Forecasting stock market return using ANFIS: the case of Tehran Stock Exchange

Maryam Ahmadifard¹, Fakhrie Sadenejad², Isa Mohammadi³, Kobra Aramesh⁴

¹ MA of management, Member of Young Researchers Club, Khorramabad Branch of, Iran.
² Master of Public Administration student at Islamic Azad University of Shooshtar, Iran.
³ Master of Public Administration student at Islamic Azad University of Shooshtar, Iran.
⁴ Master of Public Administration student at Islamic Azad University of Shooshtar, Iran.

ABSTRACT

The present study focused on the development of exact forecast methods and presentation of a good model by intelligent systems to estimate the stock return as a tool to meet the demands of the investors to cover the potential market risk to prove some tools to improve the quality of financial decisions. To do this, by the data of oil price, gold price, foreign exchange rate in free market and TEPIX in the period (2005-2009) as in dependent variables, the stock return with ANN, ANFIS models were predicted. In modeling the artificial neural networks MLP architecture under Levenberge-Marqurdt algorithm with three-layer (4-20-1) was used. In neural fuzzy systems by hybrid learning algorithm and Takagi-Sugeno Fuzzy Inference system after 25 iterations with two membership functions for each input variable (8 membership function) of (gbellmf) type, with sigmoid activation function, the optimal model was designed. To evaluate the validity of the proposed model, the results of two models were compared based on (MSE, NMSE, MDAPE, MAPE, R²) error evaluation indices. The results of the study showed the efficiency of both models in better prediction and better performance of fuzzy-neural systems to artificial neural networks.

Key words: forecasting, artificial neural network, system fuzzy, the adaptive neural fuzzy inference system (ANFIS), stock return

INTRODUCTION

To achieve the economic development, investment is one of the important factors. In taking investment decisions, achieving good returns is the first and the most important factor of investors. The stock return of the companies includes the changes of stock price during a period and revenue of stock including cash stock, dividend and issue right. (Robert A. Haugen (2001) Thus, it is a good criterion compared to stock price during a period for evaluation of the companies. The investors aimed to maximize the expected return. Although they want to reduce risk, return evaluation is the only logical way (before risk evaluation) that the investors can do to compare the different investments (Charles P Jones 2009). Thus, the evaluation and prediction of market return can be an effective aid in taking logical decisions of the investors. Normally, the information of stock is various, incomplete, ambiguous and indefinite. Thus, the prediction of its economic performance in future is challenged. One of the most important challenges is defining the effective variables on stock return. Some of the researchers as Azar and Karimi (2009) and Ronstin and Pavelzik (2005) by accounting information predicted stock return and stated that there is no fixed relation between stock return and accounting ratios at various industries level and this correlation is changed by various reasons and the relation is weak. The recent review of literature showed that stock market return is predicted based on the macro financial and economic variables. According to Chen et al.
(1986) in testing the validity of the Arbitrage Pricing theory, it was found that macro-economic variables were associated causally with stock return. The present study among the effective factors on stock return considered important macro-economic variables. Another challenge in stock return forecast is the selection forecast method. During the recent decades, data-based models of artificial intelligence were successful present in management and financial issues as a technique in prediction and presented many articles in this field. For the first time, White (1997) applied neural networks for stock return forecast in stock market. White predicted the daily return on stock of A.B.M Company by artificial neural networks and only used “past prices” variable in the time series. After the initial study of White (1988), the neural networks were considered in financial field and various studies were done in this regard. Renu (2010) predicted stock return by artificial neural networks and believed that in financial markets field, more than 80% of the studies for the design of optimum network, applied multi-layer Perspectron neural network. Karyel et al. (2005) compared the linear models of stock return prediction (Fafa and French, 1992) and forecast non-linear models (neural network model and genetic algorithm). The results showed that there was a significant difference between linear and non-linear models and totally, linear models are better than non-linear models. Trinkel (2005) predicted annual stock return of three great business public joint stock companies by fuzzy neural systems, and then compared the forecast ability of this method with linear model (ARIMA). The results of his study showed that non-linear model (ANFIS) was preferred to linear model (ARIMA) in stock return forecast. Atsalakis (2009) in a study investigated the forecast techniques of stock market forecast in more than 100 papers and they focused on neural and fuzzy neural networks from the view of input data for networks design, forecast method and performance measure criteria. The results of the study showed that soft computation techniques namely artificial neural networks and fuzzy-neural were accepted to evaluate the stock market behavior at macro level and the accuracy of the predictions was suitable. He believed that stock return forecast is complex due to the market fluctuations requiring the implementation of exact models. The solution to this problem is artificial intelligence systems providing useful tools for complex environment non-linear behavior forecast such as stock market. Melek (2010) predicted stock return by neural fuzzy systems (ANFIS) in Istanbul stock market. The study aimed to respond whether neural fuzzy systems are able for exact stock return forecast? To do this, by macro-economic variables as gold price, interest rate, dollar rate, inflation (consumer price index) and industrial production index were considered as independent variables (input) during January 1990 to December 2008 based on the evaluation criteria of performance ($R^2$-RMSE) and covariance, designed the optimal network. The empirical results showed that the designed model with accuracy 98% was a useful tool for monthly stock return forecast in Istanbul stock exchange market. The studies showed that neural networks are in efficiency learning and adaptation but negative feature of black box is dedicated to it. Fuzzy logic is not very effective in learning but approximate reasoning is presented as an advantage (Agrawal et al. 2010). The present study aimed at the limitations of two models and by fuzzy neural systems, a new developed instrument of data-based model and can learn artificial neural network with language interpretation presented by fuzzy systems and presented an applied model to predict financial variables with good accuracy to improve the quality of financial decisions to create a link between financial management and artificial intelligence.

**MATERIALS AND METHODS**

This study was applied in terms of aim and as the study is quantitative based on existing data for future prediction and there was no manipulation in independent variables, the study in terms of method and nature was causal retrospective method. To select the independent variables among the effective variables on stock return based on arbitrage theory and the studies in Iran (Raei, 2003) and abroad (Melek, 2010) and based on the underlying conditions on Iran stock market, macro-economic variables including foreign exchange, oil price, gold price and stock index (as a representative of total values of the companies listed on stock exchange market) were considered for stock return forecast. The data collection was done in two sections. First, library method was done by books, journals, conference articles, scientific sites, and the information of theoretical basics and review of literature. In the second stage, the quantitative data of the
study were collected during 5 years since 2005 to 2010 via the searching in study resources and the sites in internet as followings.

<table>
<thead>
<tr>
<th>Table 1- Study data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time series</strong></td>
</tr>
<tr>
<td>Stock return</td>
</tr>
<tr>
<td>Stock return</td>
</tr>
<tr>
<td>Gold world price</td>
</tr>
<tr>
<td>Gold world price</td>
</tr>
<tr>
<td>TEPIX</td>
</tr>
</tbody>
</table>

The data analysis method in the present study was based on artificial intelligence techniques (ANN, ANFIS) by MATLAB (2008) software. To understand the proposed models, in accordance with the estimation of the designed models and method, the techniques were described briefly. The design of ANN model: In artificial neural network, the data structure is designed by programming acting as neuron or node as the basis of the biological neural network and is consisting of dendrites, axon, cell and synapse. This simulation is observed in the following figure:

Table 2- The correspondence of artificial neural network and biological neural network

<table>
<thead>
<tr>
<th>Biological neural</th>
<th>Artificial neural</th>
</tr>
</thead>
<tbody>
<tr>
<td>dendrites</td>
<td>Inputs (independent variables)</td>
</tr>
<tr>
<td>axon</td>
<td>Output (dependent variable)</td>
</tr>
<tr>
<td>synapse</td>
<td>Relation weights</td>
</tr>
<tr>
<td>Activity function</td>
<td>Cell body</td>
</tr>
</tbody>
</table>

The layer, its outputs are the final output of the network is called output layer. All the other layers are intermediate layer or hidden layers (Degmar & Tres, 2011). The hidden layers receive inputs gradually from the origin and turn into the final output. Black box is called this name as for the hidden layers. Because their outputs are not observed completely. After the network is built (in the form of a computer program), learning process is started. During neural network training, a set of data is considered. First, dividing this set of the data into training and test set. First training network is trained by training set and then by test set, the network is evaluated to fine that the network can cope with the data that is shown or not. During the training, the error of optimum output and real output is measured. The training is aimed to reduce the error via weights adjustment. During the training, the error should be reduced and training is stopped when this reduction is less. The development process of neural network program applied in this study is including 9 steps presented by (FREEMAN, 1991) and the implementation process of these steps is shown in Figure 4. The parameters of the designed model are shown in Table 3.
Ahmadifard et al

The estimation of ANFIS model: In introduction of fuzzy neural systems, it should be said that fuzzy neural system. Is a neural network, its inputs are fuzzy sets and for system training, membership functions and fuzzy parameters are required. Fuzzy logic is a kind of logic replacing the conclusion method in the brain of human beings and to express the ambiguity in the form of a number, a function for membership in a set is introduced dedicating to each element, a real number ranging 0,1. This number shows membership degree of an element to the required set. The membership of zero element shows that the required element is out of the set while 1 shows that the element is in the set (Renu, 2010). Fuzzy systems are knowledge or rule-based systems. A simple if-then in fuzzy logic is having the following structure:

\[ B \text{ If } X \text{ is } A \text{ Then } Y \text{ is } \quad (1) \]

The first part of fuzzy rule (A is X) is called antecedent or premise and the second part is (Y is B), consequent or conclusion. A, B are determined values in fuzzy set with X, Y (parent set) (Hagan, 2004). Based on the type of rules, there are three types of fuzzy inference systems including Mamdani, Takagi-Sugeno and Tsukamoto. The only difference is the consequent part of fuzzy rules. In Mamdani system, the antecedent and conclusion are both linguistic variables. In Sugeno system, the conclusion is a mathematical term and it is not linguistic. In Tsukamoto fuzzy model, the conclusion part of any fuzzy rule is denoted by fuzzy set with uniform membership function (Atsalakis, 2010). Each of the fuzzy systems and artificial neural networks has merits and demerits. Fuzzy system can use human language and can use the experiences of experts and other people but cannot learn. Indeed, by observation data, fuzzy system cannot be trained but neural networks are able for self-programming by data set and they are implicit and cannot use human language. For the first time Jang (1993) applied the linguistic power of
fuzzy systems and neural network training and presented a system called Adaptive Neuron-Fuzzy Inference System (ANFIS). This is done by specific changes in artificial neural network components. For example, while ordinary neural networks are consisting of similar neurons, the constituent neurons of fuzzy neural networks are heterogeneous and fuzzy neuronal networks consist of various neurons with different calculation features. The fuzzy neuron concept is imagined as each fuzzy neuron denotes a linguistic variable as average, low, etc. Thus, neuron output shows a membership function, shows the input vector belonging degree to a linguistic item and using ANFIS systems by Takagi-Sugeno fuzzy system as 5-layer forward networks structure. Parameters modification is done only in layer 1, layer 4. Training these parameters is a two-stage process, in the first stage, the parameters are constant and the information is forward propagated to the fourth layer as the conclusion parameters are defined via the least squares estimator. Then, in the next stage or return way, the conclusion parameters are constant to propagate the error between the parameters based on gradient descent method. Adjusting ANFIS parameters can be done by error back propagation algorithm alone or as a combination of back propagation and least squares and the learning algorithm is called hybrid. (Zanchetin, 2010)

The important points required in ANFIS modeling are: 1- The type of function membership, 2- The number of membership functions for each input, 3- The number of EPOCHs, 4- Training algorithm, 5- Step size (learning rate) and training data. Based on four input variables in the present study, for each input, a separate membership function should be determined. To do this, via continuous change of various membership functions and the number of membership functions, for each input, two generalized bell-shaped built-in membership function (gbellmf) are defined. This function is determined by three parameters with soft and non-zero advantage in all the points. The number of input time series in the fuzzy neural networks is 2720 observations, of which 2176 observations are considered as input and 544 observations are output data, 90% of which, 1956 data are considered as training data and 220 data as test data. The optimal fuzzy neural model was used after 25 iterations.
To compare the forecast ability of artificial neural network and fuzzy neural network, Mean Square Error (MSE), Mean Absolute Percentage of Error (MAPE), the coefficient of determination (R2), Normalized Mean Square Error (NMSE) and median of the absolute predictive error (MDAPE). The efficient model is the one in which R2 is close to 1 and the rest of measures are close to 0.

**Table 6- The comparison of two models based on common measures**

<table>
<thead>
<tr>
<th>Model</th>
<th>MSE</th>
<th>MDAPE</th>
<th>MAPE</th>
<th>NMSE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural network (mlp)</td>
<td>1.1405</td>
<td>2.72</td>
<td>3.6443</td>
<td>0.1928</td>
<td>0.9981</td>
</tr>
<tr>
<td>(anfis)</td>
<td>1.2685</td>
<td>2.3904</td>
<td>3.6119</td>
<td>0.1835</td>
<td>0.9982</td>
</tr>
</tbody>
</table>

**Discussion and Conclusion**
Based on the changes in the current world, namely in developing countries encountered with various threats, these countries require suitable solutions to solve their economic problems to use their natural facilities and wealth. One of the important solutions is development of investment (1). Development of investment absorbed non-efficient capitals and directed them to the productive economic sectors and based on the direction of the investors (based on risk and return), the investment is guided in the industries with more profit and less risk and this leads into the optimum allocation of the resources. The main motivation of the investors is achieving good return. If we can predict investment return and present some models for it, more reliable conditions are created in capital market developing the investment in financial markets. There are a few researches on return forecast or stock price in TSE and mostly statistical techniques were used for forecasting. The recent studies showed that due to the complex nature of the time series including stock return, using non-linear flexible networks as artificial neural networks and fuzzy systems compared to the linear and no n-linear classic models for forecast, present better results and based on the ambiguity, non-linearity and uncertainty of TSE, the present study provided a good instrument to improve the forecast in stock market by artificial intelligence methods including neural networks and fuzzy neural systems as non-linear and free of the model. The empirical results of the study showed that 2-Fuzzy neural networks were more precise than artificial neural network and forecast error was less than 5%. 3- It was defined that the process of building a neural network model, a systematic process is not defined and it is an error and trial process and in this process, the quality of the world done and the provided model depend upon the available time for doing more tests and the interest of the researcher to conduct the tests and the more complete the tests, the more the chance of having access to the best model. 4- ANFIS has also some limitations: Only in Sugeno systems, ranging 0, 1, the system has one output and all the rules should have unit weight and it is mostly affected by the problem range. The more the number of the data, the better the results. 5-Based on the wide range of artificial neural network, there are many issues in financial markets complex issues and some of them are not dealt in this study and they can be used in further studies. One of the recommendations is the design of various neural networks with more time horizon, different learning algorithms, forecast by fundamental analysis variables and considering all internal factors (profitability, liquidity, etc.) and external factors (Macro economy indices, etc.) to make the link between financial management and artificial intelligence to promote the stock market more stable.

Acknowledgment

My gratitude goes to the lecturers and authors who helped me with promoting the quality of journal.

REFERENCE


