the errors. The result tends to place more emphasis on larger errors.

11. Testing ANN with new inputs

The evaluation of the NN models was based on unseen data. The testing process was done repeatedly on the test data and the models that performed the best with each the performance measure were selected. This resulted in the future value of the stock index for the next day or next week. The testing data consisted of 10% of the whole data. By feeding unseen data to the ANN, we are able to see the generalization power of ANN.

12. Results of the research

In the three-layer network, there are 9 inputs neurons and one output. Using the trial and error method, experiments were done with a number of hidden units in both first and second hidden layers. The number of neurons in each hidden layer was from 1 to 16. So, there are many combinations of topology that can be done with this. Experiments for finding the topology were done from the next day prediction.

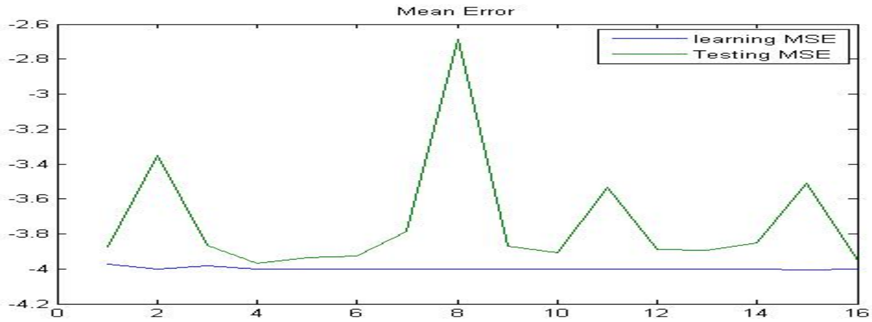
After the experiments were done, it was found that the ANN with 3 hidden units on the first hidden layer and 4 hidden units on the second achieved the best performance. Thu s, it was decided that topology 3-4-1 would be used for prediction.

The results of MSE gained from different topologies are listed in the tables below:

Table2. Results of (MSE) from 3-x-1 topology

Architecture	3-1-1	3-2-1	3-3-1	3-4-1	3-5-1	3-6-1
MSE	0.00013391	0.00044467	0.0001362	0.0001070	0.00011593	0.0001178
Architecture	3-7-1	3-8-1	3-9-1	3-10-1	3-11-1	3-12-1
MSE	0.00016359	0.000205	0.0001343	0.0000124	0.0002938	0.0001290
Architecture	3-13-1	3-14-1	3-15-1	3-16-1		
MSE	0.0001282	0.00014099	0.0003093	0.0001133		

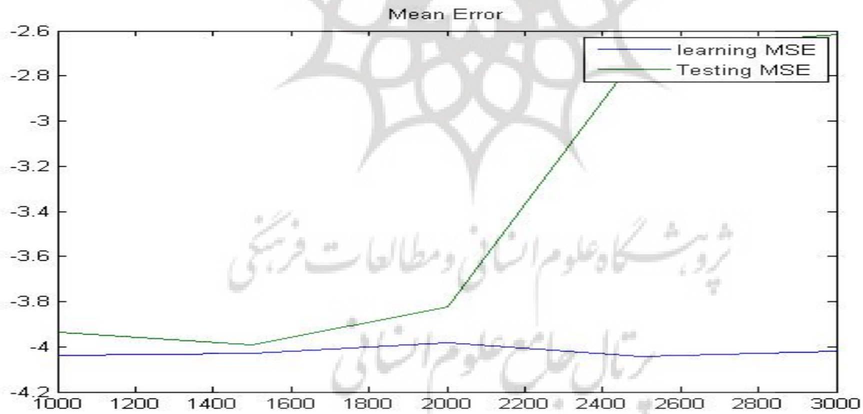
Figure11. Mean Square Error (MSE) from 3-x-1 topology



(In the above figure, horizontal axis measures number of epochs and vertical axis measures Mean Square Error)

The second parameter to be tuned is the epochs. The epoch values were determined experimentally through trial and error approach. The lowest amount of MSE was achieved when the epoch size was 1500.

Figure12. Epochs from 3-x-1 topology



(In the above figure, horizontal axis measures number of epochs and vertical axis measures Mean Square Error)

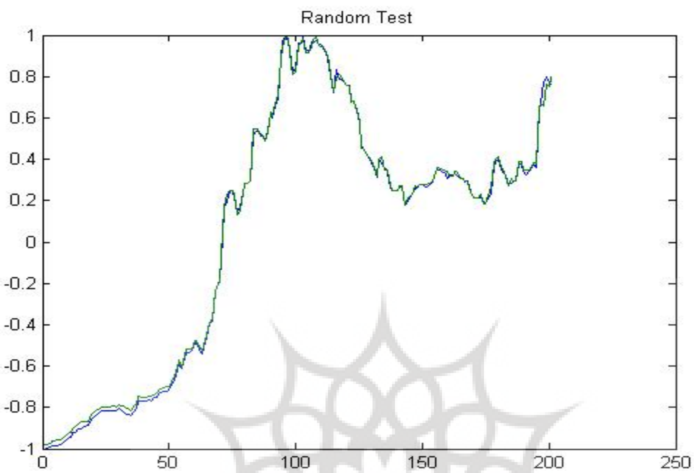
13. Conclusions

The results of the three assembling method of data are presented as follow.

13.1. Results of random selection

The figure below shows the results of testing 200 unseen data. The vertical axis represents 200 trading days and the horizontal axis demonstrates the amount of TEPIX that is normalized between 1 and -1. (The green line is the output of the ANN and the blue is actual output.)

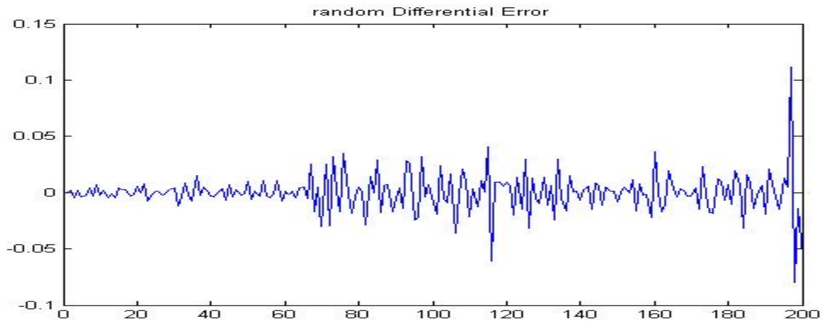
Fiure13. Next day prediction results of random selection



(In the above figure, horizontal axis measures 200 trading days and vertical axis measures the TSE index which is normalized between -1 and +1)

The differential error of the ANN is shown in figure 14.

Figure14. Differential error of random selection

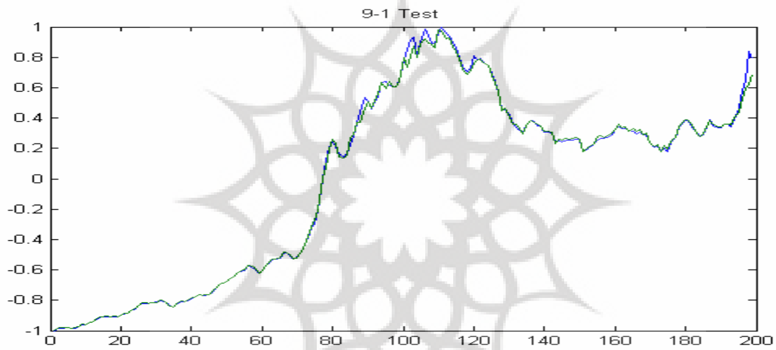


(In the above figure, horizontal axis measures 200 trading days and vertical axis measures of random differential error)

13.2 Results of deterministic selection

The figure below shows the results of testing 200 unseen data.

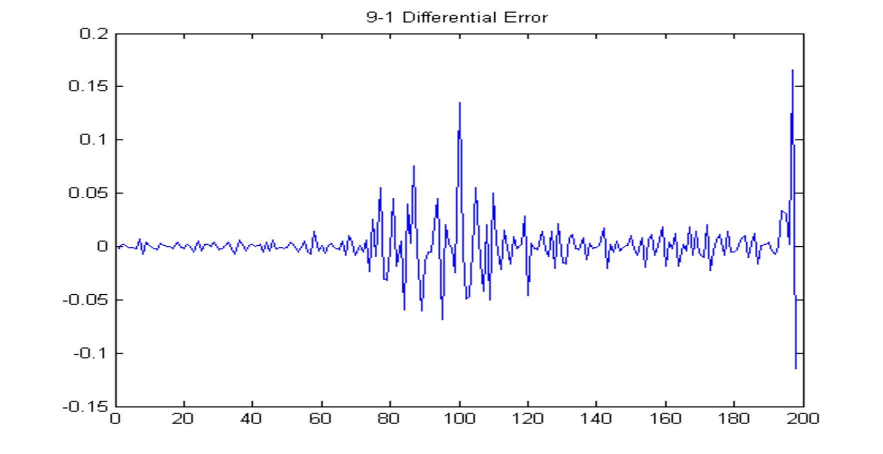
Figure15. Next day prediction results of deterministic selection



(In the above figure, horizontal axis measures 200 trading days and vertical axis measures the TSE index which is normalized between -1 and +1)

The differential error of the ANN is shown in figure 16.

Figure16. Differential error of deterministic selection method

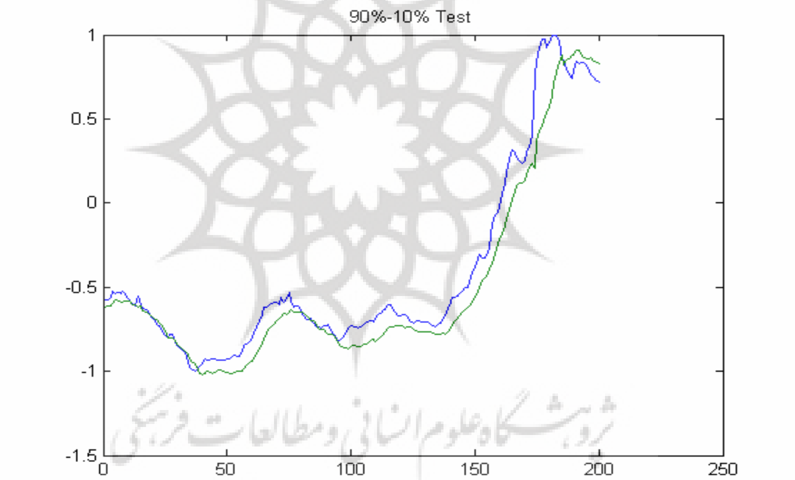


(In the above figure, horizontal axis measures 200 trading days and vertical axis measures of random differential error)

14.2 Results of consecutive selection

The figure below shows the results of testing 200 unseen data.

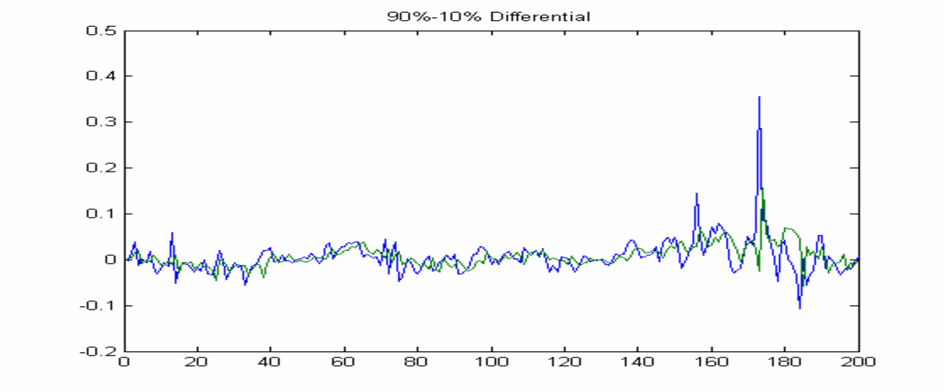
Figure17. Next day prediction results of consecutive selection



(In the above figure, horizontal axis measures 200 trading days and vertical axis measures the TSE index which is normalized between -1 and +1)

The differential error of the ANN is shown in figure 18.

Figure18. Differential error of consecutive selection method



(In the above figure, horizontal axis measures 200 trading days and vertical axis measures of random differential error)

This study tried to forecast the stock price index, specifically the Tehran Stock Exchange Price Index (TEPIX). A review of the literature showed that an ANN has large potential as a predictive model. Thus, an ANN was employed for the prediction of the TEPIX.

In addition ANN could indicate that there are relationships between TEPIX and the fundamental variables such as Rial Dollar exchange rate, oil price, gold price and trading volume and the technical variables such as Moving Average (MA), Moving Average Convergence/Divergence (MACD) The Relative Strength Index (RSI), Momentum (MOM) and Rate of Change (ROC).

In conclusion, an ANN predicted the stock index when proper and accurate data was supplied to the model. In general, an ANN was able to predict short-term future values (for the next day) of the TEPIX. As a result, an ANN was able to capture the relationship between the fundamental and technical data and TEPIX. In addition, results show that the fewer the hidden units, the better it is for ANN to generalize and learn the data rather than to memorize them.

References:

1. Brown, S.J., Otsuki, T., (1993). Risk premia in Pacific-Basin capital markets. Pacific-Basin Financial Journal, Volume 1, Issue 3, September, 235-261.
2. Clare, A.C., Thomas, S.H., (1994). Macroeconomic factors, the APT and the UK stock market. Journal of Business Finance and Accounting 21, 309–330.
3. Daniels. H., & Kamp, B. (1999). Application of MLP Networks to Bond Rating and House Pricing. Neural computing & Applications, 8, 226-234. Retrieved June 2004 from Springer-Verlag Link database.
4. Dastgir, M., S.H. Sajadi, V. Khodadadi, and P. Khalili, (2009), Studying the Impact of Risky Factors on Returns of the Petrochemical Companies, Journal of Economics Science (JES), Vol.9, No.1 (32).
5. Ham, F.M. & Kostanic, I. (2001). Principles of Neurocomputing for Science and Engineering. (intl. ed.). Singapore: McGraw-Hill Higher Education
6. Hellstrom T. & Holmstrom K. (1998, August 9). Predicting the Stock Market (Tech. Rep. Series Ima-TOM-1997-07). Vasters, Sweden: Malardalen University. Retrieved August 4, 2003, from Applied optimization and Modeling (TOM) at the Department of Mathematics and Physics Web site: http://www.cs.umu.se/~thomash/pub.htm
7. Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. Neural Networks, 4, 251-257
8. Iskandar, Nadia F. (2005). An Artificial Neural Networks Approach for Short Term Modeling of Stock Price Index. Master's Thesis. University of Regina, Saskatchewan, Canada.
9. Kaastra, K. & Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. Neurocomputing, 10, 215-236.
10. Kalyvas, E. (2001). Using Neural Networks and Genetic Algorithms to predict stock market returns. Unpublished Master's thesis. University of Manchester, Manchester, United Kingdom. Retrieved February 2, 2003 from Citeseer.IST database.
11. Kohara, K., Ishikawa, T., Fukuhara, Y., & Nakamura, Y. (1997, March). Stock Price Prediction using Prior Knowledge and Neural Networks. Intelligent Systems in Accounting, Finance and Management, 6(1), 11-22. Retrieved February 2004 from Wiley Interscience Journals database.
12. Leigh, W., Purvis, R., & Ragusa, J.M. (2002). Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network and genetic algorithm: a case study in romantic decision support. Decision Support Systems, 32, 361-377. Retrieved March 2004 from ScienceDirect Elsevier Science Journals database.

13. Lucas, C. & Karimi M.R. (2004). Forecasting the trends of Tehran Price Index (TEPIX) with the help of Fuzzy Logic Approach. Management Conference.
14. Jones, C. M., & Kaul, G. (1996). Oil and the stock markets. Journal of Finance 51, 463–491.
15. Jorion,P., (1990). The exchange-rate exposure of U.S. multinationals. Journal of Business 63, 331_345.
16. Malkeil, B. G. (1999). A Random Walk Down Wall Street. New York, London: W.W. Norton & Company.
17. McQueen,G., Roley, V.,(1993). Stock prices, news, and business conditions. Review of Financial Studies 6, 683–707.
18. Morelli, D., (2002). The relationship between conditional stock market volatility and conditional macroeconomic volatility Empirical evidence based on UK data. International Review of Financial Analysis 11, 101–110.
19. Murphy, John J. (1999). Technical Analysis of the Financial Markets. A Comprehensive Guide to Trading Methods and Applications. New York Institute of Finance.
20. Nag, A.K., & Mitra, A. (2002). Forecasting daily foreign exchange rates using generically optimized neural networks. Journal of Forecasting, 21, 501-511. Retrieved January 2004 from Wiley Interscience database.
21. Picton, P.D. (1994). Introduction to neural Networks. Chippenham, Wiltshire: MacMillan Press.
22. Podding, T. (1995). Bankruptcy Prediction: A Comparison with Discriminant Analysis. In A.P. Refenes (Ed.). Neural Networks in the capital markets. (pp. 311-324). England: John Wiley & Sons.
23. Refenes, A.P. (1995). Testing Strategies and Metrics. In A.P. Refens (Ed.). Neural Networks in the capital markets. (pp. 67-76). England: John Wiley & Sons.
24. Refenes, A.P., & Zaidi, A. (1995). Managing Exchange-Rate Prediction Strategies with Neural Networks. In A.P. Refenes (Ed.). Neural Networks in the capital markets. (pp. 213-220). England: John Wiley & Sons.
25. Ross, S.A., Westerfield, R. W., Jaffee, J.F., & Roberts, G.S. (1999). Corporate Finance. (2nd Canadian ed.). Whibty, ON: McGraw-Hill Ryerson Ltd.
26. Sadegh-vaziri, B. & Yazdan-panah (2006). Examining the Impact of Macro-economic variables on Tehran Price Index (TEPIX), Master's thesis, Alzahra University, Iran.
27. Sajadi, S.H., H. Farazmand, and H. Alisofi, (2010), Investigating the relationship between macroeconomic variables and stock index returns in cash: Tehran Stock Exchange, Journal of Economics Science (JES), Vol.10, No.2 (39).

28. Salim, F. & Sarebanha M.R. (2006). Evaluating the Efficiency of Tehran Price Index (TEPIX). Master's thesis, International University of Imam Khomeini, Iran.
29. Sarle, W.S., (Ed.). (1997a), Neural Network FAQ, part 1 of 7: Introduction. Retrieved June 2003, from periodic posting to the Usenet newsgroup comp.ai.neural-nets, ftp://ftp.sas.com/pub/neural/FAQ.html
30. Sarle, W.S., (Ed.). (1997b), Neural Network FAQ, part 2 of 7: Learning. Retrieved June 2003, from periodic posting to the Usenet newsgroup comp.ai.neural-nets, ftp://ftp.sas.com/pub/neural/FAQ.html
31. Singleton, J.C., & Surkan, A.J. (1995), Bond Rating with Neural Networks. In A.P. Refenes (Ed.). Neural networks in the capital markets. (pp. 301-310). England: John Wiley & Sons.
32. Steiner M., & Wittkemper, H.G. (1995). Neural Networks as an Alternative Stock Market Model. In A.P. Refenes (Ed.). Neural networks in the capital markets. (pp. 137-148). England: John Wiley & Sons.
33. Tam, K.Y. & Kiang, M. (1990). Predicting Bank Failures: A Neural Network Approach. In R. Batchelor & P. Dua (Eds.) Financial Forecasting Volume I: Stock Market Forecasting. (pp. 225-242). Cheltenham, UK: Edward Elgar Pub. Ltd.
34. Weiss, Gray (1992). "Chaos Hitts Wall-Street Theory, that is" Bussiness Week November pp. 138-140
35. White, H. (1988). Economic Prediction Using Neural Networks: The case of IBM daily stock returns. In R.R. Trippi & E. Turban, (Eds.). Neural Networks in Finance and Investing: using Artificial intelligence to improve real-world performance. (Rev. ed.). (pp. 469-481). Chicago, IL: Mc Graw-Hill/Irwin Professional Publishing.
36. Wilson, R.L., & Shadra, R. (1996). Bankruptcy Prediction using Neural Networks. In R.R. Trippi & E. Turban, (Eds.). Neural Networks in Finance and Investing: using Artificial intelligence to improve real-world performance. (Rev. ed.). (pp. 367-393). Chicago, IL: McGraw-Hill/Irwin Professional Publishing.
37. Yao, J.T. & Poh, H.L. (1995, November). Forecasting The KLSE Index Using Neural Networks. IEEE International Conference on Neural Networks, Perth, Western Australia, 2, 1012-1017. Retrieved May 2003 from CiteSeer.IST database.
38. Yao, J.T. & Tan C.L. (2001). Neural Networks for Technical Forecasting of Foreign Exchange Rates. In K.A. Smith, & J.N.D. Gupta (Eds.), Neural Networks in Business: Techniques and Applications (pp. 19-207). Hershey, PA: Idea Group publishing.

39. Yao, J.T., Tan, C.L, & Poh, H.L. (1999). Neural Networks for Technical Analysis: A Study on KLCI. International Journal of Theoretical and Applied Finance, 2(2), 221-241. Retrieved May 2003 from CiteSeer.IST database.
40. Yoon, Y., & Swales, G. (1991). Predicting Stock Price Performance: A Neural Network Approach. In R.R. Trippi & E. Turban, (Eds.). Neural Networks in Finance Finance and Investing: using Artificial intelligence to improve real-world performance. (Rev. ed.). (pp. 469-481). Chicago, IL: Mc Graw-Hill/Irwin Professional Publishing.
41. Zhang. G., Patuwo, B.E., & Hu, M.Y. (1998). Forecasting with Artificial Neural Networks: The state of the art. International Journal of Forecasting, 14, 35-62. Retrieved May 2003 from CiteSeer.IST database.
42. Zirilli. J.S. (1997). Financial Prediction using Neural Networks. London, UK: International Thomson Publishing