DETECTING THE MANIPULATION OF FINANCIAL INFORMATION BY USING ARTIFICIAL NEURAL NETWORK MODELS

Güray KÜÇÜKKOCAOĞLU[1]
Yasemin KESKİN BENLİ[2]
Cemal KÜÇÜKSÖZEN[3]

ABSTRACT

Despite their widespread usage, models of accrual based methods in detecting false financial statements have been subject to significant criticism. An alternative to the accruals approach is to use binary probit and logit models and some other multivariate statistical techniques where they combine accruals and some other financial ratios and/or indexes. The objective of this paper is to explain the historical evolution of the accrual based methods where they provide some evidence of earnings management practices and than extend to some other alternative methods in detecting manipulative practices in financial reporting. This paper also, introduces a new method that has been widely used in detecting financial distress companies. An Artificial Neural Network Model, which is based on the concept of using artificial neurons, to estimate the manipulative financial reporting practices of the companies listed in the Istanbul Stock Exchange (ISE). The results indicate that the proposed Artificial Neural Network Model outperforms the traditional statistical techniques used in earnings manipulation practices.

Key words: Earnings Management, Financial Ratios, Artificial Neural Network Model.

1. Introduction

It is quite hard to detect manipulation of financial information from publicly available financial statements. Academicians, who have less opportunity to access companies' information and have less authority as compared to regulatory bodies, are trying to generate some models to detect companies exercising earnings management practices[4]. In the earnings management literature, these models have a methodology through which falsified financial statements are found out as classifying indicators on the manipulation of financial information. In this context, first, Healy (1985) presented a model to the literature which is evaluating the effects of managements' nondiscretionary accruals and accounting policy changes. Following the work done by Healy (1985), models focus on the aggregate accruals to evaluate nondiscretionary accruals and try to estimate aggregate accruals which are determined as the difference between publicly available net income and net cash flow from operations. Estimated aggregate accruals are used in regression analyses with the variables like income (or cash receipts from customers), indicating working capital needs (i.e. receivables, inventory, trade credits), and gross tangible assets, indicating normal

[1] Başkent University, Faculty of Economics and Administrative Sciences, Management Department B219, Bağlıca Kampüsü, 06530, Ankara – Turkey + 90 (312) 234 10 10 / 1728, gurayk@baskent.edu.tr
[2] Gazi University, Industrial Arts Educational Faculty, Department of Management Education.
[3] Capital Markets Board of Turkey, Head of Accounting Standarts Department.
[4] The terms "earnings management", "manipulation of financial information", "earnings manipulation", "fraud", "falsified financial statements" are all used interchangeably to describe the act of intentional misrepresentation of a firm's financial statement by managers.

depreciation. According to Jones (1991), in a regression analyses, unexpected accruals constitute the unexplained part of the aggregate accruals (Küçüksözen and Küçükkocaoğlu, 2005).

As mentioned above, in the models which are designed to detect manipulation of financial information (or which are used to predict the possible act of earnings management practices), primarily, aggregate accruals are focused. In some studies, unexpected aggregate accruals (i.e. accrual amounts excessing the requirements of companies' activities) are estimated through indexing accrual amounts to the total assets or income amounts directly and abnormal accrual amounts accumulating in years are considered as indicator of earnings management practices.

In some studies, aggregate accruals are segregated into two parts one of which consists of accruals required by activities (i.e. nondiscretionary accruals) and the other one consists of accruals not required by activities (i.e. discretionary accruals) are indexed to companies' total assets or sales. Yearly trends in these indices are considered as the indicators of various practices of manipulation.

After the accrual based studies, ongoing studies in the literature are based on Logit and Probit models. It is observed that after the Beneish's (1997) first probit study on earnings management practices, different models with different calculations are presented to detect the earnings management practices on financial information. To contribute to this field of study, we have applied the Artificial Neural Network model along with the Beneish's indices on the companies listed in the Istanbul Stock Exchange (ISE).

To the best of our knowledge, first study using neural networks to determine manipulation of financial information was done by Fanning, Cogger and Srivastava (1995). Later on, Fanning and Cogger (1998) demonstrated that 8 out of 20 variables (trade receivables/sales, trade receivables/total assets, inventory/sales, tangible assets/total assets, total liabilities/equity, sales/total assets) have reasonable explanatory power in determining manipulation of financial information through changing the extent of the database which was used in their first study. As will be explained in the following parts of this study, a study of Küçüksözen and Küçükkocaoğlu (2005), in which similar ratios were tested on ISE, and a study of Spathis (2002), in which similar ratios were tested on Athens Stock Exchange, showed that this kind of ratios are not sufficient enough to explain the manipulation of financial information. Contrarily, according to these studies, it is observed that the indices which were suggested by Beneish (1997) have more reasonable explanatory power.

In the following parts of this study, the models which focus on accruals as determining the manipulation of financial information will be mentioned. After that, various models which include alternative approaches and include analyses based on the information derived from financial statements will be presented. In section three, we applied neural networks model to determine earnings management practices through an empirical study. In the fourth and fifth section, results of this study will be summarized.

2. Models Related to Manipulation of Financial Information

This section of the study will focus on the accrual based models which have been first presented into academic literature by Healy in 1985 and have been later on become the research study of many academicians. In addition to Healy's study, models whose objectives are to predict manipulation of financial information through financial ratios and indexes, with alternative approaches, will be presented.

2.1. Accrual Based Models

The accrual based models which have been first introduced into academic literature by Healy (1985) and later improved by DeAngelo (1986), and Jones (1991), used under different calculation methods, along with different names. This section of the study focuses on the accrual based models which have been started by Healy in 1985 and improved by other academicians.

2.1.1. Healy Model

Healy, in his study of 1985, created the hypothesis that managers who get bonus schemes via premiums manipulate financial information by using total accruals in order to increase the amount of their incentives. Healy tested this hypothesis using the following model.

$NDA_t = 1/n\Sigma_\tau(TA_\tau / A_{\tau-1})$

NDA = Nondiscretionary accruals
TA = Total accruals
A = Total assets

2.1.2. DeAngelo Model

In 1986, DeAngelo studied the hypothesis that managers manipulate financial information in order to show the value of shares understated at the time when a publicly held company will be converted to a private company via buying out all the company's stocks from investors. DeAngelo tested this hypothesis using the following model.

$NDA = TA_{t-1}/ A_{t-2}$

NDA = Nondiscretionary accruals
TA = Total accruals
A = Total assets

2.1.3. Jones Model

Jones (1991), tested whether American companies understated their profit via manipulation of financial information, in order to benefit from the protection of customs such as increase in customs tariffs or restriction of quotas within the sector that the company belongs, at the time that USA Commerce Commission examined their records.

$TA_{it} / A_{it-1} = \alpha_i [1/A_{it-1}] + \beta_{1i} [\Delta REV/A_{it-1}] + \beta_{2i} [PPE/A_{it-1}] + \varepsilon_{it}$

TA = Total accruals,
A = Total assets,
ΔREV = Change in Revenues,
PPE = Gross Plant, Property and Equipment

2.1.4. Modified Jones Model

According to Jones (1991) model, not only in the period of manipulation of financial information but also in the period of prediction, the indifference between the decision of nondiscretionary accruals and sales revenue is assumed. According to the study of Dechow, Sloan and Sweeney (1995), this model measures nondiscretionary accruals with errors; therefore, this assumption leads to problems in calculating nondiscretionary accruals. In this regard; instead of using only change in income, the use of change in income will be used by subtracting net change in receivables (current year's receivables minus previous year's receivables). In other words, change in income will be adjusted by taking the change in receivables into consideration. In this context, in modified Jones model, it is implicitly assumed that changes in the amount of sales on credit are generated from the manipulation of financial information. This assumption is based on the acceptance that the use of implicit rights in defining income generated from sales on credit can be implemented easily compared to defining income in cash sales as well as manipulation of financial information can be realized easily via sales on credit (Küçüksözen and Küçükkocaoğlu, 2005).

$$NDA_t = \alpha_1(1 / TA_{t-1}) + \alpha_2[(\Delta REV_t - \Delta REC_t) / TA_{t-1})] + \alpha_3(PPE_t / TA_{t-1})$$

NDA = Nondiscretionary accruals,
TA = Total Assets,
ΔREV = Change in Revenues,
ΔREC = Change in Receivables,
PPE = Gross Plant, Property and Equipment

2.1.5. Industry Model

In parallel to the Jones (1991) model, industrial model loosen the assumption that discretionary accruals are constant. Instead of directly modeling the determinants of nondiscretionary accruals, this model works via the assumption that the change in these determinants are the same in the same sector that all companies belong. This method is based on the use of median values of total accruals calculated through the scaled asset sizes of companies which belong to the same sector, except for the exemplary companies that have been examined.

$$NDA_t = \beta_1 + \beta_2 median_j (TA_t / TA_{t-1})$$

NDA = Nondiscretionary accruals
TA = Total Assets

Dechow, Sloan and Sweeney (1995), tested all the accrual based models leading to the manipulation of financial information. According to the results of their study, Modified Jones model is the strongest model in determining the manipulation of financial information compared to other models above. (Küçüksözen and Küçükkocaoğlu, 2005).

2.2. Mixed Models

The following mixed models which also include total accruals, try to predict earnings management practices via converting financial statement figures into financial ratios and indexes.

2.2.1. Logit and Probit Models

A new point of view has been improved for the determination of companies that manipulate financial information, especially with the innovation of Beneish in 1997. In addition to the use of linear regression in order to determine the change in accruals, Beneish emphasized that probit and logit models which focus on other variables can be used to determine the companies manipulated their financial information. Therefore, Beneish contributed to the earnings management literature in 1997 and 1999.

2.2.1.1. Beneish Model (Probit Model)

In his modeling studies at 1997 and 1997, Beneish emphasize his idea that companies that benefit from the manipulation of financial information do not always use accruals, aggressively. As well as he focuses on the idea that different variables should be used in order to determine the manipulation of financial information in companies. These variables are based on the information located on the financial statements and include characteristics that may help to determine the manipulation of financial information created by companies. Also, they are important to determine whether companies have any transactions contrary to the generally accepted accounting principles.

In his models, the data in the sense of explanatory variables which belong to both companies that manipulate financial information and control companies that are assumed not to manipulate financial information are all part of the probit analysis. Probit analysis is a method of regression analysis that is convenient for dependent variables (M_i; dual variable; value is 1 for manipulators, 0 for control companies) used in the below equation.

In his modeling studies at 1997 and 1999, Beneish find out coefficients for each variable by making probit analysis of the data of control companies as well as companies that manipulate financial information. By using these coefficients, it is possible to calculate whether each company manipulates financial information or not. In this context, if the result of M_i is close to zero, the firm is not a manipulator. Or, if M_i is close to 1, it is a manipulator.

Within the framework of Beneish Model (1997), (1999) (Probit Model)

$M_i = \beta^i X_i + \epsilon_i$

> Mi= Dummy variable; for companies that manipulate financial information value = 1, companies that do not manipulate financial information value = 0,
> β^i = Coefficient for each independent variable within the framework of the model
> X_i = Matrix which constitute explanatory variables,
> ϵ_i = Error term

Some of the major explanatory (independent) variables in the model;

- Days Sales in Receivables Index (DSRI)
- Gross Margin Index (GMI)
- Asset Quality Index (AQI)
- Depreciation Index (DEPI)
- SG&A Index (SGAI)

- Working Capital Accruals to Total Assets (TATA)
- Sales Growth Index (SGI)
- Days in Inventory Index (DII)
- Abnormal return in stock prices

According to Dechow, Sloan and Sweeney (1996), Beneish model provided the users of financial statements the opportunity to evaluate companies from different aspects by taking the picture of those companies' financial situation and financial performance in addition to the model of Jones (1991) which focuses on manipulation of financial information through receivables. Also, variables used in this model are related not only to the determination of manipulative transactions which took place within the company but also to the determination of intent for future manipulative transactions within the company.

On the other hand, according to Beneish (1997), his model makes Jones model (1991) stronger. With in this context, this model indicates correctly the implementation of financial information manipulation within the companies that use nondiscretionary accruals at great amounts. In this regard, nondiscretionary accruals can take place for the manipulation of financial information as well as for the decision companies towards their strategic goals within the framework of operating activities.

2.2.1.2. Spathis Model (Logit Model)

Different from indexes used in the probit model of Beneish in 1997 and 1999, Spathis focuses on financial ratios in this study in 2002. Instead of probit regression, he implemented logistic regression in his analysis. Accordingly the model which was created by Spathis in 2002 based on the equation below, his model makes logistic regression analysis to control companies and companies that manipulate financial information according to the independent variables.

$$E(y) = \frac{\exp(b_0 + b_1 x_1 + b_2 x_2 + ...b_n x_n)}{1 + \exp(b_0 + b_1 x_1 + b_2 x_2 + ...b_n x_n)}$$

In this equation; dependent variable E(y) is equal to 1 for the companies that disclosed false financial information and is equal to 0 for control companies. b_0 shows the value of intersection. $b_1, b_2, \ldots b_n$ constitute the coefficients of independent variables. $x_1, x_2, \ldots x_n$ indicate independent variables below.

FFS = b_0 + b_1(D/E) + b_2(Sales/TA) + b_3(NP/Sales) + b_4(Rec/Sales) + b_5(NP/TA) + b_6(WC/TA) + b_7(GP/TA) + b_8(INV/Sales) + b_9(TD/TA) + b_{10}(FE/GE) + b_{11}(Taxes/Sales) + b_{12} (Altman Z-score)

- Debt/Equity (D/E),
- Sales/Total assets (Sales/TA),
- Net profit/Sales (NP/Sales),
- Receivable/Sales (Rec/Sales),
- Net profit/Total assets (NP/TA),
- Working capital/total assets (WC/TA),
- Gross profit/Total assets (GP/TA),
- Inventories/Total assets (INV/Sales),
- Total debt/Total assets (TD/TA),

- Financial expenses/Operational expenses (FE/GE),
- Taxes/Sales (Taxes/Sales)
- Altman Z-score (Z-score),

In 2000, Spathis made logistic regression analysis of 76 companies using their financial statements at the Athens Stock Exchange. He tried to figure out the ratios to find out the financial statements which do not reflect the reality by using some of the value on the financial statements. According to his analysis, the following ratios have explanatory power in the detection of earnings management practices; Inventories to sales (INV/Sales), Total debt to total assets (TD/TA), and Altman Z Score.

According to a similar study conducted by Küçüksözen and Küçükkocaoğlu (2005) at the Istanbul Stock Exchange listed companies, Net income to total assets and Financing expenses to total operating expenses have some have explanatory power in the detection of earnings management practices.

2.2.2. Multivariate, Multi-Criteria Models

UTADIS methodology, which is being used in financial management, default prevision, credit risk analyses, calculations of country risk, portfolio choice and management etc., was used in determining manipulation of financial information by Spathis, Doumpos and Zopounidis (2004). Spathis, Doumpos and Zopounidis (2004), used the variables which were presented in Spathis' (2002) Logit Model and established a difference curve function, and classified companies as manipulators and non-manipulators through the determination of difference curve function's upper and lower limits. Even though they claim that their study has a %100 success, we believe that the conclusions of their study have a misleading structure due to their methodology, insufficient structure and extent of the database they used in the calculations.

3. Detection of False Financial Statements through Neural Network Models

Rapid advances in computer technology have enabled very complicated calculations to be performed in an instant. Nevertheless, processes like hand-writing, speech, and visual recognition remain a difficult challenge for computers. This challenge has led scientists to develop alternative data processing systems which differ from the classical approaches used by computer. One of the first steps taken on this issue has been trying to benefit from some of the biological findings related to the operation of the human brain. The structural and operational characteristics of the neural networks inside the human brain are much too complex to be duplicated in terms sufficiently simple to facilitate a useful mathematical model. Studies of the neurophysiologists and psychologists have, however, been helpful in this regard. These types of mathematical models are named neural networks (Sungur, 1995).

Neural network refers to an artificial intelligence technology. It can produce successful results when multiple variables and complicated mutual interactions are present or where multiple solution groups obtain. Because of such features, artificial neural network technology is considered an appropriate means to be used in the financial failure field (Salchenberger, Çınar and Lash 1992; Wilson and Chong 1995; Koh and Tan 1999; Yıldız, 2001).

According to the neurophysiologists and psychologists, artificial neural networks are designed to explain the human brain's functions. On the engineers' side, the artificial neural networks are, before all else, alternative means of performing calculations. However, the bond between these two research motivations is strong. While the neurophysiologic findings constitute a source of inspiration for developing new mathematical models, the results of the studies and implementations made from such developed mathematical models have the ability to direct the neurophysiologic researches.

Robert Hecht-Nelson (1989), the first person to develop commercial artificial neural networks, identifies the artificial neural network as " a data processing system which processes the data by dynamically creating answers to the inputs delivered from outside, which is composed of simple elements associated with each other" (Yıldız, 2001). In another words artificial neural networks can be defined as a parallel and dispersed single or multi layered data processing system which is composed of many simple processor elements (artificial neurons). Each of these neurons has its own memory, capable of transaction, and communicates with other neurons through one-way signal channels (Gülseçen, 1995).

The structure of an artificial neural network, contains three main layers: the input layer where the interconnected nerves are present, the output layer, and the hidden layer.

The first (input) layer enables the intake of the exterior data into the artificial neural network. These exterior data are equivalent to the independent variables in statistics. The last (output) layer's function is to transmit the data out. Output variables are equivalent to the dependent variables in statistics. The other (hidden) layer inside the model is located between the input layer and the output layer. The nerves inside the hidden layer have no attachments to the exterior environment. They only receive the signals coming from the input layer and transmit signals to the output layer.

The determination of the number of hidden neurons inside the hidden layer is very important. The definition of the size of the network is essential for assessing the performance of the network. Increasing or decreasing the number of the hidden neurons and layers, affects the structure of the network as being complicated or simple.

One of the most important elements inside the artificial neural network is the connection between the neurons enabling them to transmit data to each other. A connection, which enables data transfer from any (i) neuron to any other (j) neuron, has a weighted value (w_{ji}). Weights reflect the relative force that is used as an input inside a neuron. Inside an artificial neural network, each connection has a different weighted value. In this way the weights affect each input of each processor element (Yıldız, 2001).

In Figure 1, inside an artificial neural network structure, the inputs are labeled X, outputs received from the hidden layer are labeled h, and the outputs obtained at the end are labeled Y (Güneri, 2001).

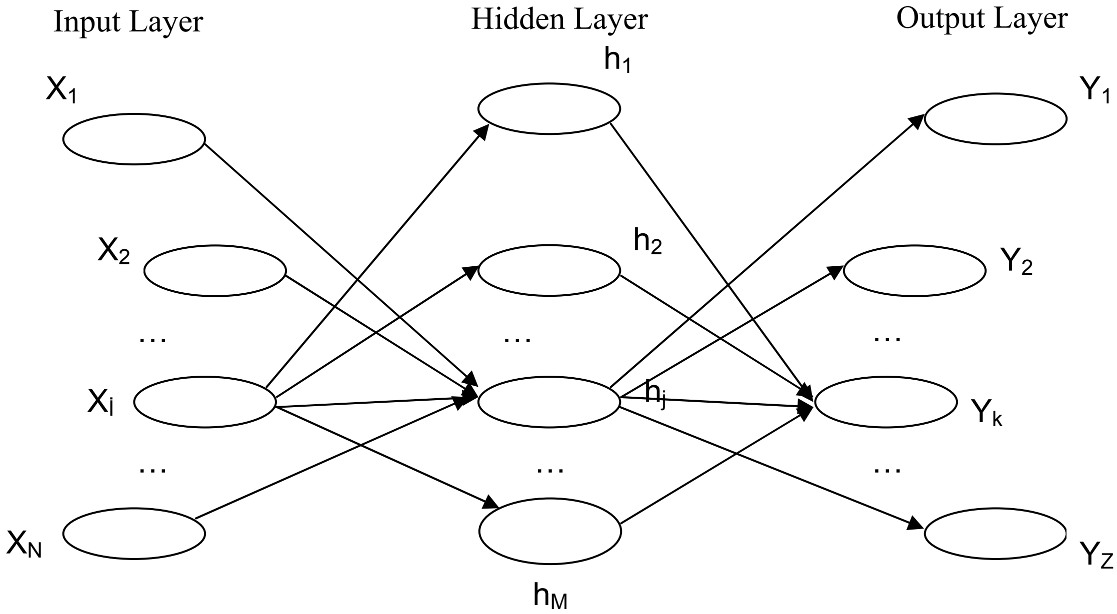

Figure 1: Structure of the Artificial Neural Network

In the artificial neural network model, the weighted values of the connections between the nerves are produced randomly by the SPSS package program. The network is tested by using these values.

The data inside the data set are randomly separated into three parts: the training, validity and test sets. The training set data are used for training the network. The validity set, is used in accordance with the weights of a classifier. The validity set is used to select the number of hidden units inside an artificial neural network. The test set is used for evaluating the performance of the training. The data are allocated 80% to the training set, 10% to the validity set and 10% to the test set.

In order to remove the effect of the measurement unit, the data are standardized such that each data point contributes to the decisions or recommendations equally. The package program used standardizes the data initially. Then, the transition function is (in this study, the sigmoid function) is selected.

The difference between the actual output values and desired output values is measured and depending upon the result, the connection weights of the network model are modified. The return passage resulting from the weights of the connections is realized by the production of the network which starts with the connections of the output layers and ends with the connections of the input layers.

The number of nerves inside a layer can be automatically determined by the networks or can be arranged to be interconnected. In many cases, increasing the number of nerves develops the performance of the multiple layer networks on the training data.

The performance of the validity data is checked in order to evaluate the effect of the number of hidden layers inside a problem. The mean absolute error (MAE) and root mean square error (RMSE) are used to determine the performance of the network structure. The value where the mean absolute error and root mean square error are minimum determines the number of hidden layers to be used. The number of hidden layers was determined to be six to mean absolute error (MAE), four to root mean square error (RMSE).

Threshold value on artificial neural network application is determined 0.50. If the estimation value is greater than the threshold value, then the related companies considered as manipulative financial reporting companies, If smaller then considered as non-manipulative financial reporting companies. After this, these classification values are compared with the actual values, and then correct classification value has been calculated.

For the training of the network, 10.000 iterations were realized. At the end of artificial neural network analysis, correct classification tables for training, validity and test sets were obtained.

3.1. Companies Included In the Analysis

In our study, non-financial 126 Istanbul Stock Exchange listed companies were chosen. Banks, insurance companies and other financial companies were excluded as in other studies. The chosen companies' balance sheets and income statements corresponding to the years 1992-2002 were analyzed in the context of our study.

Capital Markets Board of Turkey (CMBT), investigated these 126 companies' financial statements corresponding to the years 1992-2002, and detected earnings management practices in 168 observations and no signs of earnings management practices in 1.040 observations. To find out these investigations conducted by CMBT, where manipulation of financial information existed and the cases where manipulation of financial information did not exist, ISE's daily bulletins corresponding the dates between 01.01.1992 - 31.07.2004 and CMBT's weekly bulletins corresponding to the dates between 01.01.1996-31.07.2004 were investigated using some key words (financial statements, balance sheet, income statement, profit, loss, income, cost, audit report, capitalization, and restatement). According to the information gathered from these bulletins, the companies which were determined and announced to the public as having exercised manipulation of financial information as a result of the CMBT's investigations and/or the ones which received qualified audit opinion changing the values in their financial statements about their publicly available financial statements or the companies which had changed the values in their financial statements after balance sheet date were considered as the companies which exercised manipulation of financial information. In addition, the companies which had changed the information in their financial statements prepared to be registered with the Board during the investigations done by CMBT were also considered as the companies which exercised manipulation of financial information.

In the 1.040 observations which were set as control group, there might be companies which exercised manipulation of financial information but the fact is not publicly available because they may not been investigated by CMBT or the fact that they had not been closely examined by their auditors. Nonetheless, this condition is considered as type 1 error as in all other studies.

3.2. Definition of Variables and Data Resources

Most of the following explanatory (independent) variables in the model were picked from Beneish's studies (1997), (1999)[5];

i. Sales Growth Index (SGI)
ii. Days Sales in Receivables Index (DSRI)
iii. Gross Margin Index (GMI)
iv. Asset Quality Index (AQI)
v. Depreciation Index (DEPI)
vi. Sales, General and Administrative Expenses Index (SGAI)
vii. Leverage Index (LVI)
viii. Total Accruals to Total Assets Index (TATA)
ix. Days in Inventory Index (DII)
x. Financial Expenses Index (FEI)

As we have mentioned in previous paragraphs, Beneish in 1997 and 1999, pointed out that companies that benefit from the manipulation of financial information do not always constitute the companies which use accruals aggressively. Beneish emphasized the idea that different variables are required to estimate manipulation on financial information. These variables have necessary characteristics not only to determine companies' manipulations realized by using financial statements but also to figure out any illegal transactions of companies based on generally accepted accounting principles.

In this context, the independent variables that have been chosen for our study are mainly the same ones used by Beneish in 1997 and 1999. In addition to these variables, the following variables adopted to the artificial neural network model has been created by Küçüksözen and Küçükkocaoğlu in 2005; the proportion of inventory to sales (INV/Sales) and the proportion of financial expenses to sales (FEI/Sales), are given emphasis on this study as well.

For the analysis of artificial neural network, it was benefited from Neural Connection programming.

3.3. The Empirical Results

In the artificial neural network analysis, the model of the problem includes 10 input layers, due to 10 independent variables. In other words, the model is constituted by the indexes and the financial ratios used in the definition of the variables and data resources, so in the input layer ten neurons exist. In the output layer, due to the firms' classification that whether or not making manipulation of financial information, there is an output layer. Thus, only one neuron exists.

To be able to define the numbers of hidden layers, firstly, the hidden layer number is taken 1 and the errors are calculated for the 10-1-1 model. After these calculations, the number of hidden layers is increased and the mean square error (MSE) and the mean absolute error (MAE) of the data validity are calculated. The results are shown in Table 1.

---

[5] The explanation based on the calculation of indexes and proportions are given in the Appendix of our study. The data set and explanations related to 10 independent variables determined for our study are taken from the studies of Küçüksözen (2005) and Küçüksözen and Küçükkocaoğlu (2005).

Table 1. The Results of the Models Used In the Determination of the Number of the Hidden Layers

Model	MSE	MAE
10-1-1	0.305868	0.208335
10-2-1	0.307067	0.211865
10-3-1	0.306856	0.211334
10-4-1	**0.304201**	0.210488
10-5-1	0.307615	0.208804
10-6-1	0.304297	**0.206663**
10-7-1	0.307096	0.209142

Analyzing the Table 1, we can see the mean square error of the one hidden layer model (10-1-1) is 0.305868. While increasing the number of hidden layers, mean square error is increasing after the forth hidden layer. Thus, 4-hidden layer model (10-4-1) is chosen to constitute the prediction model due to its lowest mean square error.

Looking at the mean absolute error, we can see the mean absolute error of the 1 hidden layer model (10-1-1) is 0.208335. The 6 - hidden layer model (10-6-1 model) is chosen because it has the lowest mean absolute error.

As increasing the number of hidden layers, the network performance is also increasing because it shows the new characteristics of the every new hidden layer data set. It is possible to see a decrease in adding too much layer. The reason of this is the loss in the general power and the noise is going to be learned from the network data. Making error measurement from the validity data, we can decrease the danger of excess learning (Neural Connection, 1997; Güneri, 2001).

For the training of the network, 10.000 iterations are realized. The 80% of the data is training set, the 10% is validity set, and the 10% is test set. Thus, 966 data belongs to training set, 121 data belongs to validity set, and 121 data belongs to test set. To constitute the prediction model, 4- hidden layer model (10-4-1) and 6- hidden layer model (10-6-1) are treated for training, validity and test set, and the classification tables are determined. The results are shown in the tables below.

3.3.1 The Results of Four- Hidden Layer Model

Table 2. Classification of the Model for the Training Set (10-4-1)

Actual \ Prediction	non manipulative financial reporting companies (0)	manipulative financial reporting companies (1)	Total
non manipulative financial reporting companies (0)	**831**	**1**	**832**
manipulative financial reporting companies (1)	**132**	**2**	**134**
Total	**963**	**3**	**966**

Correct classification percentage for the training set is calculated as 86.231888. Incorrect classification percentage is 13.768116.

Table 3. Classification of the Model for the Validity Set (10-4-1)

Actual \ Prediction	non manipulative financial reporting companies (0)	manipulative financial reporting companies (1)	Total
non manipulative financial reporting companies (0)	**108**	**0**	**108**
manipulative financial reporting companies (1)	**13**	**0**	**13**
Total	**121**	**0**	**121**

Correct classification percentage for the validity set is calculated as 89.256195. Incorrect classification percentage is 10.743802.

Table 4. Classification of the Model for the Test Set
(10-4-1)

Actual \ Prediction	non manipulative financial reporting companies (0)	manipulative financial reporting companies (1)	Total
non manipulative financial reporting companies (0)	**100**	**0**	**100**
manipulative financial reporting companies (1)	**21**	**0**	**21**
Total	**121**	**0**	**121**

Correct classification percentage for the test set is calculated as 82.644630. Incorrect classification percentage is 17.355371.

When we combine the results for training, validity and test set, we have Table 5 according to artificial neural network applications.

Table 5. Classification of the Model According To Artificial Neural Network Applications
(10-4-1)

Actual \ Prediction	non manipulative financial reporting companies (0)	manipulative financial reporting companies (1)	Total
non manipulative financial reporting companies (0)	**1039**	**1**	**1040**
manipulative financial reporting companies (1)	**166**	**2**	**168**
Total	**1205**	**3**	**1208**

The general correction percentage according to artificial neural network applications is found as 86.175496. Incorrect classification percentage is 13.824503.

When we want to predict future situation of a new company added to the model, according to the artificial neural network application, the probability of prediction being correct is 86.175496 %.

3.3.2. The Results of Six-Hidden Layer Model

Table 6. Classification of the Model for the Training Set (10-6-1)

Actual \ Prediction	non manipulative financial reporting companies (0)	manipulative financial reporting companies (1)	Total
non manipulative financial reporting companies (0)	**832**	**0**	**832**
manipulative financial reporting companies (1)	**134**	**0**	**134**
Total	**966**	**0**	**966**

Correct classification percentage for the training set is calculated as 86.128365. Incorrect classification percentage is 13.871635.

Table 7. Classification of the Model for the Validity Set (10-6-1)

Actual \ Prediction	non manipulative financial reporting companies (0)	manipulative financial reporting companies (1)	Total
non manipulative financial reporting companies (0)	**108**	**0**	**108**
manipulative financial reporting companies (1)	**13**	**0**	**13**
Total	**121**	**0**	**121**

Correct classification percentage for the validity set is calculated as 89.256195. Incorrect classification percentage is 10.743802.

Table 8. Classification of the Model for the Test Set
(10-6-1)

Actual \ Prediction	non manipulative financial reporting companies (0)	manipulative financial reporting companies (1)	Total
non manipulative financial reporting companies (0)	**100**	**0**	**100**
manipulative financial reporting companies (1)	**21**	**0**	**21**
Total	**121**	**0**	**121**

Correct classification percentage for the test set is calculated as 82.64463. Incorrect classification percentage is 17.355371.

When we combine the results for training, validity and test set, we have Table 9 according to artificial neural network applications.

Table 9. Classification of the Model According to Artificial Neural Network Applications
(10-6-1)

Actual \ Prediction	non manipulative financial reporting companies (0)	manipulative financial reporting companies (1)	Total
non manipulative financial reporting companies (0)	**1040**	**0**	**1040**
manipulative financial reporting companies (1)	**168**	**0**	**168**
Total	**1208**	**0**	**1208**

The general correction percentage according to artificial neural network applications is found as 86.092715. Incorrect classification percentage is 13.907284.

When we want to predict future situation of a new company added to the model, according to the artificial neural network application, the probability of prediction being correct is 86.092715 %.

Both the results of 4-hidden layer model and 6-hidden layer model are summarized in Table 10.

Table 10. Four-Hidden Layer and Six- Hidden Layer Artificial Neural Network Models

10-4-1 (General Correction %)	10-6-1 (General Correction %)
86.175496	86.092715
10-4-1 (Incorrect Classification %)	10-6-1 (Incorrect Classification %)
13.824503	13.907284

The predicting power of a 4- hidden layer model (% 86.175496) is higher than the predicting power of 6-hidden layer model (% 86.092715), and incorrect classification percentage of the 4-hidden layer model (%13.824503) is lower than 6-hidden layer model. As a result of this, 4-hidden layer model is chosen as an artificial neural network model.

4. Conclusions

The models presented in the literature, try to determine manipulated financial information through separating the companies into two groups one of which contains companies having exercised manipulation of financial information and the other one of which contains companies not having exercised manipulation of financial information.

With in this context, Küçüksözen and Küçükkocaoğlu's probit model (2005), which is based on Beneish Model and uses the database that is included in this study, is designed to estimate the probability of financial information based manipulation per company through calculating the values related with the independent variables corresponding to the years between 1993-2002 for each 126 listed companies and applying these values to the equation which is generated using data of year 1997. In fact, according to the results of probit model, Küçüksözen and Küçükkocaoğlu could estimate %38 (range between %33-57) of the companies having exercised manipulation of financial information and %61 (range between %43-74) of the companies not having exercised manipulation of financial information (control group). These ratios are similar to the ratios that exist in Beneish's (1999) study, especially the ones for the companies having exercised manipulation of financial information.[6]

In this study, the companies having exercised manipulation of financial information and the companies not having exercised manipulation of financial information were separated using an artificial neural network model. According to our findings, the probability of true prediction could be %86.17 and the probability of false prediction could be %13.82. Even though these findings cannot be compared with the ones of probit model, we think that artificial neural network model enables us estimating true classification with a high probability. So we think that this model should be taken into consideration when detecting companies which exercise manipulation of financial information.

In conclusion, when the variables that are necessary to find out manipulation of financial information are known, artificial neural network approach could be used for determining the companies which will exercise manipulation of financial information.

[6] Beneish's model was estimated %37-56,1 of the companies having exercised financial information based as manipulator. This ratio is %80-92 for the control group.

REFERENCES

Beneish, M. D., (1997), "Detecting GAAP Violation: Implications for Assessing Earnings Management Among Firms with Extreme Financial Performance", Journal of Accounting and Public Policy, Vol:16, No: 3, pp: 271-309.

Beneish, M. D., (1999), "The Detection of Earnings Manipulation", Financial Analysts Journal, Vol:55, No:5, pp:24-36.

DeAngelo, L.E. (1986), "Accounting Numbers as Market Valuation Substitutes: A Study of Management Buyouts of Public Stockholders", The Accounting Review, Vol:61, No:3, pp. 400-420.

Dechhow, P.M., Sloan, R.G., Sweeney, A.P., (1995), "Detecting Earnings Management", The Accounting Review, Vol:70, No:2, pp: 193-225.

Healy, P.M., (1985), "The Effect of Bonus Schemes on Accounting Decisions," Journal of Accounting and Economics, Vol:7, pp:85-107.

Fanning K, Cogger KO, Srivastava R. (1995), "Detection of management fraud: A neural network approach", Intelligent Systems in Accounting, Finance and Management, 4: No. 2, pp:113–126.

Fanning K, Cogger KO, (1998), "Neural Network Detection of Management Fraud Using Published Financial Data", Intelligent Systems in Accounting, Finance and Management,. 7, pp:21–41

Gülseçen, S. (1995). "Yapay Sinir Ağları ile Finansal Uygulamalar ve Döviz Kuru Tahmini İçin Bir Öneri", 4.Türk Yapay Zeka ve Yapay Sinir Ağları Sempozyumu Bildirileri, 26-28 Haziran. (TAINN'95), pp:51-61.

Güneri, N. (2001). Öğrenci Başarısızlıklarının Analizinde Sinir Ağları Yaklaşımının Lojistik Regresyon Analizi İle Karşılaştırılması, (Yayınlanmamış Yüksek Lisans Tezi), Ankara: Ankara Üniversitesi.

Hecht-Neilsen, R. (1989). Neurocomputing, Massachusetts, Addison-Wesley Pub. Comp.

Jones, J., (1991), "Earnings management during import relief investigations", Journal of Accounting Research, Vol: 29, No:2, pp:193-228.

Koh, H.C., Tan, S.S. (1999), "A Neural Network Approach to the Prediction of Going Concern Status", Accounting and Business Research, Vol: XXIX, No:3, pp: 211-216.

Küçükkocaoğlu, G., Küçüksözen, C., (2005), "Gerçeğe Aykırı Finansal Tabloların Ortaya Çıkarılması: İMKB Şirketleri Üzerine Ampirik Bir Çalışma", Muhasebe ve Finansman Dergisi, Sayı 28.

Küçüksözen, C., (2005), "Finansal Bilgi Manipülasyonu: Nedenleri, Yöntemleri, Amaçları, Teknikleri, Sonuçları Ve İMKB Şirketleri Üzerine Ampirik Bir Çalışma", SPK Yayınları, No.183, Temmuz 2005, Ankara.

Küçüksözen, C., Küçükkocaoğlu, K., (2005), “Finansal Bilgi Manipülasyonu: İMKB Şirketleri Üzerine Ampirik Bir Çalışma”, "1st International Accounting Conference On The Way To Convergence" Kasım 2004, İstanbul, Muhasebe Bilim Dünyası (Mödav) Conference Booklet.

Salchenberger, L.M., Çınar, M., Lash, N.A. (1992). “Neural Networks: A New Tool for Predicting Thrift Failure”, Decision Sciences, Vol: XXIII, No: 4, pp: 899-916.

Spathis, C., (2002), “Detecting False Financial Statements Using Published Data: Some Evidence From Greece”, Managerial Auditing Journal, Vol: 17, No: 4, pp: 179-191.

Spathis, C., Doumpos M., Zopounidis, C., (2004), “Detecting falsified financial statements: a comparative study using multicriteria analysis and multivariate statistical techniques”, The European Accounting Review, Vol: 11, No: 3, pp: 509–535

Sungur, M. (Ed.) (1995). Mühendis Gözüyle Yapay Sinir Ağları, Ankara, ODTÜ.

Wilson, N., Chong, K.S. (1995). “Neural Network Simulation and Prediction of Corporate Outcomes: Some Empirical Findings”, International Journal of the Economics of Business, Vol: II, No: 2, pp: 31-51.

Yıldız, B. (2001). “Finansal Başarısızlığın Öngörülmesinde Yapay Sinir Ağı Kullanımı ve Halka Açık Şirketlerde Ampirik Bir Uygulama”, İstanbul Menkul Kıymetler Borsası Dergisi, Vol:17, pp : 51-67.

Appendix.

The functions and calculation methods of 10 independent variables that have been determined for our empirical study are explained below[7].

Sales Growth Index (SGI)

$$(SGI) = \frac{\text{Gross Sales}_t}{\text{Gross Sales}_{t-1}}$$

can be calculated as above. Sales growth does not necessarily prove the manipulation of financial information. According to the professionals, growing companies that take sales growth into account are more inclined to manipulation of financial information compared to other companies; because, the structure of debt/equity and the needs of resources create pressure on managers in order to increase sales in these companies. If a decrease in the prices of common stock is observed related to slowing down on growth of these companies, the more pressure will be seen on managers in order to manipulate financial information in such a case.

Days Sales in Receivables Index (DSRI)

$$(DSRI) = \frac{\text{Receivables}_t / \text{Gross Sales}_t}{\text{Receivables}_{t-1} / \text{Gross Sales}_{t-1}}$$

can be calculated as above. This index shows the change in trade receivables at time t by comparing them at time t-1 according to sales. As long as there is no extreme change in the policy of credit sales of the company, this index is expected to have a linear structure. An important increase in this index is based not only on the accountancy of consignment sales recorded as trade receivables and sales toward the increase in income as well as profit of the company but also on the creation of trade receivables from current accounts of group companies. These two applications are considered as the indicators of the manipulation of financial information.

Gross Margin Index (GMI)

$$(GMI) = \frac{(\text{Gross Sales}_{t-1} - \text{Cost of Goods Sold}_{t-1}) / \text{Gross Sales}_{t-1}}{(\text{Gross Sales}_t - \text{Cost of Goods Sold}_t) / \text{Gross Sales}_t}$$

can be calculated as above. If the index is greater than 1, it indicates that gross margin of the company is getting worsen. This indicator is a negative sign for future expectations of the company. In order to forward gross margin into a positive direction, it is assumed that companies will apply for manipulation of financial information within the objective of creating a picture of increase in sales revenue or decrease in cost of sales or both.

---

[7] The data set and explanations related to 10 independent variables determined for our study are taken from the studies of Küçüksözen (2005) and Küçüksözen and Küçükkocaoğlu (2005).

Asset Quality Index (AQI)

$$(AQI) = \frac{(1 - \text{Current Assets}_t + \text{PPE}_t) / \text{Total Assets}_t}{(1 - \text{Current Assets}_{t-1} + \text{PPE}_{t-1}) / \text{Total Assets}_{t-1}}$$

can be calculated as above. This index shows the change in other non current assets except current assets and plant, property and equipment within total assets by compared to previous year. If this index is greater than 1, it is an indicator that the company will capitalize its expenses instead of writing them as current period expenses on the income statement. In this context, this situation is considered as manipulation of financial information. Therefore, it is expected a positive correlation between asset quality index and financial information manipulation.

Depreciation Index (DEPI)

$$(DEPI) = \frac{\Delta\text{Depreciation}_{t-1} / (\Delta\text{Depreciation}_{t-1} + \text{PPE}_{t-1})}{\Delta\text{Depreciation}_t / (\Delta\text{Depreciation}_t + \text{PPE}_t)}$$

can be calculated as above. In our study, depreciation expenses were not directly calculated by using data from balance sheet and income statement. For this reason, depreciation expense of any period is determined as the difference between accumulated depreciation of current period and accumulated depreciation of previous period. This amount may create difference in terms of current period's depreciation expense. In this context, in depreciable assets, the change in current period will vary the amount of accumulated depreciation without affecting depreciation expense very much. Also, as it is mentioned below, this approach is going to be more appropriate to calculate depreciation expense by considering that these companies belong to reel sector as well as there is no big change in their depreciable assets.

If this proportion is greater than 1, this situation indicates that the company decreases its depreciation expenses in order to declare high profit by considering that the expected useful life of plant, property and equipment will be lengthened or the method of depreciation will be changed in such a way to reduce expenses.

On the other hand, it is expected that this index will not change very much by considering that companies which constitute our study are manufacturing companies in reel sector. In manufacturing industry, it is not expected that depreciable assets of these companies will increase or decrease very much in the context of purchases and sales. By taking this factor into account, if an important increase is observed on this index on a yearly basis, this situation is accepted as an indicator of financial information manipulation. For this reason, it is assumed that there is a positive correlation between depreciation expenses and financial information manipulation in our model.

Sales, General and Administrative Expenses Index (SGAI)

$$(SGAI) = \frac{(\text{Mkt. Sales Expenses}_t + \text{Gen. Adm. Expenses}_t) / \text{Gross Sales}_t}{(\text{Mkt. Sales Expenses}_{t-1} + \text{Gen. Adm. Expenses}_{t-1}) / \text{Gross Sales}_{t-1}}$$

can be calculated as above. It is expected that there is a correlation which will not change for a long time between marketing, sales, distribution and general administrative expenses and sales. These expenses will change according to main activities of the company; in other words, these expenses are variable expenses based on change in sales. In this context, it is accepted that sales are manipulated or expenses are under priced in case of important changes which take place in this variable, in other words, in case of an significant decrease in the proportional relationship between sales and these expenses, as long as there is no important increase in efficiency. Within the framework, it is assumed that there is a positive correlation between this index and financial information manipulation.

Leverage Index (LVI)

$$(\text{LVI}) = \frac{(\text{Long Term Debt}_t + \text{Short Term Debt}_t) / \text{Total Assets}_t}{(\text{Long Term Debt}_{t-1} + \text{Short Term Debt}_{t-1}) / \text{Total Assets}_{t-1}}$$

can be calculated as above. If this variable is greater than 1, it indicates that the proportion of obligation of the company has been increased. The reason behind this variable being in this model is to determine the manipulation of financial information applications which will provide the opportunity to get rid of any conditions on not meeting company's obligations.

Total Accruals to Total Assets Index (TATA)

$$(\text{TATA}) = \frac{\begin{array}{l}\Delta \text{Current Assets} - \text{Cash and Marketable Securities} - (\Delta \text{Short Term Debt} \\ - \Delta \text{Current Portion of the Long Term Debt} \\ - \Delta \text{Deferred Taxes and some other Legal Laibilities}) \\ - \text{Depreciation Expenses}_t\end{array}}{\text{Total Assets}_t}$$

can be calculated as above. When we calculate total assets index in such way, this shows the change in between debt-receivables and revenue-expense items within the framework of accrual basis and based on company's administrative initiatives. The reason behind this variable being in this model, is to determine any manipulation of financial information applications based on increase in revenue or decrease in expense or vice versa within the framework of accrual basis. In this context, if this variable, in other words, non-cash working capital increases or decreases dramatically, it is assumed that manipulation of financial information takes place.

Days in Inventory Index (DII)

$$(\text{DII}) = \frac{\text{Inventory}_t / \text{GrossSales}_t}{\text{Inventory}_{t-1} / \text{GrossSales}_{t-1}}$$

can be calculated as above. In order declare low or high profit, company's managers change general production overheads, cost of goods sold or inventory valuation methods such LIFO, FIFO or Weighed Average.

Financial Expenses Index (FEI)

$$(FEI)= \frac{\text{Financial Expenses}_t / \text{Gross Sales}_t}{\text{Financial Expenses}_{t-1} / \text{Gross Sales}_{t-1}}$$

can be calculated as above. In Turkey, one of the most often situation that we face on manipulation of financial information is to capitalize financing expenses by adding into accounts receivable, inventory, next year's expenses, associates, plant, property and equipment, intangible assets, and/or continuing investments instead of recording financing expenses as current period expenses on the income statement. In this regard, company's managers will be able to reach their own final results by capitalizing important portion of financing expenses by increasing profit. Or, to decrease profit, they will record financing expenses as current period expense. Due to the flexible structure of tax law on the subject of recording financing expenses as expense for current period or capitalizing them, the applications of financial information manipulation are enchained. Within this framework, it is assumed that there is a correlation between this index and manipulation of financial information.