A NEW INTELLIGENT METHOD BASE ON NEURAL NETWORK FOR STOCK PRICE INDEX PREDICTION

A. Ghezelbash 1 F. Keynia 2 M. Mozaffari Legha 3

1. Faculty of Engineering, University of Jiroft, Jiroft, Iran, teacher.ghezelbash@yahoo.com
2. Energy Research Institute, Institute of Science and High Technology and Environmental Science, Graduate University of Advanced Technology, Kerman, Iran, fkeynia@kgut.ac.ir
3. Department of Power Engineering, Shiraz University, Shiraz, Iran, mozaffari.m@shirazu.ac.ir

Abstract- Prediction of stock price index movement is regarded as a challenging task of financial time series prediction. An accurate prediction of stock price movement may yield profits for investors. Due to the complexity of stock market data, development of efficient models for predicting is very difficult. In this paper, we proposed a new intelligent method base on Neural Network to forecast stock price index in the stock market. The proposed methods consist of two sections. In section one, a new two-stage feature selection algorithm is used to obtain the best-input data set and section two consists of an improved hybrid neural network algorithm to predict the stock price index variation. The proposed model results on Tehran Stock Market main index show the proposed method abilities.

Keywords: Index Stock Market Price, Artificial Neural Network, Prediction, TEPIX.

I. INTRODUCTION

Today’s investment in the stock market is interesting because of its high returns over time. Stock markets are affected by many highly interrelated economic, social, political and even psychological factors, and these factors interact with each other in a very complicated manner [1]. Tehran Stock Exchange from April 1990 attempted to calculate and publish its price index named TEPIX. Already number of accepted companies in stock tableau over 336 companies and worth of Tehran Stock Exchange nearly 65 billion dollars.

The stock exchange is an official capital market. That buy and sell stock companies or bonds of public or private accredited institutions, under regulations and certain rules is conducted. Stock Exchange as the pulse the country's economic regarded is economic analysts [2]. Indexes are exponents that are responsible for clarifying general conditions price, or return among the all or groups of listed companies. Therefore could be the index called a cursor that indicates the general level specific parameter (typically price or yields) among the group of variables (all or a group of companies) which changes its compared to with date source is considered.

The benchmark compared to the index value at any time on the source, is basis number. Normally basis number is considered 100 and by divisions Index value at any time on the basis number, can be growth, value desired parameters achieved within the groups of variables. Changes in a company's stock price determined easily and with regard to the stock price trend is obtained, but consider trend general price level of all the stock companies present in the stock market is done by the Main index price.

Refenes et al. [3] indicate that conventional statistical techniques for prediction have reached their limitation in applications with nonlinearities in the data set. Artificial neural networks (ANN), a computing system containing many simple nonlinear computing units as neurons interconnected by links, is a well-tested method for financial analysis on the stock market. Neural networks have been shown to be able to decode nonlinear financial time series data.

The first time White [4] of neural networks is used to predict the stock market. He was followed by the question whether neural networks are capable of nonlinear time series rules and champion of unknown movements in asset prices and changes of price are identified. Purposes white of this paper to show this that how a neural network predictor is able to do it. He proved this by offering an example of the daily prices of at IBM.

After the initial study White in 1988, neural network methods in the financial field various studies were performed in this field in the world. Chiang et al. [5] used an error back-propagation network for prediction pure asset prices of investment companies at year-end. They were compared network data and their results with results obtained from traditional econometric techniques and found that Neural Networks significantly better use of regression techniques when the data is low.

Aiken and Bsat [6] of a predictor neural network who was trained by GA, have used for the prediction interest rate treasury of the United States and concluded that neural networks can be appropriate for this work. Chan [7] had predicted financial time series by using predictor neural network and daily data of Shanghai Stock Exchange.
To speed up the convergence of the gradient descending algorithm and multiple linear regression was used to determine the weight. They concluded that neural networks could be completely satisfactory to the time series to predict and choice of weights in their approach leading to become less calculating costs. Lendasse [8] by using neural network methods forecast index, input of network includes two types of data were exogenous and endogenous who exogenous economic data include international stock price indexes (SBF 250, Topix, S&P 500) conversion rates (dollar/mark, dollar/yen) and the interest rate (quarterly and treasury interest rates) and endogenous data included index values in different history. They concluded from their study that use of neural networks perform well than linear methods.

Egeli et al. [9] attempted to predict the daily index of Istanbul Stock Exchange (ISE), the network inputs where the conversion rate of dollar/lira on before the day, index value on before the days, the nightly interest rate and five virtual variants for five days a week. His research concluded that perform neural networks to predict more accurately than a moving average for 5 days and 10 days. Research on Iranian business excellence in Tehran’s Stock Exchange, like Roza Gary Ahangar, Mahmoud Yahyazdehfars & Hassan Pournaghshband [10] in 2010 estimated the stock prices of activated companies in Tehran (Iran) stock Exchange.

This paper shows that without the use of extensive market data useful and proper prediction can be made. It begins with a general discussion of the possibilities of common stock price forecasting in an emerging stock market, like Tehran’s Stock Exchange (T.S.E). It is followed by a section neural network subsequently, a section is devoted to a case study on the stock price Index in Tehran’s Stock Exchange, pointing to the promises and problems of such an experiments.

For making this prediction, we have used of the nine indexes provided by the official website from the Tehran stock market [11], which include main index, First market index, Second market index, Financial index, Industry index, Weighted average of the top fifty companies, Simple average of the top fifty companies, Price index and cash returns and Free float index. For this purpose, we have used of data related to nine indices during the all workdays of Tehran Stock Exchange in 2011 and 2012 and the first 100 days of 2013, that’s a total of including 550 business days.

Review this nine indexes during 550 days shows this issue that these indicators do not have any significant relationship together. Course results obtained with a logical definition of each of these indicators are quite consistent [12]. For this reason, we tried to help the neural network in MATLAB software to derive meaningful relationships to data from 550 days. This paper done two types of forecasts in the Tehran Stock Exchange. In the first phase, we have predicted Main index during the recent 100 days, Due to data 550 days before that. In the second phase, which was staged original and useful of our project.

Each of the eight sub-indexes (First market index, second market index, financial index, industry index, weighted average of the top fifty companies, simple average of the top fifty companies, price index and cash returns and free float index) are our forecast for the next business day in the stock market. We estimate main index (the main and most important stock index) model based on eight indexes the past 500 days.

II. ARCHITECTURES OF PROPOSED INTELLIGENT METHOD

Action choice of index actual to predict the behavior of stock index and designed model PS is very difficult. ANNs were used in each phase, have different architectures, which are described separately in each stage.

As Zekic [13] states in her literature research for stock price prediction with ANNs, networks trained with a back-propagation algorithm with one hidden layer perform better. This idea is supported by many other sources.

According to Heaton [14] there is no theoretical reason to use more than two hidden layers and for most of the problems where the function being approximated contains a continuous mapping from one finite space to another, one hidden layer is enough. Chenoweth and Obradovic [15] prove that networks with large number of inputs do not necessarily perform better and generally lead to worse performance because of the amount of noise, while Steiner and Wittkemper [16] state that networks with fewer underperform compared to models that have more.

As opposed to the myth that neural networks are best used, including every data that seems to be handy and let it train for a large amount of time, a moderate amount of time with selectively chosen data using proper termination of training is known to perform better. These studies show that the number of inputs should be balanced so that they should not be numerous to degrade the performance of the network while being enough in numbers to contain information that is going to lead the network to find the hidden patterns.

Determine the number of layers, the number of inputs, outputs and the number of hidden layers and the number of neurons in each layer of an ANN is the most important design issues. There are several approaches or rules of thumb for choosing the number of neurons in hidden layer(s) while designing an ANN topology. The ones retrieved from the literature are [17]:

\[
\frac{(i+o)}{2}, \text{as defined by Man-Chung et al (2000)}
\]

\[
2^{l+1}, \text{as defined by Azoff (2003)}
\]

\[
\frac{2}{3}i(i+o), \text{as defined in the Neural Ware Software Manual}
\]

\[
\frac{2}{3}i+1, \text{as defined by Heaton (2005)}
\]

\[
k_i+1, \text{as defined by Freisleben (2005)}
\]

\[
\sqrt{i(i+o)}, \text{as defined by Freisleben (2009)}
\]

\[
\ln(i), \text{as defined by Gencay (2011)}
\]
where, $i$ the number of inputs, $o$ is the number of outputs and $k$ is the number of hidden layers. This paper used these relations in the construction of neural networks predictor.

### III. THE STOCK MARKET INDEXES PREDICTION

#### A. The First Phase, Prediction Main Index

Tehran Stock Exchange from March 1990 attempted to calculate and publish their price index names are TEPIX, this total time index included 52 companies that listed companies were. This index indicates that the total market value than the base year (or 1991) is several times. The Main index is wide application in country macroeconomic and investment market investors and is calculated using Equation (1) [18].

\[
TEPIX = \frac{\sum_i P_{oi}q_{oi}}{D_i} \times 100
\]

where, $P_{oi}$ is the price of $i$th company at time $t$, $q_{oi}$ is the number of shares outstanding $i$th company at time $t$, $D_i$ is the base value of time $t$ when equal source was $\sum P_{oi}q_{oi}$, $P_{oi}$ is price company $i$ at the time of source, $q_{oi}$ is number of shares outstanding company $i$ at the time of source and $n$ is number companies eligible index.

Relation 1 shows this subject is which to calculate Main index of the stock market, any of the other indexes does not use. So this phase our prediction is a very important. Our goal in the first stage, the prediction total index is based on eight main index stock market (First market index, Second market index, financial index, Industry index, weighted average of the top fifty companies, Simple average of the top fifty companies, Price index and cash returns and free float index).

The training algorithm such that for 450 days, every 8 indexes (value of all rows per day $X$ in First matrix) bring the Main index amount on that day (as value of column $X$ in second matrix) has entered into the ANN. This means that our ANN in input layer, eight neurons and in the output layer have one neuron (Figure 1) [19]. For testing ANN, a matrix with 8 rows and 100 columns (i.e. the eight sub-indexes in 100 days) into making it and then output results are compared with the Main Index on that day.

The ANN used in the output layer has one neuron. The only remaining issue, the number of neurons in an ANN layer is hidden. We created eight of ANN according to section II. ANNs obtained have come in Table 1. We used the neural network toolbox in MATLAB software to implement ANNs, the output matrix of ANNs (The value predicted for Main index during 100 days) with real data matrix (amount offered by the Tehran Stock Exchange, TSE, for Main index of during 100 days) are compared in Table 1. To examine the performance of the developed models, MAPE criteria were used to calculate errors.

The criterion mean absolute percentage error (MAPE), according to Equation (2), shows the mean absolute error that can be considered as a criterion to model risk to use it in real-world conditions [20].

\[
MAPE = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{x_i - \hat{x}_i}{x_i} \right|
\]

where, $X_i$ is the actual value and $\hat{x}_i$ is the forecast value.

Table 1. Compare the performance of each ANNs is built

<table>
<thead>
<tr>
<th>The number of hidden layer neurons</th>
<th>Neural network architecture</th>
<th>The error value (MAPE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>8-4-1</td>
<td>1.2569</td>
</tr>
<tr>
<td>5</td>
<td>8-5-1</td>
<td>0.9513</td>
</tr>
<tr>
<td>17</td>
<td>8-17-1</td>
<td>2.0145</td>
</tr>
<tr>
<td>6</td>
<td>8-6-1</td>
<td>1.1267</td>
</tr>
<tr>
<td>7</td>
<td>8-7-1</td>
<td>2.3651</td>
</tr>
<tr>
<td>3</td>
<td>8-3-1</td>
<td>0.7521</td>
</tr>
<tr>
<td>2</td>
<td>8-2-1</td>
<td>1.0083</td>
</tr>
</tbody>
</table>

Figure 1. The architecture of three-layered feed-forward ANN

For this reason, we provide data related to 550 days, we used data from the 450 days for training the ANN and data from the 100 days for test and conclusions of the ANN. I have done for training ANN: First, we create two matrixes. The first matrix with 8 rows 450 columns and the second matrix has one row and 450 columns. Each row is one of the indexes and each column the value of index 8 in a day these two matrixes are training our ANN.

Figure 2. Reply get ANN into MATLAB software
Accordance Table 1, we found that the most ideal ANN architecture, for our issue, structure 8-7-1 is, with fault $\alpha = 0.7521$ ability to extract the main index based on eight indexes for every day. This neural network after 738 replications and with the regressions shown in Figure 2 nearest reply to actual data with minimum error is. This paper used the training functions and transfer function for this ANN, following: TRAINGDM, TRAINGDA, TRAINGDX and transfer function LOGSIG.

Figure 3 the actual value Main index and its value predicted by ANN shows. At this value of the Main Index during the days 01/02/2013 until 28/05/2013 (100 days) were compared.

B. The Second Phase, Prediction All Indexes in the Next Business Day

To predicting each of the nine indexes stock market in TSE for next business day, we carefully movements Stock Exchange indexes in the recent period of 504 business days (about two years) appointed under profoundly study. Difficult to predict the index next day when more showed, we found that TSE in the days and months of the year, would experience different movements and many external factors influences these indexes. We decided to instead examine the external parameters, with long-term process indexes, new methods for prediction each of the TSE indexes we devised. Review of indexes movements, so it was for us that indexes in the period 100 days are almost identical and monotonous movements of their show (Figure 4). To note that knowing the next day's index value, the rights activist decision to buy or sell its stock, has greatest effect, more will appear.

As the Figure 4 shows, some of the indexes have been some days that are quite abnormal and out of range. If these values abnormal can be into ANN, (due to having too much difference with the rest of the data), cursed are unusually and far from truth responses. Example in Figure 4, the Price index and cash returns has two failures unusual. Solution to resolve this problem, pre-processing data prior entering the ANN (normalization data).

At this stage, we used a completely innovative and new approach to prediction indexes next stock day by ANNs. Initial data included of each eight index values during the recent 504 business days. We were innovation these basic raw data, we partition into two isodimorphous different matrixes.

$\sum_{i=0}^{7} \omega_{i,k} x_{i,t}$

$N_{k,t} = 1/(1 + e^{-n_{k,t}})$

$P_{q,t} = \rho_{q,0} + \sum_{k=0}^{7} \rho_{q,k} N_{k,t}$

$Y_{i} = \gamma_{0} + \sum_{q=0}^{7} \gamma_{q} P_{q,t}$

where, $\{x\}$ is the set of input variables, $\{y\}$ is the output variable, $i$ number of neuron input layer or the same number of simultaneous inputs. $k$ is the number of neurons hidden layer. Number $i$' input variables $\{x\}$. Sentence fixed $\omega_{k,0}$ is affected by a transfer function. Finally, to make the output $\hat{y}$, the response of the network sum is to sentence fixed $\gamma_{0}$.

ANNs constructed in this phase are three-layer networks. The input layer has 100 neurons, the number of hidden layer neurons is chosen according to Section II, and output layer neurons have the same amount of stock index is in next business days. ANNs constructed, and the results predicted for next business days in the Table 2.
E. Prediction of Main Indexes to Next Business Days

The only issue remaining, estimated Main index according to eight indexes in TSE is predicted for the next business days, to do this, we perform to the first phase, which is fully described in the Section A. The strong point of our innovative approach to predict nine indexes, combining of both approaches is mentioned in the sections A and B (Figure 7).

IV. CONCLUSIONS

In this paper, improved ANN models for stock price prediction have been proposed. In this models, for prediction each of the TSE indexes, we not used any parameters outside of stock market. Thus, the modeling error was decreased because of the match between the real data and those used for modeling. Satisfactory performance of the proposed model was demonstrated through comparisons with real stock price data and the results of other ANN models.

Predictions performed in this paper, for major reasons are better than other approaches do. Firstly, dollar exchange rate in Iran due to political is unrealistic, for the reason we used only from internal parameters of the stock market. Similarly, due to absence of Iran in WTO, fluctuations of formal internationally indicators such as S&P 500 and Dow Jones Stock Index, not effective in movements of Iran stock market indexes. The second reason for our success is related to many studies been done on indexes movements, so that we could detect periods in stock market which indexes have almost uniform motion.

The third reason for performance of this method simultaneous use of two methodologies for forecasting the movement of indexes, so that the one series of past prices of indexes, we extracted two completely different partition as the input of ANNS. The last advantage this approach using a large number of ANNS with various structures and different training algorithms. We forecasts values for each indexes the day 28/5/2013, were compared with actual values on the same day in the Table 3 and have brought them each rate errors. Our review shows that this issue is in which months year that symmetrical is the religious months of Muharram and Ramadan, TSE indexes, compared to other months, movements are more different.

According to Table 2 to determine the best architecture and training functions to obtain each index is in Table 3.

D. Prediction of Seven Indexes to Next Business Days

We’ve done for seven indexes the remaining (Second market index, Financial index, Industry index, weighted average of the top fifty companies, simple average of the top fifty companies, price index and cash returns and Free float index) phases first market index. Where it knowing the value of each of the eight indexes in TSE is very important for next business day, in training phase ANNs, in the software MATLAB, we used a variety of training functions, by nearest reply output to have the real data.

Replies ANNs with different architectures and different training functions are shown in Table 2. To obtain this answer, we made number of eighty-eight ANNS, and for each unique index, we choose a particular architecture. Table 3 also shows the error rate built ANNS, is in terms of percentage.

Table 3 shows the error rate built ANNs, were compared with actual values on the same day in the Table 3 and have brought them each rate errors. Our review shows that this issue is in which months year that symmetrical is the religious months of Muharram and Ramadan, TSE indexes, compared to other months, movements are more different.

According to Table 2 to determine the best architecture and training functions to obtain each index is in Table 3.

![Figure 7. Combining both approaches](image-url)
As future work, the economic significance of this study can be evaluated using a day-to-day portfolio management simulation system, taking into consideration the transaction costs resulting from trades (buy and sells). The combination of internal and external parameters of the stock market simultaneously as the inputs to ANN can have a large effect on the reducing amount output error of ANN. For each of the companies in the stock market prediction prices of next business day are very effective on the profit and loss of investors.

To compare the performance of an artificial neural network to linear regression, a regression equation was computed from the same data used for training the neural network to linear regression, a regression equation was tested to determine which tool is the better predictor. The results show that neural network has best performance for predicting stuck price network has was significantly better able to explain the relationship between Inputs and output.

For prediction more accurately of the next business days and also increase the number of days foreseeable (with acceptable error rate), following actions can be done:

- Increase the number of days that used price of indexes in those days as the ANNs input parameters.
- Modify the structure of ANNs that are used.
- Use of hybrid methods for reducing the training time of ANNs.
- Review of index behavior patterns in weeks and months of the same in several consecutive years. (Investors behavior for buy and sale of stocks in months of the year is different).

**REFERENCES**


BIOGRAPHIES

Ali Ghezelbash was born in Kerman, Iran, in 1983. He received the B.Sc. degree in Computer Engineering (software) in 2008 and the M.Sc. degree in Information Technology in 2012. He is a member of the Iranian Association for Management of Technology (IRAMOT) and Informatics Society of Iran. He has presented several journal papers. His research interests are in soft computing and artificial intelligent, especially optimization applications of neural network.

Farshid Keynia was born in Kerman, Iran. He received the B.Sc. degree in Electrical Engineering from Shahid Bahonar University, Kerman, Iran, in 1996, and the M.Sc. and Ph.D. degrees in Electrical Engineering from Semnan University, Semnan, Iran, in 2001 and 2010, respectively. Currently, he works in Graduate University and Advanced Technology, Kerman, Iran. His researches focus on forecasting methods in deregulated electricity markets, and artificial intelligent algorithms.

Mahdi Mozaffari Legha was born in Kerman, Iran, in 1987. He received his B.Sc. from Yazd Branch, Islamic Azad University, Yazd, Iran in 2009 and the M.Sc. degree from Saveh Branch, Islamic Azad University in 2011. Currently he is the Ph.D. student of Power Engineering in Shiraz University, Shiraz, Iran. He is a Trainer of Power Electrical Engineering at Kerman Branch, Islamic Azad University, Kerman, Iran. He is interested in the stability of power systems and electrical distribution systems. He has presented more than 10 journal papers.