Öneri. C.3. S.14.
Haziran 2000. ss.23-27.

A PURPOSED APPLICATION OF ARTIFICIAL NEURAL NETWORKS IN FINANCIAL FORECASTING

İsmail Hakkı ARMUTLULU[1], Gürkan SERHADLIOĞLU[2]

[1]M.Ü.İ.İ.B.F. İşletme Bölümü, Doçent Dr., [2]Akbank T.A.Ş. Müfettişi

***Özet:** Hisse senedine ait alım, satım ve elde tutma kararlarının verilmesi yapay sinir ağlarının kullanılabileceği bu tip alanlardan biridir. Bu çalışmanın amacı yapay sinir ağlarının borsada alım-satım veya bekleme kararlarının verilmesinde ne şekilde kullanılabileceğinin incelenmesinden ibarettir. Bu amaçla öncelikle yapay sinir ağlarının tanımı ve ne şekilde çalıştıkları incelendikten sonra oluşturulan modelin performansı IMKB 30'a kayıtlı üç senet üzerinde gösterilmeye çalışılacaktır.*

I. INTRODUCTION

For stock evaluation currently there are many techniques available. Especially with the recent improvements in computer technology new techniques have been introduced. Some of them are artificial neural networks, genetic algorithms, and fuzzy systems [1].

The main purpose of this essay is to examine the potential of artificial neural networks in financial forecasting. For this purpose the performance of an artificial neural network will try to be shown on randomly selected stocks which are enlisted in İMKB 30 list. Before examining the details of the application and discuss the findings, we believe it will be wise to give some introductory information about artificial neural networks (ANN).

ANN is statistically based mathematical model and tries to emulate the learning process of a natural brain [1]. Their working process is the simplest copy of the process of a biological neuron in the brain [2]. In the biological model aggregated impulse of all the stimuli are processed in the center part of the neuron. The condition for the neuron to be activated is that the total impulse must be higher than a threshold value. According to experts, for a specific stimulus, the pattern between the most activated neurons changes in order to process this stimulus. For the same stimulus or similar stimuli again the same neurons will be activated. This process represents how the brain learns. With that property experts believe that brain stores the information fore later use. However, if the stimulus, which induced that pattern, does not occur in the future or rarely occurs, once firmly established connection pattern between the neurons will begin to loose or totally disappear. This means that the learned information is lost or vaguely remembered [3]. ANN is built by connecting one or more artificial neurons. But the main difference is that we use synaptic weights to replace dendrite [4]. For impulse aggregation we multiply each impulse with the correspondent weight. If the sum is greater than a given threshold value the sum is introduced to the activation function [5]. The purpose of this function is to restrict the fluctuation of the aggregated value to the lowest or the highest value. The result of this function represents the answer of the model for the given stimulus. There are different types of activation functions but mostly used ones are sigmoid and hyperbolic functions. Figure.1 shows an example of an artificial neural network. ANN can be classified by their learning algorithms and structure [2]. The fact that where we will use ANN constitutes an important factor in determining which learning algorithm will be used. ANN applications areas range from financial applications to speech recognition. Another factor is the availability level of statistical information about the input and output values. If there is enough information about input- output values then it is suggested that supervised learning algorithm should be used. Otherwise unsupervised type of learning algorithm is used [6]. In supervised learning algorithm output of the model can be controlled and if the discrepancy between the model output and real output is greater than the desired value, then synaptic weights are readjusted in order to diminish this discrepancy. After this first adjustment the same input values are reintroduced to the same model and again the control is made. If the discrepancy is still higher than the desired value weight adjustment is performed once more. These processes end until the gap between the model output and actual output is zero or diminishes to an acceptable value. Because we can't change the connection pattern between the artificial neurons from case to case we simulate this process by changing the synaptic weights. Before adjusting the synaptic weights first we should calculate the error value for each of the artificial neurons in the model. These values are calculated form the right to left [3]. The blew equation is depicting the weight adjustment process.

$$W_{new} = W_{old} + \eta * \Delta W \ [6]$$

where η is called as a known constant [4, 13]. Choosing this constant high fastens the learning process but at the same time the quality of the learning process is diminish. Suggested η constant for sigmoid functions is 0.1, for hyperbolic function 0.2 [1, 8]. In order to initialize the learning process we give random values for the synaptic weights from a range, which can be calculated with the following formula.

$$\left[-\frac{2.4}{Fi}, +\frac{2,4}{Fi}\right]$$

where F_i represent the number of inputs for a specific unit. Thus far we have tried to give some information about "delta learning algorithm" [9]. Another important aspect of this process is to determine the structure of the model and when to end the weight adjustment process. Whit the rule of thump 20 iteration is said to be the optimum iteration number [6, 7]. In general we can find the optimum iteration number with trials.

According to their structure ANNs can be classified as feed forward or feedback. The best example for feed forward neural networks is perceptron. In this model artificial neurons are located in layers which are parallel to each other. Best example for feedback artificial neural networks is Hopfield neural network [7].

The difference between perceptron and multi-layer perceptron is, as the name suggest, there are one or more hidden layers between the input layer and the output layer in multi-layer perceptron. With that additional feature multi-layer perceptron gains the ability to solve "XOR" problems.

To decide how large we will build a model we can use two main techniques. One of them is to initially build a large model and then calculate the importance of each artificial neuron to the performance of the whole model. Then we can eliminate the least important units. For the second technique we add one unit at a time and measure the incremental change of the performance of the model. According to experts, extra unit apart from the optimum size of the model creates very little additional value to the whole performance [4] The connection pattern among the neurons can be different if there is some specific information about the input values. For example in the existence of such a case each unit in the model can be in contact with the nearest neighbor [6]. Thus far we have tried to examine the "supervised learning" algorithm. As we mentioned before there is another learning process. This unsupervised learning algorithm is closest possible learning algorithm to the brain learning process [2]. In this algorithm the most common used learning algorithms are based on simulation based learning algorithms. According to most of the expectations this model will be the foundation of the future computers which will not only have thinking ability but emotions as well.

As for every model ANN has some weak and strong points. The most advantageous of this model is its flexibility [2]. With this property this model has an ability to adapt individual cases. On the other hand its most disadvantageous side is the difficulty of evaluation of its results. In order to eliminate this problem, studies on "Nuero Fuzzy" has been gaining momentum [10]. Some advantageous of ANN are as follows

- Artificial Neural Networks are sensitive to both the qualitative and quantitative properties of the variables.

- They are useful to establish some sort of relationship between the output and input sets when it is very difficult or some times impossible to establish such a link.

- There is no necessity for variables to fallow a certain distribution. The sample size affect may not pose an important aspect. Multi-correlation, although it will be wise to eliminate some variables, which have high correlation among themselves to improve the performance of the purposed model, is considered to be less important for ANN applications.

Apart from these advantageous there are some disadvantageous of ANNs. ANNs are dependable to the variables used in the training session. In some cases, the model can emphasize some variables as if they were important, but actually they are not. Because of economical and time limitations, training time is limited. And if the built model runs contrary to a well known and commonly accepted theory, then we may get only reliable performance if the application variables are similar to that of the training variables. The last but not the least the result of ANN can not be regarded as reliable or even correct in some cases. Even if we use the some variables we may not get the same results. So far we have tried to introduce the concept of ANN in basic. Because our purpose is to evaluate stock performance by using an ANN, it will be wise to look at some aspects of stock evaluation from the technical analysis point of view. So we may get some information about which variables can be used in our model.

Generally speaking there are three main analysis techniques in stock evaluation. These are fundamental analysis, technical analysis and effective market behaviors. For fundamental analysis one has to have substantial amount of knowledge in economics [11, 16]. The main idea lies on the theory of technical analysis is the belief that stock prices contain a lot of information about the stock [3, 11]. However it will be very normal to say that some companies misbehavior about dissent disclosures of company assets and liabilities may harm this belief. Recent developments in computer science and communication facilities, ordinary people who are interested in investing in stock exchange can perform technical analysis with the available computer software.

II. APPLICATION WITH THREE STOCKS

In our application we have chosen three stocks from IMKB 30 list randomly and tried to make some predictions about the very near feature performance of the stock by using an ANN. For this purpose a multi-layer perceptron was chosen.

In our model there are five input variables and three clusters. The input set represents the closing values of each session. The input values of highest and lowest prices, seven-session moving average price, closing price and the percentage figure of the number of stocks, which have higher closing price values have been compared to the previous session. Because our model can only except binary values, price figures will be divided accordingly to be between [0,1] interval. Our variables are ratio type variables. The main reason that we wanted to use real number variables is that ANNs are sensitive to both qualitative and quantitative aspects of the variables. We could also use nominal variables. In that sense we could ask some questions which would have either "Yes" or "No" answers. Few example questions would be; "has the transaction volume of the stock increased?", "has the index increased?" or "has the gap between highest and lowest price of the stock widen?". If the answer is "yes" than we assign "1" otherwise "0" as of variable. There are also some techniques to choose the most suitable variables [1]. Apart from those techniques this job can be done by trial. Transaction value was also considerate to be included in our input set. But later it was dropped because the added value of this variable to the overall performance of the model was considered to be very low. The main purpose of this model is to classify the input sets into three different clusters. These clusters represent buy, sell, or hold decisions. Numerical presentation of these groups are "100", "001" and "010" respectively. Those numerical values are to be used as target values during the training session. It is very simple to determine those target values by comparing the closing prices of two consecutive sessions. For example, if the price of the second session is higher than the previous one then the target value will be "100". If there is no change then the target value will be "010" and finally if there is a drop in the stock price then the target value will be "001". During the training session the input values of each session are considerate as single input set. Following steps summarize the very basic of working process of this model.

- Fist the model is trained with 10 input sets. During that stage by using back propagation learning algorithm our model stores the knowledge by changing the synaptic weights.

- After the end of the 11th session we introduce the input values of this session to the model. But this time there will be no change on the synaptic weights. We will get only the results of the model.

- During the decision making process we compare the computed target values of the end of the training session and the results of the 11th session. By doing so we can make some forecasts about the 12th session.

III. CONCLUSION

As we have mentioned before the most difficult part of the applications which involve the artificial neural networks is the difficulty level to interpret the results given by the model. In order to extract some rules we have to observe the behavior of the model. For this purpose we observed the performance of our model for a two-month time interval. After that we came up with a basic rule. At the end of our thesis study we draw a table to demonstrate this rule. In that we can make some decisions about the very future performance of the selected stock by just comparing the computed figures of the related cluster figures. In this comparison we compare the membership values of the buy, hold or sell decision clusters with the membership values of the same clusters which we get after introducing the input values of the 11th session. In this comparison we focus just on the figure "1" which represents the membership. Now in order to clarify this rule we will create a scenario. Assume that 10th sessions target value is "100". So we will end up with the membership figures such that the first one will be very close to "1" and the others will be very close to "0". After introducing the input values of the 11th session if there is a drop in the figure which is close to "1" then this can be interpret as the price of the stock will not continue its rise and will drop. If in this example the target value of the 10th session were "001" instead of "100" as the only change then we would say the price of the stock will stop falling and begin to rise. Of course if the computed figure for the "1" increases further more this time the on going trend of the price will continue its progress without any change. If there is no significant change of the figure we can say that price levels will be the same for the next session.

Those decisions are also spotted on a price graphic for each of the stock in order to see the performance of the model more clearly. ***Those price graphics and the decision table, which shows the decision rules, can be found at the end of this article.***

As we said earlier one of the input values in our model is the percentage value of number of stocks which have higher closing price value for session. Instead employment of β constant for each stock could be more suitable. Like this example other related values can also be tried for better performance. This process of trying different input values for better performance is being studied. In case of stock split the price scale of any stock can change. This could cause a problem, because the price value will change dramatically and this can lead to confusion. In this case the usage of the model should be postponed to get enough input data after split for training session But for the availability of some techniques to make some adjustments between the stock prices before the split and after the split mandatory delay can be eliminated.

REFERENCES

[1] Deboeck, J. G., **Neural, Genetic and Fuzzy Systems for Chaotic Financial Markets**, John Willey & Sons, Inc., New York, 1994.

[2] ---------, http: // www.altavista.com / Phuriwat (Neuralnet & Expert System).

[3] Wong, F.; D., Lee, "A Hybrid Neural Network For Stock Selection", **In Proceedings of the 2nd Annual International Conference on Artificial Intelligence Applications**, New York, 1993.

[4] Haykin, S., **Neural Networks**, Prentice Hall, New Jersey, 1996.

[5] Trippi, R. R.; Lee, K. J., **Artificial Inteligence in Finance & Investing**, Irwin Professional Publishing, Chicago, 1996.

[6] Wong, F., "Time Series Forecasting Using Back Propagation Neural Networks.", **Neuro-computing**, 2, 1991.

[7] Zeidenberg, M. Neural Network Models in Artificial Intelligence. , Ellis Horwood, New York, 1990

[8] Zimmermann, H., "Applications of Boltzman Machine in Economics", **Methods of Operations Research**, Vol. 60/61.

[9] Wong, F.S.; P.H., Goh; B.K., Quek, "Fuzzy Neural Systems for Stock Selection", **Financial Analysts Journal**, Jan.-Feb., 1992.

[10] Altrock, V. C., **Fuzzy Logic & Neuro Fuzzy Applications in Business &Finance**, Prentice Hall PTR, New Jersey, 1997.

[11] Francis, C. J., **Investments Analysis and Management**, McGraw Hill, New York, 1991.

[12] Gately, E., **Neural Networks for Financial Forecasting**, John Willey & Sons Inc., New York, 1996.

[13] Ross, J. T., **Fuzzy Logic with Engineering Applications**, McGraw Hill, New Mexico, 1995.

[14] ---------, "*Yapay Zeka ve Finansal Uygulamaları*", **Global Bülten**, Global Menkul Değerler A.Ş. Yayınları, İstanbul, Temmuz – 1999.

[15] Sarı, Y., **Borsada Göstergelerle Teknik Analiz**, Alfa Basım Yayım Dağıtım, İstanbul, 1998.

[16] Yuan, B. O.; Klır, J. G., **Fuzzy Sets and Fuzzy Logic Theory and Applications**, Prentice Hall, New Jersey, 1994.

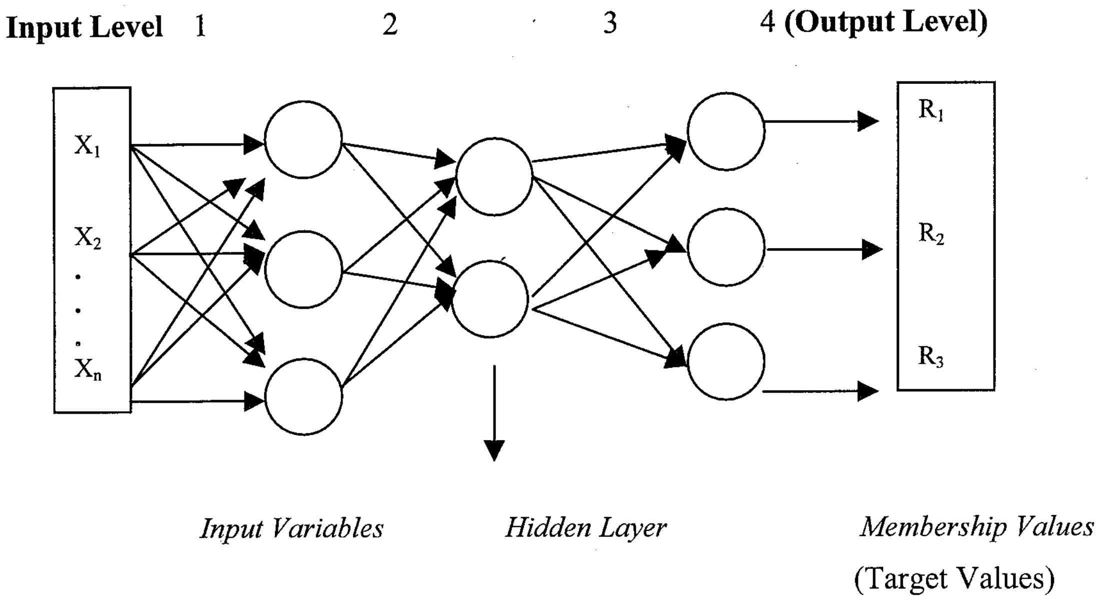

Figure.1: An Artificial Neural Network.

Table.1. Decision Table

The last target value in the 10-session training set.	1 0 0	0 1 0	0 0 1
Possible events after entering the input values at the end of the 11th session.	• A decrease in the value of "1" can indicate that the price increase of the stock will stop and movement will turn downward. • An increase in the value of "1" can indicate that the price increase for the stock will continue. • A very small change in the value of "1" (0 point 1 or 2 percent change) can indicate that the price level for the stock will not change.	• A decrease in the value of "1" can indicate that the price level for the stock will drop, otherwise increase.	• A decrease n the value of "1" can indicate that the drop in the price level will stop and price movement will turn upward. • An increase in the value of "1" can indicate the downward trend in the price level will continue. • A very small change in the value of "1" (0 point 1 or 2 percent change) can indicate that the price level for the stock will not change.