Süleyman Demirel Üniversitesi
İktisadi ve İdari Bilimler
Fakültesi Dergisi
Y.2009, C.14, S.2 s.157-170.

Suleyman Demirel University
The Journal of Faculty of Economics
and Administrative Sciences
Y.2009, Vol.14, No.2 pp.157-170.

THE USE OF DATA MINING TECHNIQUES IN DETECTING FRAUDULENT FINANCIAL STATEMENTS: AN APPLICATION ON MANUFACTURING FIRMS

HİLELİ FİNANSAL TABLOLARIN TESPİTİNDE VERİ MADENCİLİĞİ TEKNİKLERİNİN KULLANIMI: İMALAT FİRMALARI ÜZERİNE BİR UYGULAMA

Yrd.Doç.Dr. H. Ali ATA[*]
Yrd.Doç.Dr. İbrahim H. SEYREK[**]

ABSTRACT

Detection of fraudulent financial statements is a very important issue for auditors. Because of the difficulty of detection of such fraudulent financial statements, several techniques, both qualitative and quantitative, are being employed by auditors. In this study, a set of data mining techniques not widely known to auditors are used to help the detection of financial statement fraud. The study is done based on the data from 100 manufacturing firms listed in ISE. The results show that leverage ratio and return on assets ratio are important financial ratios in detecting financial statement fraud.

ÖZET

Hileli finansal tabloların tespiti denetçiler için oldukça önemlidir. Bu tür hileli finansal tabloların tespit edilmesi oldukça zor olduğundan, denetçiler nicel ve nitel birçok teknik kullanmaktadırlar. Bu çalışmada denetçiler tarafından yaygın olarak bilinmeyen bazı veri madenciliği teknikleri, finansal tablolardaki hileleri tespit etmeye yardımcı olmak üzere kullanılmıştır. Çalışma İMKB'de işlem gören ve imalat sektöründe faaliyet gösteren 100 firmanın bilgilerine dayalı olarak gerçekleştirilmiştir. Araştırma sonucunda kaldıraç oranı ve aktif karlılık oranının finansal tablo hilesini tespit etmede önemli finansal oranlar olduğu belirlenmiştir.

Fraud, Financial reporting, Fraudulent financial statements, Data mining.
Hile, Finansal raporlama, Hileli finansal tablolar, Veri madenciliği

[*] Gaziantep Üniversitesi, İİBF. İşletme Bölümü, ata@gantep.edu.tr
[**] Gaziantep Üniversitesi, İİBF. İşletme Bölümü ,seyrek@gantep.edu.tr

1. INTRODUCTION

With the high amount of reported fraud cases and the damages it causes to companies and investors, it becomes increasingly more important to detect such fraudulent activities. Fraud is one important reason for the failure of many companies and it especially causes damage in capital markets because investors, creditors and financial analysts base their decision on the publicly available financial statements. In capital markets, the existence of fraudulent financial statements causes a big threat to investors' trust to companies and affects their investment decisions.

Fraud is a broad concept but basically there are two types of fraud seen in practice. The first one is the misappropriation of assets and the second one is the fraudulent financial reporting[1]. Misappropriation of assets involves such things as outright theft, embezzlement, padding of expense accounts, misuse of company assets, etc. On the other hand, fraudulent financial reporting involves an intentional distortion of the financial statements such as reporting sales that did not happen, reporting income into the current year that actually belongs in the next year, capitalizing expenses improperly or reporting an expense in the next year that should be reported in the current year. Financial statement fraud is typically conducted by management or with their consent and knowledge. Therefore, financial statement fraud can be seen as management fraud and is defined by Elliott and Willingham[2] as: "The deliberate fraud committed by management that injures investors and creditors through materially misleading financial statements."

It is the auditors' responsibility to determine whether the financial reports are in accordance with GAAP, assess the risk of fraud in financial reports, and identify the presence of fraud in financial statements. On the other hand, the detection of fraud in financial statements is a difficult task which requires more than using standard auditing procedures. Therefore, auditors need new tools and techniques to simplify auditing tasks and help them in detecting such fraudulent financial statements. In this regard, computer based tools can be a great helper. In this study, data mining is used in determining the fraudulent companies.

2. LITERATURE REVIEW

There are many examples of companies, both public and private, which have suffered due to fraud committed by finance or related departments in the organization. This situation has led many researchers to do empirical researches on the subject. Most of the previous studies on financial statement fraud were about fraud risk factors known as red flags in the

---

[1] Louis BRAITTA, *Audit Committee Handbook*, NJ, USA, John Wiley &Sons, 2004, p.36

[2] R.K. ELLIOTT, and J.J. WILLINGHAM, *Management Fraud: Detection and Deterrence*, Petro celli Books, New York, 1980, p.4

literature. Many researchers (Romney et al.[3], Loebbecke et al.[4], Heiman-Hoffman et al.[5], Abdul Majid and Tsui[6], Apostolou et al.[7]) have tried to determine the factors which are related to fraudulent financial statements by using these red flag indicators.

In general, financial statement fraud techniques works by overstating assets and revenues and generally this is done by recording revenues prematurely or by using fictitious records[8]. The experience of auditors is an important factor for finding out financial statement fraud. For example, Moves and Hasan[9] found that the experience of auditors and their previous success in detecting fraud are important factors. Similarly, Knapp and Knapp[10] studied the effects of audit experience and risk assessment using analytical procedures and their findings indicate that audit managers are more effective than audit seniors in assessing the risk of fraud with analytical procedures. Another study by Ansah et al.[11] suggests that the size of the audit firms, the auditor's tenure and years of experience are important factors in detecting fraud.

In detecting fraudulent financial statements, analytical procedures including ratio analyses have been used extensively (Blocher[12], Calderon and Green[13], Loebbecke et al.[14]). For example, Persons[15] studied 103 pairs of

---

[3] B.M. ROMNEY, W.S. ALBRECHT, and D.J. CHERRINGTON, "Auditors and The Detection of Fraud", *The Journal of Accountancy*, May, 1980, pp. 63-69.

[4] K.J. LOEBBECKE, M.M. EINING, and J.J. WILLINGHAM, "Auditors' Experience with Material Irregularities: Frequency, Nature, and Detectability", *Auditing: A Journal of Practice & Theory*, Vol. 9, No. 1, 1989, pp. 1-28.

[5] B.V. HEIMAN-HOFFMAN, P.K. MORGAN, and M.J. PATTON, "The Warning Signs of Fraudulent Financial Reporting", *Journal of Accountancy*, October, 1996, pp. 75-86.

[6] G.F.A. ABDUL MAJID, and J.S.L. TSUI, "An Analysis of Hong Kong Auditors' Perceptions of The Importance of Selected Red Flag Factors in Risk Assessment", *Journal of Business Ethics*, Vol. 32, 2001, pp. 263-274.

[7] B. APOSTOLOU, J. HASSELL and S. WEBBER, "Management Fraud Risk Factors: Ratings by Forensic Experts", *The CPA Journal*, October, 2001, pp. 48-52.; B. APOSTOLOU, J. HASSELL, S. WEBBER, and G.E. SUMMERS, "The Relative Importance of Management Fraud Risk Factors", *Behavioral Research in Accounting*, Vol. 13, 2001, pp. 1-24.

[8] M.S. BEASLEY, J.V. CARCELLO, and D.R. HERMANSON, *Fraudulent Financial Reporting: 1987–1997.An Analysis of U.S. Public Companies*, COSO, New York, 1999, p.32

[9] G.D. MOYES, and I. HASAN, "An Empirical Analysis of Fraud Detection Likelihood", *Managerial Auditing Journal*, Vol. 11, No. 3, 1996, pp. 41-46.

[10] C.A. KNAPP, and M.C. KNAPP, "The Effects of Experience and Explicit Fraud Risk Assessment in Detecting Fraud with Analytical Procedure", *Accounting, Organizations and Society*, Vol. 26, 2001, pp. 25-37.

[11] S.O. ANSAH, G.D. MOYES, P.B. OYELERE, and D. HAY, "An Empirical Analysis of The Likelihood of Detecting Fraud in New Zealand", *Managerial Auditing Journal*, Vol. 17, No. 4, 2002, pp. 192-204.

[12] E. BLOCHER, *The Role of Analytical Procedures in Detecting Management Fraud*, Institute of Management Accountants, Montvale, N.J., 2002.

[13] T.G. CALDERON, and B.P. GREEN, "Signaling Fraud by Using Analytical Procedures", *Ohio CPA Journal*,Vol. 53, No. 2, 1994, pp. 27-38.

[14] LOEBBECKE et al., pp. 1-28.

[15] O.S. PERSONS, "Using Financial Statement Data to Identify Factors Associated with Fraudulent Financial Reporting", *Journal of Applied Business Research*, Vol. 11, No. 3, 1995, pp. 38-46.

fraud and non-fraud firms which were matched by industry and time period using logistic regression analysis and found that financial leverage, capital turnover, asset composition and firm size were significant factors influencing the likelihood of fraudulent financial reporting.

Recently, new models are being used in predicting financial statement fraud. For example, Eining, Jones, and Loebbecke[16] studied the effect of using expert systems on the performance of auditors in detecting fraud. They found that auditors using expert systems were better at finding fraud and were more consistent in their decisions. Green and Choi[17] developed a model based on neural networks. The results of the study show that neural networks can be a very useful tool in detecting financial statement fraud. Similarly, Fanning and Cogger[18] used a neural network to develop a model based on eight variables in detecting management fraud. They compared the performance of their model to other models based on discriminant analysis and logistics regression and claimed that their model performed better. Bell and Carcello[19] have developed a logistic regression model using fraud risk factors to predict the likelihood of financial statement fraud. The significant risk factors they found were weak internal control environment, rapid growth, inadequate or inconsistent profitability, undue emphasis on meeting earnings projections, dishonest or overly evasive management, ownership status (public vs. private), and interaction between a weak control environment and an aggressive attitude towards financial reporting. Spathis[20] also used logistic regression model to detect falsified financial statements. The results of this study suggest that there is a potential in detecting falsified financial statements by using published financial statements.

In a recent study by Sun and Lee[21], researches used data mining techniques in predicting financial distress. They combined attribute-oriented induction, information gain and decision tree models in their study using 35 financial ratios and data from 135 pairs of firms. The results of their study show that data mining techniques are feasible and valid for detecting fraudulent financial statements. Similarly, data mining techniques were used

---

16 M.M. EINING, D.R. JONES, and J.K. LOEBBECKE, “Reliance on Decision Aids: An Examination of Auditors’ Assessment of Management Fraud”, *Auditing: A Journal of Practice & Theory*, Vol.16, No.2, 1997, pp. 1-19.

17 B. P. GREEN, and J. H. CHOI, “Assessing the Risk of Management Fraud through Neural-Network Technology”, *Auditing: A Journal of Practice and Theory*, Vol.16, No.1, 1997, pp.14–28.

18 K. FANNING, and K. COGGER, “Neural Network Detection of Management Fraud Using Published Financial Data”, *International Journal of Intelligent Systems in Accounting, Finance & Management*, Vol.7, No.1, 1998, pp. 21–24.

19 T. BELL, and J. CARCELLO, “A Decision Aid for Assessing the Likelihood of Fraudulent Financial Reporting”, *Auditing: A Journal of Practice & Theory*, Vol.9, No.1, 2000, pp. 169–178.

20 C.SPATHIS, “Detecting False Financial Statements Using Published Data: Some Evidence From Greece”, *Managerial Auditing Journal*, Vol.17, No.4, 2002, pp. 179–191.

21 J. SUN, and H. LI, “Data Mining Method for Listed Companies’ Financial Distress Prediction” *Knowledge-Based Systems*, Vol.21, No.1, 2006, pp. 1–5.

for fraud detection in a study by Kirkos, Spathis and Manolopoulos[22]. In this study several techniques were compared by their performances and also the important financial ratios in detecting fraudulent financial statements were determined. A recent study by Chen and Du[23] also used data mining techniques and artificial neural networks using data from 68 firms in Taiwan Stock Exchange. They developed a financial distress model using financial and nonfinancial data. The results of the study suggest that artificial neural networks are better at predicting financial distress of companies compared to traditional statistical techniques. These studies show that data mining techniques are gaining popularity among researches in detecting financial statement fraud and related areas.

2.1. Data Mining

With the advent of information technology which enables and facilitates the collection, storage and processing of large amount of data, organizations now are in a better position to benefit from the data they collect from their processes, customers, partners and environment. But the huge amount and sophistication of data collected requires the use of advanced techniques, like data mining, to extract meaning from the raw data and use it for purposes beneficial to the organization.

Data mining is the term used to describe the analysis of data to discover previously unknown relationships that provide useful information[24]. It is used in several industries including healthcare, finance, retail, telecommunications, and others to solve problems and to improve the various aspects of the business. Data mining basically works by analyzing data and generating descriptive and predictive models which help to solve the problems. There are several classes of data mining applications used in practice. Larose[25] classifies data mining tasks in six categories including description, estimation, prediction, classification, clustering and association. Description is basically used to describe patterns and trends in data while in estimation the value of a numerical dependent variable is tried to be estimated based on one or more independent variables. Similar to estimation, classification techniques are used to find the value of a dependent variable based on some manipulating variables but in this case the dependent variable is categorical rather than a numerical value. Fraud detection is a common example of a classification problem. In prediction, the values to be found exist in the future like the prediction of the price of a stock in upcoming months or which team will be the winner of the cup. In clustering, one tries to

---

[22] E. KIRKOS, C. SPATHIS, and Y. MANOLOPOULOS, "Data Mining Techniques for the Detection of Fraudulent Financial Statements", *Expert Systems with Applications*, Vol. 32, No.4, 2007, pp. 995–1003.

[23] W.S. CHEN and Y.K. DU, "Using Neural Networks and Data Mining Techniques for The Financial Distress Prediction Model", *Expert Systems with Applications*, Vol. 36, 2009, pp. 4075–4086

[24] D. HAND, H. MANNILA, and P. SMYTH, *Principles of Data Mining*. MIT Press, Cambridge, MA, 2001, p.1

[25] D. T. LAROSE, *Discovering Knowledge in Data: An Introduction to Data Mining*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2005, pp. 11-17

find similarities among entities and try to group them. Finally, in association type of data mining applications, the objective is to determine relationships among several attributes of given entities. Many techniques are developed for the above mentioned classes of problems which come from various disciplines like artificial intelligence, pattern recognition, machine learning and statistics. Some of these techniques are decision trees, neural networks, etc.

Although there are several commercially available software packages with user-friendly graphical interfaces making complicated data mining tasks seemingly trivial to apply, it should be noted that data mining is a process with several steps during which careful human intervention and interpretation are required. To summarize, data mining consists of the following fundamental steps which needs to be carefully applied in order to generate meaningful results from the analyses performed. The first step in a data mining task is to define the business problem clearly. Then, the data are collected and prepared for modeling. This step is very time consuming and basically consists of data cleaning and manipulation. After that, a data mining model most suitable for the problem at hand is chosen and applied. In this phase generally data are split into training and validation sets. The training data are used to derive rules and formulas while the validation data is used how well the generated rules work on a different set of data. Based on the results of model application, the performance of the model is evaluated and either it is modified or deployed to solve the business problem.

3. RESEARCH METHODOLOGY

In order to use data mining techniques to detect the fraudulent financial statements, we used data from 100 Turkish companies listed in Istanbul Stock Exchange. Half of the companies showed some type of fraudulent financial statements in the auditors' reports, and the other half didn't. Since fraud detection can be considered a classification problem, we used two classification techniques: decision trees and neural networks which are briefly explained below.

3.1. Decision Trees

Decision trees are commonly used data mining algorithms for classification type of problems. Here, the algorithm tries to construct a tree structure consisting of nodes and branches connecting the nodes. The purpose of the tree is to classify the data according to the discrete values of a target variable using several predictor variables. Nodes of a decision tree represent the testing points of predictor variables. Based on the result of testing at a node, the tree may split into more decision nodes at lower levels or into leaf nodes. Root node is the first node of the tree so it is the node where the main split occurs. On the other hand, leaf nodes, where no more splitting occurs, represent classifications. Decision trees provide a hierarchical decision model and are generally easy to interpret. On the other hand, the generated decision tree model may be complex making it difficult to understand. This complexity may be due to over fitting and memorizing the training data which reduce generalizability of the resulting model[26]. There are several decision tree algorithms which try to develop best tree structure to fit the data. Most commonly used decision tree algorithms are ID3, C4.5 and CART.

3.2. Neural Networks

Neural networks have been used in several fields and are very important tools of data mining. Neural networks try to mimic how the human brain processes information. Similar to neurons in our brains, a neural network consists of several interconnected input streams whose inputs are fed into an activation function to produce its outputs. Unlike some other methods, neural networks provide a nonlinear model. This nonlinear modeling capability makes it ideal to solve complex processes in several areas[27]. On the negative side, neural networks are difficult to interpret and generally take longer time to train. Like other data mining techniques, neural networks are first trained using example inputs by which network learns the patterns and relationships, and then the learned results can be generalized and used on fresh data.

3.3. Variables in the Models

In order to determine the important financial ratios for the detection of fraudulent financial statements, we used the financial statements of 100 manufacturing firms listed in Istanbul Stock Exchange in 2005. Half of these firms are remarked as fraudulent based on the publicly available independent auditors' reports and the remaining 50 as non-fraudulent. In determining the fraudulent financial statements, the auditors' reports are very important because it is a summary of all the work done by the auditor during the auditing process and the conclusions they have reached. With the help of auditing reports, the financial situation of the audited firm, the results of their activities and the conformity of cash flows to generally accepted accounting rules are revealed. Therefore the opinions of the independent auditors are very important in determining whether the financial statements of firms are fraudulent or not.

The variables used in the study are calculated using the balance sheets and the income statements of the selected firms. In the selection of these variables, the previous studies in the literature were used. The selected variables are basically related to the liquidity, financial situation, the efficiency of the activities, and the profitability of the firm. In this regard, we determined 24 financial ratios to be included in the study. These variables and the reasons for selecting them are explained below.

For the firms, the high amount of debt structure is considered to be related to fraud[28]. Therefore, Total Debt to Total Assets (TD/TA), Times Interest Earned Ratio (TIER), Interest Expense / Operating Expenses (IE/OE)

---

[26] LAROSE, pp. 107-109

[27] M.P. WALLACE, "Neural Networks and Their Application in Finance", *Business Intelligence Journal*, July, 2008, pp. 67-76.

[28] PERSONS, pp. 38-46.

and the Earnings Before Interest and Taxes (EBIT) ratios are included in the analyses. With the increase of the ratio of debts in capital structure, the firm's liquidity risk increases but with effect of lever, the welfare level of company executives improves. Therefore, company executives want the ratio of debts to increase in the capital structure of the firm. This situation indicates that high level of debt usage increases the probability of fraud in financial statements.

Fraud in financial statements is done deliberately by the management, and it is carried out by manipulating the accounts of balance sheet and income statements. Several studies indicate that management has more power on manipulating inventory accounts[29]. In this regard, showing inventory at a cost lower than its actual value in the accounts or recording obsolete inventory are some examples of inventory manipulation. In order to see such manipulations, Inventory to Sales (IS), Inventory to Total Assets (ITA) and Turnover Rate of Inventory (TRA) variables are used in the study. Moreover, because of the effect of inventory costs, the relationship between sales and the cost of goods sold become vulnerable to manipulation. Therefore, Gross Profit to Net Sales (GP/NS) and Gross Profit to Total Assets (GP/TA) ratios are also used. The fraudulent recording of sales before they are earned can be seen as additional accounts receivable[30]. Therefore we included the ratios Account Receivable to Sales (AR/S) and the Turnover Ratio of Receivables (TRA) in our analyses.

For company managers, the increase of the welfare level of stakeholders is an important indication of their success. In order to maximize the benefits of shareholders, the profitability ratios should raise. Therefore, company executives, may manipulate the profitability ratios to improve shareholders benefits which results in fraud in financial statements[31]. In the analysis the variables which are used in calculating profitability ratios like Net Income to Net Sales (NI/NS), Operating Profit to Net Sales (OP/NS) and Net Profit to Net Sales (NP&NS) are also used.

While the increase of profitability ratios contributes to the welfare levels of shareholders, it causes a risk for the liquidity of the company because increase in profitability reduces solvency of the company. Therefore it is necessary to make an optimal balance between profitability and liquidity. In studies related to financial fraud detection, many researchers used liquidity ratios[32]. Some of these ratios are Current Assets to Current Liabilities (CA/CL), Quick Assets to Current Liabilities (QA/CL) and Cash to Current Liabilities (C/CL). Also studies indicated that fraud is related to company size and the Total Assets were used as a measure of size. In this study logarithm of Total Assets (LGTA) is used as a variable to investigate the

---

[29] H. SCHILIT, *Financial Shenanigans: How to Detect Accounting Gimmicks and Fraud in Financial Reports*, New York, USA, McGraw- Hill., 2002.

[30] FANNING and COGGER, pp. 21–24.

[31] S. L. SUMMERS, and J. T. SWEENEY, "Fraudulent Misstated Financial Statements and Insider Trading: An Empirical Analysis", *The Accounting Review*, Vol.73, No.1, 1998, pp. 131–146.

[32] KIRKOS et al., pp. 995–1003.

relationship between company size and fraud. In addition to above variables, Taxes to Sales (T/S) ratio is included in the model considering the relationship between taxes and fraud.

Before applying data mining techniques to selected variables, we performed t tests on each variable, using fraud and non-fraud companies as the grouping variable. For non-significant differences, the variable is excluded and not used in the following analyses since those variables wouldn't make a contribution to the final analyses. The results of the t tests and the selected variables to be used in data mining methods are shown in Table 1.

Table 1: The Result of T Test Analyses

Variable*	Mean-NF	Mean-F	t-value	Sig
LGTD	7,69	7,71	-0,135	0,893
D/E	0,5	0,22	0,074	0,941
TIER	3,62	0,89	5,149	0,000
TD/TA	0,29	0,71	-5,084	0,000
I/S	0,13	0,24	-2,713	0,009
I/TA	0,12	0,18	-3,050	0,003
GP/NS	0,27	0,11	4,990	0,000
NP/TA	0,09	-0,06	6,755	0,000
LTD/TA	0,3	0,31	-0,212	0,833
NP/NS	0,11	-0,14	4,149	0,000
LGTA	8,28	7,94	2,728	0,008
WC/TA	0,47	0,46	0,355	0,723
CA/CL	3,31	1,22	5,754	0,000
C/CL	1,18	0,14	4,435	0,000
OP/NS	0,14	-0,11	5,901	0,000
QA/CL	2,50	0,75	6,110	0,000
TRE	1,99	3,67	-0,576	0,567
TRA	1,26	0,96	1,496	0,140
TRAR	158,62	13,40	1,103	0,276
TRI	15,50	6,89	2,738	0,008
AR/S	0,14	894,15	-1,000	0,322
T/S	-0,03	-0,03	0,016	0,987
GP/TA	0,25	0,12	5,351	0,000
IE/OE	0,09	0,33	-3,601	0,001

* Significant variables are shown in bold

4. ANALYSIS AND RESULTS

Using the selected variables mentioned above, we developed two classification models based on decision trees and neural networks. In both methods, almost half of the randomly selected cases (47 out of 100) were used for model training and the remaining cases (53 out of 100) were used for model validation.

The decision tree model constructed is shown in Figure 1 and the performance of the model is shown in Table 2. In the training phase, the tree correctly classifies 95,74% of cases correctly and 4,26% incorrectly. The model shows a good performance on the test data. The constructed decision tree correctly classifies 36 out of the 53 cases correctly which means a 67,92% correct classification.

Table 2: Classification Performance of Decision Tree

Classification	Training Data Set		Validation Data Set	
	Number	Percentage	Number	Percentage
Correct	45	95,74%	36	67,92%
Wrong	2	4,26%	17	32,08%
Total	47		53	

As seen from Figure 1 the tree splits at the root with TD/TA variable. If the TD/TA ratio is greater than 0,497, the firm is classified as fraudulent. This means that companies with TD/TA greater than 0,5 is very likely to commit fraud. This result is consistent with similar studies in the field and shows that excessive debt usage increases the likelihood of financial statement fraud. Therefore, the ratio of Total Debt to Total Assets is one of the main ratios for auditors to check for the detection of a possible fraudulent activity. For those companies with TD/TA less than 0,497, the second important variable is Net Profit to Net Sales. If NP/NS is greater than 0,03 than it is classified as non-fraud. Therefore, auditors should check for the values of profit margin and turnover rate because these values affect the company's expected profitability on their investments. The companies with active profitability value of less than 3%, the probability of fraud is high. In these companies, the amount of inventory investments in total assets is very important. According to decision tree, companies with Inventory to Total Assets ratio less than 8,8% are fraudulent companies. If this ratio is above 8,8%, then the next important variable to look at is Gross Profit to Total Assets ratio. In that case, if the GP/TA ratio is greater than 18,6%, then the company is classified as fraudulent. This makes sense because it is a practice in some fraudulent companies to increase the amount of gross profit in total assets by recording costs less than their actual values even though the amount of inventory in total assets is high.

The other data mining technique employed is Neural Networks. The results of the neural network analysis are shown in Table 3. As the table shows in the training phase, the model correctly classifies 91,49% of the cases correctly and 8,51% incorrectly. For the validation data set, the model shows a good performance and correctly classifies 41 out of 53 cases which means 77,36% of the cases. This result indicates for our data neural network model performs better than decision tree model.

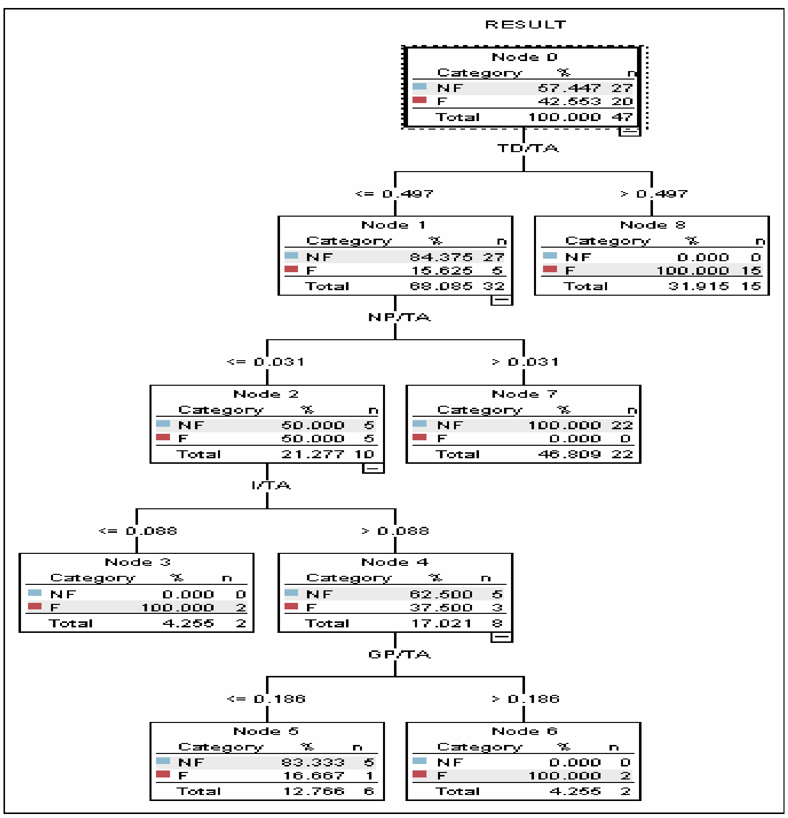
Table 3. Classification Performance of Neural Network

Classification	Training Data Set		Validation Data Set	
	Number	Percentage	Number	Percentage
Correct	43	91,49%	41	77,36%
Wrong	4	8,51%	12	22,64%
Total	47		53	

On the other hand, the results of neural networks are not as easily interpreted as decision trees. According to the results of neural network analyses top five most important variables in detecting fraud are as follows. Current Assets to Current Liabilities ratio is an important ratio showing the liquidity of the firm. Another important ratio is Operating Profitability to Net Sales ratio. OP/NS is a measure used to assess the success of firm in its main activity area. The third ratio is Total Debts to Total Assets ratio which is used to assess the financial situation of the company. TD/TA also affects profitability of the firm due to its leverage effect. One another important ratio is Inventory Expenses to Operating Expenses ratio. If the company uses too much debt, then IE/OE ratio decreases its profitability. Finally, Net Profits to Total Assets ratio is important and it shows the profits on its investments and it is an indicator of success with respect to financial performance.

In summary, the results of both analyses show that leverage ratio and return on assets ratios are the fundamental ratios in revealing financial statement fraud.

Figure 1: The Structure of the Generated Decision Tree Model



5. CONCLUSION

Financial statement fraud is a big concern for contemporary businesses, so companies place great importance to struggle with this problem. In order to prevent the damages caused by fraud, accountants and auditors should use new and innovative techniques to detect financial statement fraud. In this study, data mining techniques which are not widely known to auditors were used to help the detection of financial statement fraud. The study shows that application of data mining techniques on several financial ratios derived from the financial statements of companies can help auditors in fraud detection such that they can use the results of these analyses as an a early warning signal for the possible engagement of financial statement fraud. The detection of fraud indicators in financial statements has an important effect in determining financial statement fraud.

The use of data mining techniques has been increasing in several disciplines and there are many applications of these techniques for businesses including finance. In this study, two commonly used data mining techniques which are frequently used for classification type of problems, namely decision trees and neural networks, were used for the detection of financial statement fraud. The results of the study show that the use of high amount of debt in capital structure of companies is the most important indicator for the possibility of fraud in financial statements. In this regard, for companies with leverage ratio greater than 50%, it is very likely to commit financial statement fraud. The reason for this finding may be the fact that the use of more debt by company executives expecting that it will increase capital profitability actually increases leverage ratio and this situation increases the probability of financial statement fraud. Another important sign of fraud is return of assets. As the returns on investments of a company increases, the possibility of that firm to commit fraud decreases. Therefore, the auditors should pay attention to return ratio of assets in addition to leverage ratio and should be aware that profitability less than 3% indicates a high risk of fraud.

Other than these two main ratios, the other important warning signals are the amount of inventory in total assets, the of ratio gross profit to total assets, the ratio of current assets to current liabilities, the ratio of operating profits to net sales and the ratio of financial expenses to operating expenses.

Even though there are several techniques suggested in the literature to be employed by auditors to detect financial statement fraud, most of these studies also indicate that excessive amount of debts, manipulating inventory and receivables accounts and fraudulent operations on profitability ratios are very important issues to consider in fraud detection. Therefore, auditors should be aware of these activities in addition to checking financial ratios when they try to detect financial statement fraud.

BIBLIOGRAPHY

1. Abdul Majid, G.F.A. and Tsui, J.S.L., "An Analysis of Hong Kong Auditors' Perceptions of The Importance of Selected Red Flag Factors in Risk Assessment", *Journal of Business Ethics*, Vol. 32, 2001, pp. 263-274.
2. Ansah, S.O., Moyes, G.D., Oyelere, P.B. and Hay, D., "An Empirical Analysis of The Likelihood of Detecting Fraud in New Zealand", *Managerial Auditing Journal*, Vol. 17, No. 4, 2002, pp. 192-204.
3. Apostolou, B., Hassell J. and Webber, S., "Management fraud risk factors: ratings by forensic experts", *The CPA Journal*, October, 2001, pp. 48-52.
4. Apostolou, B., Hassell J., Webber, S. and Sumners, G.E., "The Relative Importance of Management Fraud Risk Factors", *Behavioral Research in Accounting*, Vol. 13, 2001, pp. 1-24.
5. Beasley, M.S., Carcello, J.V. and Hermanson, D.R., *Fraudulent Financial Reporting: 1987–1997. An Analysis of U.S. Public Companies*, COSO, New York, 1999.
6. Bell,T. and Carcello, J., "A Decision Aid for Assessing The Likelihood of Fraudulent Financial Reporting", *Auditing: A Journal of Practice & Theory*, Vol.9, No.1, 2000, pp. 169–178.
7. Blocher, E., *The Role of Analytical Procedures in Detecting Management Fraud*, Institute of Management Accountants, Montvale, NJ., 2002.
8. Braiotta, L., *Audit Committee Handbook*, NJ, USA, John Wiley&Sons, 2004.
9. Calderon, T.G. and Green, B.P., "Signaling Fraud by Using Analytical Procedures", *Ohio CPA Journal*, Vol. 53, No. 2, 1994, pp. 27-38.
10. Chen, W.S. and Du, Y.K. "Using Neural Networks and Data Mining Techniques for The Financial Distress Prediction Model", *Expert Systems with Applications*, Vol. 36 , 2009, pp. 4075–4086
11. Eining, M.M., Jones, D.R. and Loebbecke, J.K., "Reliance on Decision Aids: An Examination of Auditors' Assessment of Management Fraud", *Auditing: A Journal of Practice & Theory*, Vol.16, No.2,1997, pp. 1-19.
12. Elliott, R.K. and Willingham, J.J., *Management Fraud: Detection and Deterrence*, Petro celli Books, New York, 1980.
13. Hand, D., Mannila, H., and Smyth, P., *Principles of Data Mining*. MIT Press, Cambridge, MA, 2001.
14. Fanning, K.and Cogger, K., "Neural Network Detection of Management Fraud Using Published Financial Data", *International Journal of Intelligent Systems in Accounting, Finance & Management*, Vol.7, No.1, 1998, pp. 21–24.

15. Green, B.P. and Choi, J.H., “Assessing The Risk of Management Fraud Through Neural-Network Technology”, *Auditing:A Journal of Practice and Theory*, Vol.16, No.1, 1997, pp.14–28.

16. Heiman-Hoffman, B.V., Morgan, P.K. and Patton, M.J., “The Warning Signs of Fraudulent Financial Reporting”, *Journal of Accountancy*, October, 1996, pp. 75-86.

17. Kirkos, E., Spathis, C. and Manolopoulos, Y. , “Data Mining Techniques for The Detection of Fraudulent Financial Statements”, *Expert Systems With Applications*, Vol. 32, No.4, 2007, pp. 995–1003.

18. Knapp C.A. and Knapp, M.C., “The Effects of Experience and Explicit Fraud Risk Assessment in Detecting Fraud with Analytical Procedure”, *Accounting, Organizations and Society*, Vol. 26, 2001, pp. 25-37.

19. Larose, D.T., *Discovering Knowledge in Data: An Introduction to Data Mining*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2005, pp. 11-17

20. Loebbecke, K.J., Eining, M.M. and Willingham, J., “Auditors’ Experience with Material Irregularities: Frequency, Nature, and Detectability”, *Auditing: A Journal of Practice & Theory*, Vol. 9, No. 1, 1989, pp. 1-28.

21. Moyes, G.D. and Hasan, I., “An Empirical Analysis of Fraud Detection Likelihood”, *Managerial Auditing Journal*, Vol. 11, No. 3, 1996, pp. 41-46.

22. Persons, O.S., “Using Financial Statement Data to Identify Factors Associated with Fraudulent Financial Reporting”, *Journal of Applied Business Researc*, Vol. 11, No. 3, 1995, pp. 38-46.

23. Romney, B.M.. Albrecht W.S. and Cherrington, D.J., “Auditors and The Detection of Fraud”, *The Journal of Accountancy*, May, 1980, pp. 63-69.

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

25. Schilit, H., *Financial Shenanigans: How to Detect Accounting Gimmicks and Fraud in Financial Reports*, New York, USA, McGraw-Hill, 2002.

26. Summers, S.L. and Sweeney, J.T., “Fraudulent Misstated Financial Statements and Insider Trading: An Empirical Analysis”, *The Accounting Review*, Vol. 73, No.1, 1998, pp. 131–146.

27. Wallace, M.P., “Neural Networks and Their Application in Finance”, Business Intelligence, *Business Intelligence Journal*. July, 2008, pp. 67-76.

28. Sun, J. and Li, H.,“Data Mining Method for Listed Companies’Financial Distress Prediction” *Knowledge-Based Systems*, Vol. 21, No.1, 2006, pp. 1–5.