Assistant Professor Fatih ECER, PhD
E-mail:fecer@aku.edu.tr
Afyon Kocatepe University, Turkey

ARTIFICIAL NEURAL NETWORKS IN PREDICTING FINANCIAL PERFORMANCE: AN APPLICATION FOR TURKEY'S TOP 500 COMPANIES

Abstract. *Performance assessment which is conducted to determine whether business resources are efficiently used is of great importance for managers, business partners and investors. In this study, we classified Turkey's top 500 companies between 1993 and 2009 with respect to their financial performance using artificial neural networks. For this purpose, the variable profit before tax was determined as dependent variable and the variables sales, equity, assets, export, and number of employees were determined as independent variables to predict financial performance of the companies. Multi-layer perceptrons and radial basis functions among artificial neural network models were used in predictions. The study also used decision trees, which is another intelligent technique. This study aimed to compare prediction abilities of two different intelligent techniques. It was found that artificial neural network models showed a better performance than decision trees in terms of their accuracy rates and that they can be used in predicting financial performance. Furthermore, it was found that multi-layer perceptrons had better prediction ability than radial basis functions.*

Keywords: *Financial performance, artificial neural networks, decision trees, Turkey.*

JEL Classification: C45, C53, L25

1. INTRODUCTION

Accurate prediction of business performance is crucial for business partners, managers and investors as it reveals how efficient enterprise resources are used. In addition, the investors, who plan to invest in the companies, incline to invest in the companies with a high performance.

Principle objective of performance assessment is to determine whether enterprise resources are efficiently used. In addition to this, it is an important tool to determine whether right decisions were taken in the past and to set the financial structure of the company. Business performance consists of the combination of financial and non-financial performance. Profit, efficiency, innovative activities, and sales data can be used for financial performance assessment (Düzakın and Düzakın, 2007; Ecer, Ulutagay and Nasiboglu, 2011; Hafeez, Zhang, Abdelmeguid, Malak and Iqbal, 2000; Delcea and Simion, 2011). In literature, Beaumont and Schroeder (1997), Güleş, Türkmen and Özilhan (2011), Kim and Arnold (1993), Kotha and Swamidass (2000), and Sun and Hong (2002) analysed performance assessment.

The use of models which accurately predict financial performance might help business owners and managers to take measures in a timely manner and might prevent investors to make losses. Good or poor performance of companies depends on the interaction of various variables with non-linear relationship. Thus, the principle reason for the failure of obtaining good results from predictions using models such as linear regression analysis is the complex relationships between the variables. Models which can analyse complex relationships should be used to perform better predictions. Artificial neural networks (ANNs), which fall into intelligent techniques category, can yield quite good results in cases of non-linear relationships between variables. ANNs are successfully applied in stock selection prediction in finance (Krishna Kumar, Subramanian and Rao, 2010), inflation prediction (Choudhary and Haider, 2012; Moshiri, Cameron and Scuse, 1999), firm failures prediction (Youn and Gu, 2010), stock market index prediction (Guresen, Kayakutlu and Daim, 2011; Kumar, 2009; Merh, Saxena and Pardasani, 2011; Yudong and Lenan, 2009), and credit risk assessment (Boguslauskas, Mileris and Adlytė, 2011; Han and Wang, 2011).

The aim of this study is to show the usability of ANN models in predicting financial performance. Two of ANN models, Multi-Layer Perceptrons (MLP) and Radial Basis Functions (RBF), were used in the study. Financial performance was also predicted using decision trees which is also one of the intelligent techniques to determine how good predictions ANN models produced. It should be emphasized that financial performance prediction in the study refers to classification of firms as successful or unsuccessful ones. Accurate classification rates realized in test sets were used to compare prediction performance of the models. This rate was obtained by dividing the number of correctly categorized companies by the number of all companies.

The study consists of seven chapters. The second chapter of the study includes a literature review. The models used in the study were briefly explained in chapter three. Variables used in data set and analyses were analysed in chapter four. Application was explained in chapter five and findings were discussed in chapter six. A general evaluation was made in the last chapter.

2. PREVIOUS STUDIES

The intelligent techniques covered are decision trees, fuzzy logic techniques, data envelopment analysis, quadratic programming, ANNs, Bayesian models, soft computing, and support vector machines. Among intelligent techniques, ANNs are widely used in studies in finance. While some of the studies used a single method, some of the studies used more than one method. Examples of the previous studies are presented below.

Costea, Eklund, Karlsson and Voineagu (2009) classified Scandinavian Telecommunication Companies as failed or successful. First they used the Self-

Organizing Map (SOM) algorithm to cluster the companies, constructing a two-dimensional U-matrix map. They then attached the class labels to each data row from the dataset and applied two classification methods Multinominal Logistic Regression (MLR) and Decision Tree (DT) analyses to develop class prediction models. They found that while MLR is more optimistic than DT, the two methods results were very similar.

Coats and Fant (1993) proposed an ANN model as an alternative method of the same ratios used by Multivariate Discriminant Analysis (MDA). They compared the results of ANNs and MDA using the data of 282 firms. They showed that ANNs outperformed MDA and revealed MDA could be considered equivalent to a special case of ANNs when the input variables are linearly separable.

Ince (2006) indicated usability of ANNs in portfolio management and found that maximum yield can be obtained by using ANNs.

Avcı and Çinko (2008) analysed the effectiveness of ANNs in predicting daily index returns using different ANNs models. The researchers found that none of the ANNs were superior to other ANNs.

Ban and Mazıbaş (2009) used ANNs, discriminant analysis and logistic regression analysis to classify the banks with financial failure and successful banks in Turkey and reported that ANNs showed a better performance than other models.

Kumar (2009) carried out a study to predict index values of USA and Chinese stock markets and found that when compared to ARIMA models, ANNs predicted index values better.

Youn and Gu (2010) used ANNs and logistic regression analysis in predicting whether the firms were successful or not and indicated that ANNs were a better tool for prediction. Furthermore, they emphasized that logistic regression analyses can be preferred for prediction.

On the other hand, in a study which was carried out to determine stock market indexes, Merh et al. (2011) compared ARIMA models and ANNs to determine the stock exchange index and found that ARIMA models made better predictions than ANNs.

3. PREDICTION MODELS

Intelligent techniques are successfully used in prediction, classification, optimization, generalization and association. In this section, we describe some intelligent techniques that are used in this paper in predicting financial performance. The present study used ANNs and decision trees among intelligent techniques. What follows are brief descriptions of the prediction models.

3.1. Decision Tress

Decision trees have a tree-like structure consisting of root, branches, and leaves. It can visually show independent variables affecting dependent variables in the form of a simple structure. Strong aspects of decision trees are that they don't require large calculations; they can be easily applied to complicated data; their results are easy to interpret and they are not affected by deficient data during analysis.

Some of the decision tree models include ID3, C4.5, CHAID and C&RT/CART. C&RT, which is also termed as classification and regression trees, is a widely used decision tree model. It allows for classification of data by dividing them into two homogenous sets. This study used C&RT model as it attempted to classify the companies as the ones with good performance and the ones with poor performance.

3.2. Artificial Neural Networks

ANNs methodology was explained by Amari (1977), Feldman and Ballard (1982), Korhonen (1984), Lippman (1987), Minsky and Papert (1988), Grossberg (1988), Wasserman and Oetzel (1990), Burke (1991), Kuan and White (1994), Hagan, Demuth and Beale (1995) and Bishop (1996) in detail. In this chapter, ANNs will be explained briefly.

An ANN consists of inputs, weights, summation functions, activation function and outputs. The main purpose of the network is to produce a set of output in response to a set of input. While performing this task, the network is first trained with the examples shown to it and then its performance is tested. Training of the network refers to identification of weight values of the connections. Weight values are at first designated randomly. During the training process, weight values are changed and optimum weights are found. Thanks to the found weights, the network gains the ability to make a generalization. In the training process, the data are divided into two sets as training set and test set. Training is completed when the networks begin to produce correct answers to all examples shown to them. Then, the performance of the networks is tested on the basis of their responds to new examples shown to them, which are included in the test set. Even if a trained network is shown examples containing incomplete data, the network continues to operate. Since information is spread to the network, even if some neurons break down, the network continues to operate. The network designates the weight value of the variables; however, the users don't know what these weights mean. For this reason, the information of the network cannot be interpreted. In addition to this, other disadvantages of ANNs include obtaining an acceptable solution rather than the optimum solution; its unsuitability for the identification of parameters such as momentum coefficient and learning coefficient and lack of rules which determine the structure of the network. ANNs are often thought of as "black box" as it is very difficult to understand how they solve a problem. However, the fact that obtained results are satisfactory has eliminated the criticism developed for ANNs methodology.

3.2.1. Multi-Layer Perceptrons

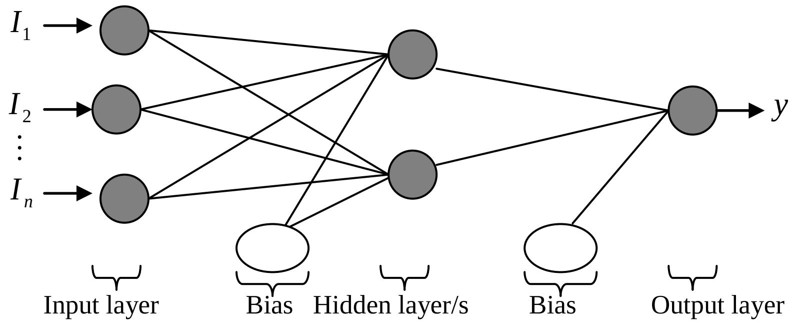
MLP networks presented in Figure 1 were developed by Rumelhart, Hinton and Williams (1986). This is a feedforward network since the information moves only in forward direction. MLP network's training process is usually realized with the backpropagation (BP) algorithm. The BP algorithm seeks to minimize error between the observed output and the desired output. MLP network contains three types of layers: the input layer, one or more hidden layers, and the output layer.

In an MLP network, generally a non-linear sigmoid or hyperbolic tangent function is used in the hidden layer. Sigmoid function can be expressed as follows:

$$\Psi(x) = \frac{1}{1+e^{-x}} = \frac{e^{x}}{e^{x}+1} \qquad (1)$$

Sigmoid function is used in classification problems as it allows for output values to receive values at $(0,1)$ range.

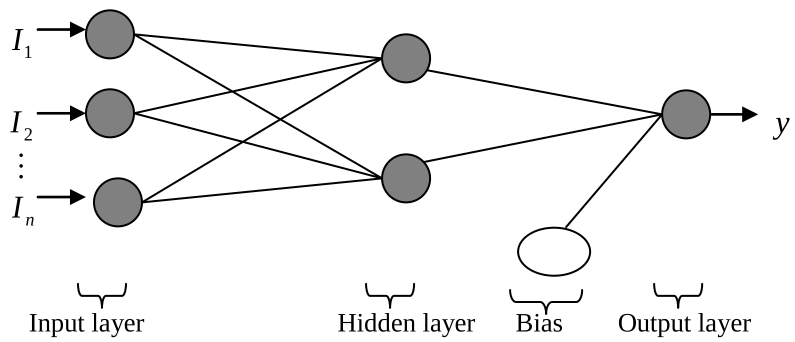
Figure 1: A MLP architecture



3.2.2. Radial Basis Functions

Although RBF networks indicated in Figure 2, which are feedback networks, resemble MLP network, its main distinction from MLP is that it reaches solution more rapidly due to fast data processing process. In addition, it has a single hidden layer and uses both linear and Gaussian functions as activation function.

Figure 2: A RBF Network

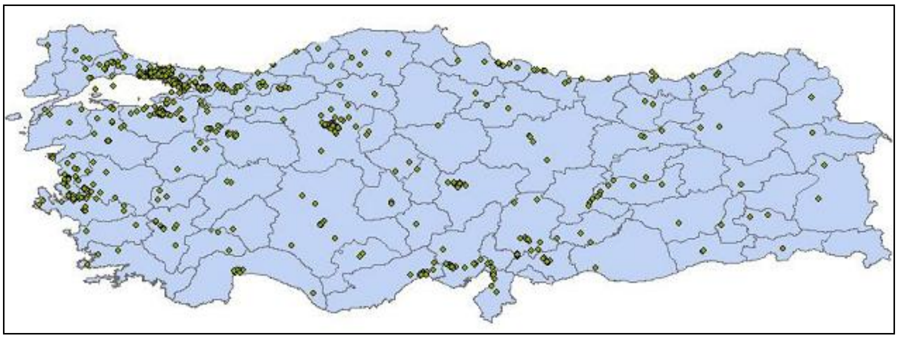


The working principles of RBF networks can be explained as follows: When a data comes to the input layer, the RBF network produces a centre vector. The distance between the vectors obtained from input data and centre vector are calculated and these distances can be transformed in the hidden layer using activation function. Thus, output of each hidden layer is in the form of a function of the distance between centre vector and input vector. Then, the output values of the hidden layer and their weights are multiplied and the obtained value is written in its corresponding place in activation function. As a result, output values are obtained. Following this, errors are calculated and the model starts learning process to reduce errors. It should be emphasized that the learning process takes place only in the hidden layer and therefore the solution is reached quickly.

4. DATA SET AND VARIABLES

The data used in analyses belong to the top 500 industrial enterprices in Turkey based on sales from production between 1993 and 2009. The data were supplied from Istanbul Chamber of Industry (ISO). Geographic locations of the companies included in top 500 in 2009 are presented in Figure 3.

Figure 3: Distribution of companies throughout Turkey. (Ecer et al., 2011)



The data set of used in the analyses consisted of sales, equity, assets, export, number of employees, and profits before tax of the companies.

Input and output values used in the study can be summarized as follows:

Input 1- Equity capital: The value of an ownership interest in property, including shareholders' equity in a business

Input 2- Export: A value of export income in a year

Input 3- Assets: A value of total assets

Input 4- Number of employees: Average number of employees in a firm in a year

Input 5- Sales from productions: Income that a firm receives from its normal business activities

Output- Profit: A value of profit before tax in a year

Since the data announced by ISO are in unit of Turkish Liras, the data were converted into US Dollars based on the exchange rate of the corresponding year announced by Turkish Central Bank to standardize all data. Statistical values are presented in Table 1.

Table 1: Statistical data of the top 500 companies in Turkey (1993-2009)

N	Statistics	Sales (million $)	Equity (million $)	Assets (million $)	Exports (million $)	Employee	Profit (million $)
8500	Mean	187.56	73.61	159.80	48.07	674.48	10.95
	Std. dev.	605.84	301.01	570.13	174.59	1880.24	53.23
	Std. error	6.57	3.26	6.18	1.89	22.79	0.60
	Median	73.84	20.18	55.83	13.02	427	3.30

Profit variable, which is the dependent variable, should be categorical to determine prediction performance of the models. In this context, while converting the profit variable, which is in fact a constant variable, into a categorical variable, the companies with a profit of 1 standard deviation above the profit average were categorized as the companies with a good performance; while the companies with a profit of 1 standard deviation below the average were categorized as the companies with poor performance. According to Table 1 that profit average is 10.95 and the standard deviation is 53.23. Therefore, the companies with a good performance were found to be those with a profit of greater than $10.95+53.23=64.18$. On the other hand, the companies with a poor performance were found to be those with a profit of smaller than $10.95-53.23=-42.28$ value. In summary, the companies included in the study were categorized as the ones with a good performance and poor performance as follows.

$$\text{Poor performance if profit} < -42.28$$
$$\text{Good performance if profit} \geq 64.18$$

The companies with a profit between -42.28 and 64.18 were excluded from the analysis. The remaining 499 examples were included in the analyses. In addition, of the 499 examples, 144 were included in the companies with a poor performance, while 355 were included in the companies with a good performance.

5. APPLIYING THE INTELLIGENT TECHNIQUES

It was observed that in analyses performed with both ANNs models and decision trees, the best result was obtained when the 70% of the data was spared for training and 30% was spared for test. For this reason, the data were divided into two sets at specified rates in all analyses.

In analyses performed with MLP network, sigmoid function was used as activation function and BP algorithm was preferred as learning algorithm. The best result was obtained when, among model parameters, learning rate was determined as 0.4 and momentum coefficient was determined as 0.9. In this context, the results of the analysis are presented in Table 2.

Table 2: Classification of MLP network

		Poor	Good	Accuracy rate (%)
Training set	Poor	46	6	94.7
	Good	2	98	
Test set	Poor	14	5	92.1
	Good	0	44	

According to confusion matrix in Table 2, MLP network separated the companies with a good performance from those with a poor performance at a high rate in the training set (94.7%) and test set (92.1%). The model realized this performance with one input, one hidden and one output layer. Hidden layer of the network included 4 neurons. Furthermore, the error rate of classifying the companies with poor performance as companies with good performance (Type I error) was found to be 26.3%, the error rate of classifying the companies with good performance as the companies with poor performance (Type II error) was found to be 0% and general error rate of the model was found to be 7.9%. It was observed that based on error rates, MLP network classified all of the companies with a good performance with zero error.

Analysis results using RBF network are presented in the confusion matrix in Table 3. There were 7 neurons in hidden layer of the network architecture. RBF network classified the financial performance of the companies in the test set with an 87.3% accuracy, which can be considered as a very good ratio. According to confusion matrix, the Type I error rate of the network was 26.3%, Type II error rate was 5.6%, and general error rate was 12.7%.

Table 3: Classification of RBF network

		Poor	Good	Accuracy rate (%)
Training set	Poor	37	15	85.6
	Good	8	100	
Test set	Poor	14	5	87.3
	Good	2	34	

Analysis results performed with C&RT among decision tree models are presented in Table 4. The model performed a classification in the test with 86.1% accuracy. Type I error of the model was found to be 36.4%. In other words, C&RT model evaluated more than one third of the companies with a poor performance as the companies with a good performance and classified them wrongly. According to confusion matrix, Type II error and general error rate of the model were 6.7% and 13.9% respectively.

Table 4: Classification of C&RT

		Poor	Good	Accuracy rate (%)
Training set	Poor	73	38	86.7
	Good	10	241	
Test set	Poor	21	12	86.1
	Good	7	97	

6. DISCUSSIONS

Accurate classification results and error rates obtained from the test set are summarized in Table 5.

Table 5: Results obtained on the basis of test set

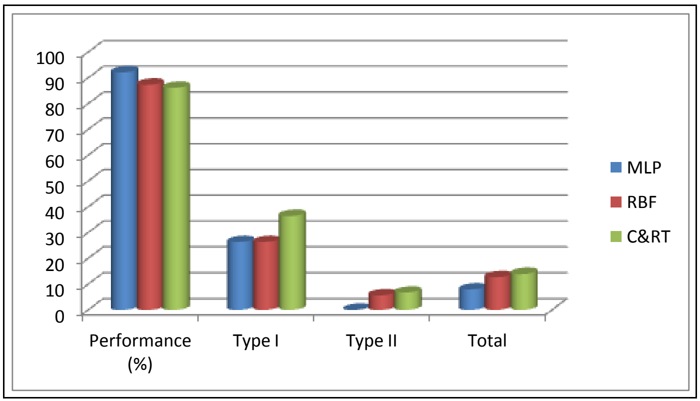
			Errors (%)		
Model		Accuracy rate (%)	Type I	Type II	Overall
ANNs	MLP	92.1	26.3	0	7.9
	RBF	87.3	26.3	5.6	12.7
Decision tree	C&RT	86.1	36.4	6.7	13.9

Comparing the three classification techniques in terms of accuracy rates, we can state that while MLP is more optimistic than RBF and C&RT, the three methods results are very similar. However, it was observed that ANNs classified the companies more accurately as the companies with good performance and poor performance. Considering error rates of the models, it was observed that MLP network had the lowest error in terms of general error rates. MLP was followed by RBF network and C&RT model respectively. There is a striking detail in Table 5 that Type I errors of the models are higher than Type II errors. In other words, the models failed to classify the companies with actually a poor performance as the ones with a poor performance accurately enough. However, the models classified the companies with actually a good performance as the ones with a good performance at a satisfactory level. Another striking finding was that Type II error of the MLP network was found to be 0. In other words, MLP network classified all of the companies with a good performance accurately.

7. CONCLUSION AND EVALUATION

The study compared prediction performance of different intelligent techniques using data of the largest 500 companies in Turkey in the 1993 and 2009. It was found that ANNs and decision trees had the ability to successfully model non-linear relationships. As indicated in Figure 4, the models achieved to classify the companies with an accuracy of more than 86%. Comparison of ANNs and decision trees showed that ANNs were slightly more successful than decision trees. It was also found that error rates of ANNs were lower than those of decision trees. The results of the study were found to be consistent with the finding in the literature pointing out to the superiority of ANNs over other methods. Our findings revealed that both intelligent techniques can be used in predicting financial performance and that if one had to be preferred, ANNs can make more accurate predictions. Future studies might make predictions by using different intelligent techniques.

Figure 4: Accuracy rates and errors of models



REFERENCES

[1]**Amari, S. I. (1977),** ***Neural Theory of Association and Concept-Formation.*** *Biological Cybernetics*, 261, 175-185;
[2]**Avcı, E., Çinko, M. (2008),** ***Forecasting the Index Returns with Artificial Neural Network Models: An Application in the Emerging European Stock Markets,*** *İktisat İşletme ve Finans*, 23 (266), 114-137;
[3]**Ban, Ü., Mazıbaş, M. (2009),** ***Bank Failure Prediction with Artificial Neural Networks: A Comparative Application to Turkish Banking System***, *İktisat İşletme ve Finans*, 24 (282), 27-53;
[4]**Beaumont N.B., Schroeder, R.M. (1997),** ***Technology, Manufacturing Performance and Business Performance amongst Australian Manufacturers,*** *Technovation*, 17 (6), 297-307;
[5]**Bishop, M. B. (1996),** ***Neural Networks for Pattern Recognition,*** *Oxford University Press, Englang;*
[6]**Boguslauskas V., Mileris R., Adlytė R. (2011),** ***The Selection of Financial Ratios as Independent Variables for Credit Risk Assessment,*** *Economics and Management*, 16, 1032-1038;
[7]**Burke, L. I. (1991),** ***Introduction to Artificial Neural Systems for Pattern Recognition,*** *Computers Operations Research*, 18 (2), 211-220;

[8]**Choudhary, M. A., Haider, A. (2012)*, Neural Network Models for Inflation Forecasting: An Appraisal,*** *Applied Economics*, 44 (20), 2631-2635;
[9]**Coats, P.K. and Fant, L.F. (1993)*, Recognizing Financial Distress Patterns Using a Neural Network Tool***, *Financial Management*, 22 (3), 142-156;
[10]**Costea A., Eklund T., Karlsson J., Voineagu V. (2009)*, Financial Performance Analysis of Scandinavian Telecommunication Companies Using Statistical and Neural Network Techniques,*** *Economic Computation and Economic Cybernetics Studies and Research, ASE Publishing;* 4, 87-104;
[11]**Delcea C., Simion C. (2011)*, The Incidence of Bankruptcy's Syndrome on Firm's Current Situation and Future Evolution,*** *Economic Computation and Economic Cybernetics Studies and Research, ASE Publishing,* 2, 137-152;
[12]**Düzakın, E., Düzakın, H. (2007)*, Measuring The Performance of Manufacturing Firms with Super Slacks Based Model of Data Envelopment Analysis: An Application of 500 Major Industrial Enterprises in Turkey,*** *European Journal of Operational Research*, 182, 1412-1432;
[13]**Ecer, F., Ulutagay, G., Nasiboglu, E. (2011)*, Does Foreign Ownership Affect Financial Performance? An Industrial Approach,*** *Middle Eastern Finance and Economics*, 14, 152-166;
[14]**Feldman, J. A., Ballard, D. H. (1982)*, Connectionist Models and Their Properties,*** *Cognitive Science*, 6, 205-254;
[15]**Grossberg, S. (1988)*, Nonlinear Neural Networks: Principles, Mechanism, and Architectures,*** *Neural Networks*, 1, 17-61;
[16]**Güleş, H. K., Türkmen, M., Özilhan D. (2011)*, Tekstil ve Otomotiv Yan Sanayi Küçük ve Orta Ölçekli İşletmelerinde Üretim ve İşletme Performansı,*** *İktisat, İşletme ve Finans*, 26 (304), 79-112;
[17]**Guresen, E., Kayakutlu, G., Daim, T. U. (2011)*, Using Artificial Neural Network Models in Stock Market Index Prediction,*** *Expert Systems with Applications*, 38 (8), 10389-10397;
[18]**Hafeez, K., Zhang, Y. B., Abdelmeguid, H., Malak, N., Iqbal, S. (2000)*, Firm Competence Evaluation Framework using AHP,*** *First International Conference on Systems Thinking in Management*, 225-230;
[19]**Hagan, M. T., Demuth, H. B., Beale, M. (1995)*, Neural Network Design,*** *PWS Publishing Company, Boston, USA;*
[20]**Han, Y., Wang, B. (2011)*, Investigation of Listed Companies Credit Risk Assessment Based on Different Learning Schemes of BP Neural Network,*** *International Journal of Business & Management*, 6 (2), 204-207;
[21]**Ince, H. (2006)*, Portfolio Selection with Artificial Neural Networks,*** *İktisat İşletme ve Finans*, 21 (241), 114-127;
[22]**İstanbul Sanayi Odası, *Turkey's Top 500 Industrial Enterprices*** *http://www.iso.org.tr/tr/web/StatikSayfalar/Besyuz_Buyuk_Gecmis_Yillar.aspx, Assessed on 18th May 2012;*

[23]**Kim, J. S., Arnold, P. (1993),** ***Manufacturing Competence and Business Performance: A Framework and Empirical Analysis,*** *International Journal of Operations and Production Management*, 13, 10, 4-25;
[24]**Korhonen, T. (1984),** ***Self-Organization and Associative Memory,*** *Springer Verlag, New York, USA;*
[25]**Kotha, S., Swamidass, P. M. (2000),** ***Strategy, Advanced Manufacturing Technology and Performance: Empirical Evidence from U.S. Manufacturing Firms,*** *Journal of Operations Management*, 18, 257-277;
[26]**Krishna Kumar, M. S., Subramanian, S., Rao, U. S. (2010),** ***Enhancing Stock Selection in Indian Stock Market Using Value Investment Criteria: An Application of Artificial Neural Networks***, *IUP Journal of Accounting Research & Audit Practices*, 9 (4), 54-67;
[27]**Kuan, C.-M., White, H. (1994),** ***Artificial Neural Networks: An Econometric Perspective,*** *Econometric Reviews*, 13, 1-91;
[28]**Kumar, M. (2009),** ***Nonlinear Prediction Of The Standard & Poor's 500 And The Hang Seng Index Under A Dynamic Increasing Sample,*** *Asian Academy of Management Journal of Accounting & Finance*, 5 (2), 101-118;
[29]**Lippmann, R. P. (1987),** ***An Introduction to Computing with Neural Nets,*** *IEEE ASSP Magazine* (April);
[30]**Merh, N., Saxena, V. P., Pardasani, K. R. (2011),** ***Next Day Stock Market Forecasting: An Application of ANN and ARIMA,*** *The IUP Journal of Applied Finance*, 17 (1), 70-84;
[31]**Minsky, F., Papert S. (1988),** ***Perceptrons,*** *MIT Press, Cambridge;*
[32]**Moshiri, S., Cameron, N. E., & Scuse, D. (1999),** ***Static, Dynamic and Hybrid Neural Networks in Forecasting Inflation,*** *Computational Economics*, 14, 219-235;
[33]**Rumelhart, D. E., Hinton, G. and Williams, R. (1986),** ***Learning Representation by Back-Propagating Errors,*** *Nature*, 323, 9, 533-536;
[34]**Sun, H., Hong, C. (2002),** ***The Alignment between Manufacturing and Business Strategies: Its Influence on Business Performance,*** *Technovation*, 22, 699-705;
[35]**Wasserman, P. D., Oetzel, R. M. (1990),** ***Neural Source: The Bibliographic Guide to Artificial Neural Networks,*** *Van Nostrand Reinhold, New York;*
[36]**Youn H. and Gu Z. (2010),** ***Predicting Korean Lodging Firm Failures: An Artificial Neural Network Model along with a Logistic Regression Model,*** *International Journal of Hospitality Management*, 29, 120-127;
[37]**Yudong, Z., Lenan, W. (2009),** ***Stock Market Prediction of S&P 500 via Combination of Improved BCO Approach and BP Neural Network,*** *Expert Systems with Applications*, 36, 8849-8854.