Predicting Qualified Audit Opinions Using Financial Ratios: Evidence from the Istanbul Stock Exchange

Alpaslan Yaşar
Assistant Professor
Department of Accounting and Finance
Osmaniye Korkut Ata University
Turkey

Emre Yakut
Assistant Professor
Department of Management Information Systems
Osmaniye Korkut Ata University
Turkey

M. Murat Gutnu
PhD Student
Department of Accounting and Finance
Osmaniye Korkut Ata University
Turkey

Abstract

The purpose of this study is to predict qualified audit opinions by using discriminant, logit and C5.0 decision tree based on twelve financial ratios. The sample consists of 110 firm-years data that includes 55 qualified opinions firm-year observations and 55 unqualified opinions firm-year observations listed in the industry index of Istanbul Stock Exchange (ISE) for the period 2010-2013. The results show that the variables X_{10} (retained earnings to total assets) is the significantly most effective variables to identify audit opinions by all of the models used in the study. Other significant variables in the analysis are found to be X_6 (equity to total liabilities), X_5 (total liabilities to total assets), X_{12} (net income to equity), X_9 (net income to total assets), X_3 (Working capital to total assets), X_7 (net sales to total assets). The classification results of the models indicate that C5.0 algorithm of decision tree has the greatest classification accuracy rate for explaining unqulified and qualified opinions of the firms, compared to discriminanat and logit models.

Keywords: qualified opinions, data mining, C5.0 decision tree, discriminant, logistic regression.

Jel Codes: M41, M42, C53

1. Introduction

In auditing literature, there has been an increasing interest to develope models for identfying type of audit opinions. In this context, various statistical methods applied to predict type of audit opinions. In recent years, especially, data mining classification techniques such as decision trees (e.g., C4.5, C5, CART algorithm), neural networks have been used to predict unqualified or qualified opinions. Laitinen and Laitinen (1998) classified prior studies on qualified audit report information into the three groups. The first goup studies has been used audit report information to construct bankruptcy prediction models (Keasey &Watson, 1987; Hopwood et al., 1989). Studies in the second class, have been constructed bankruptcy prediction models for making audit qualifications (Koh, 1991). In third category studies, several statistical models have been developed to predict audit report qualifications (Dopuch et al., 1987; Keasey et al., 1988; Laitinen and Laitinen, 1998; Gaganis et al., 2007; Kirkos et al. 2007). This study falls into the third category of the previous study groups above.

The aim of the study is to develope models and identify the effective financial ratios in predicting the qualified audit opinions by using Discriminat, Logit and C5.0 Decision Tree techniques in Istanbul Stock Exchange. Additionally, we identified the significant factors by each models and also compared the predictive accuracy of the models. The statistical models developed in this study benefit mainly for auditors in prediction of audit opinion type issued by other auditors in similar circumstances, when evaluating potential clients, in determining the scope of an audit for existing clients, to control quality within firms, and as a defense in lawsuits (Dopuch et al., 1987: 447-449). This study is the first study that develops the models for predicting qualified opinions in Turkey where institutional environment is different from the developed countries. The study contribute to literature by comparing multivariate discriminant and logit models to C5.0 algorithm of decision tree (as a classification technique of data mining) in an emerging market of Turkey where there has been different institutional environment from the developed countires. The remainder of the paper is organized as follows. The next section contains review of prior literature. In section 3, research methodology is discussed. Empirical results are presented in section 4. Conclusion are presented in the final section of the study.

2. Previous Research

Keasey et al. (1988) examined financial, organisational and auditor variables to explain small companies' audit qualification bu using multivariate logistic regression model in the UK. The results suggest that companies with audited by big audit firms, have a prior year qualification, a secured loan, declining earnings, large audit lags and few non-director shareholders are more likely to receive qualified opinions. Dopuch et al. (1987) used a probit model to predict based on financial and market variables to predict auditors' decisions to issue qualified audit opinions. The results showed that current year loss, stock returns minus industry returns, and change in ratio of receivables to total assets are significant variables in the prediction of audit qualifications. They also indicate that going concern opinions had the highest rate of accuracy in the prediction. Laitinen and Laitinen (1998) developed a multivariate logistic model to identify qualified audit opinions of Finnish firms publicely traded in Helsinki Stock Exchange. The model showed that the likelihood of receiving a qualification is larger, the lower the growth of the firm, the lower the share of equity in balance sheet and the smaller the number of employees. Spathis (2003) tested the combinations of financial and no-financial variables to predict the ability to discriminate between the choices of a qualified or unqualified audit report by using logistic and ordinary least squares regression models on a sample of Greek companies. Results showed that the qualification decision is associated by financial information (such as financial distress), and by non-financial information (such as firm litigation) with a accurate classification rate of 78%.

Spathis et al. (2003), developed a model to explain the audit qualifications of publicely traded Greek companies and compared the results of multicriteria decision aid classification method (UTADIS) to logisic regression and discriminant analysis. The results of the UTADIS model suggested that receivables to sales are the most important factor for the classifying the firms and also net profit to total assets and working capital to total assets are found to be significant. Specifically, the study showed that high receivables to sales, low net profit to total assets and low working capital to total assets describe audit qualifications. Doumpos et al. (2005) implemented support vector machine (SVM) to develop models (linear and non-linear) that may support auditors in issuing a qualified opinions if auditors conclude that the financial statements fail to represent the actual position of a firm. The results showed that all SVM models were capable of distinguishing between qualified and unqualified financial statements with satisfactory accuracy. Gaganis et al. (2007a) used probabilistic neural networks (PNNs) to develop a model that explains audit qualifications on a sample of publicely listed UK companies for the period of 1997-2004. The results showed that the performance of probabilistic neural networks model had higher explanatory power in explaining audit qualifications compared to artifical neural networks models and logistic regression.

Gaganis et al. (2007b) investigated the efficiency of k-nearest neighbours (k-NN) in developing models for estimating auditors' opinion, compared to models developed with discriminant and logit analyses. The results indicate that average classification accuracy of the k-nearest neighbours models are high and so the model can be more efficient than the discriminant and the logistic models. Finally, the results were mixed concerning the development of industry specific models, as contrary to general models. Kirkos et al. (2007) used three Data Mining classificaation techniques (C4.5 Decision Tree, Multilayer Perceptron Neural Network and Bayesian Belief Network) to develop models capable of identifying qualified auditors' reports in publicely listed UK and Irish firms.

The true predictive power of the models according to ten-fold cross validation results, the Bayesian Belief Network achieved the highest classification accuracy for estimating qualified audit opinions. The multilayer perceptron neural network model achieved a marginally lower performance and also the decision tree model achieved the lowest performance. Pasiouras et al. (2007) investigated the potentials of developing multicriteria decision aid models for reproducing the auditors' opinion on the financial statements of the firms based on a sample of private and public companies in the UK. The results showed that the two multicriteria decision aid techniques achieved almost equal classification accuracies and were both more efficient than discriminant and logit analyses. Saif at al. (2012) applied a support vector machine from data mining to find rules and identify audit opinions in Tehran Stock Exchange for the years 2001 through 2007. The results indicated that comprehensibility of new algorithm was better than that of support vector machine. The study obtained the 30 rules with 20 variables that help auditor for the audit opinion prediction. Valipour at al. (2013) examined the factors that affect audit reports and the possibility of predicting audit reports using meta-heuristic methods (i.e. neural networks, adaptive neuro-fuzzy inferences system-ANFIS and genetic algorithm), on asample of firms listed in Tehran Stock Exhange for the period of 2005-2011. The results indicated that net profit to sales ratio, current ratio, quick ratio, inventory turnover, collection period, and debt coverage ratio variables had the greatest effect on auditor's opinion. In addition, the results showed that the best performance among the tested models were the ANFIS model with fuzzy clustering and least-square back propagation algorithm.

3. Research Methodology

3.1. Sample Selection

The initial sample of the study consists of 140 Turkish firms that publicly traded in the industry index of Istanbul Stock Exchange (ISE) for the years 2010 through 2013. We excluded the financial companies (i.e. banks, insurance, factoring, leasing and holding companies) as in the audit literature. Additionally, we excluded 10 firms that have an outliers. After these exclusions, the sample size reduced to 520 firm-year observations of 130 firms that contained 55 qualified opinions firm-year observations and 465 unqualified opinion observations. The final sample selected by randomly and consists of 110 firm-years data that consist of 55 qualified opinions firm-year observations and 55 unqualified opinions firm-year observations for the years 2010-2013. The sources of financial ratios and audit opinions data in the sample are financial statements and audit reports in annual reports that gathered from the Istanbul Stock Exchange.

3.2. Variables

In this study, in order to identify the effective variables in the prediction of audit qualification, we selected twelve financial ratios that used especially in the Beaver (1966), Altman (1968), Zmijewski (1984) bankruptcy models and used some variables that used in the study of Landsman et al. (1999). These financial ratios are selected from the ratios of liquidity, financial structure, operating and profitability ratios. Additionally, the dependent variable in the analysis is type of audit opinions that equal to "1" if the firm receives a qualified opinions; and 0 if the firm with unqualified opinion. Thus, the variables list of the research is presented in Table 1.

Table 1. List of Variables

Y	Type of audit opinions (1 if audit opinion is qualified; 0 otherwise)
X_1	Current assets to current liabilities
X_2	Cash from operating activities to total assets
X_3	Working capital to total assets
X_4	Receivables plus inventories to total assets
X_5	Total liabilities to total assets
X_6	Equity to total liabilities
X_7	Net sales to total assets
X_8	Net sales to equity
X_9	Net income to total assets
X_{10}	Retained earnings to total assets
X_{11}	Earnings before interest and taxes to total assets
X_{12}	Net income to equity

3.3. Methods

In this study, we used discriminant, logistic and C5.0 decision tree technique of data mining methods to identify qualified opinions using financial ratios and so identify the effective financial ratios on qualified audit opinions and compared the accurate classification performance of these models.

3.3.1. Discriminant Analysis

Discriminant analysis that generates an equation which will minimize the possibility of misclassifying cases into their respective groups, has several purposes such as investigating differences between groups, determining the most parsimonious way to distinguish groups and classifying cases into groups. The fundamental assumptions of discriminant function analysis, like the assumptions of multiple regression, are multivariate normal distribution of predictor variables, homogeneity of variance-covariance matrices and non-multicollinearity. Thus, the form of discriminant function is (Burns and Burns, 2009:590-591):

$$\mathbf{D = v_1X_1 + v_2X_2 + v_3X_3 = \ldots\ldots\ldots v_iX_i + a}$$

Where:

D = discriminate function;
v = the discriminant coefficient;
X = respondent's score for that variable;
a = constant;
i = the number of predictor variables.

3.3.2. Logistic Regression

Logistic regression, also called a logit model, is used regularly when there are only two categories of the dependent variable. Logistic regression requires fewer assumptions and is more statistically robust compared to discriminant analysis. Logistic regression does not assume a linear relationship between the independent and dependent variables. The form of logistic regression equation is as follows (Burns and Burns, 2009: 590-591):

$$\text{logit } [p(x)] = \log \left[\frac{p(x)}{1 - p(x)} \right] = a + b_1 x_1 + b_2 x_2 + b_3 x_3 \ldots \qquad (1.1)$$

When equation (1.1) is rearranged, p can be calculated as the following formula:

$$p = \frac{\exp^{(a+b_1x_1+b_2x_2+b_3x_3\ldots)}}{1+\exp^{(a+b_1x_1+b_2x_2b_3x_3\ldots)}}$$

3.3.3. Data Mining

Data mining is the process of discovering models from large scale data (Zaki and Meira, 2014). Specific application areas of data mining are characterization and comparison, classification and prediction, and cluster analysis. Data mining includes several classifications method like Decision Trees, Bayesian Belief Networks, Neural Networks, Rough Sets, Support Vector Machines, and Genetic Algorithms. Especially, Decision Trees and Bayesian Belief Networks have advantage in the decision making process (Kirkos et al., 2007: 184).

In this study, we used the C5.0 algorithm of Decision Trees methods. A Decision tree consist of internal nodes and leaf nodes. Internal nodes represent the decisions coreesponding to the hyperplanes or split points. Leaf nodes symbolize regions or partitions of the data space, which are labeled with the majority class. Decision Trees have several advantages. One of the advantages of this method is that they produce models that are relatively easy to interpret. In particular, a tree can be read as set of decison rules (Zaki and Meira, 2014:483). Other advantages of Decision Trees are having no assumption about the distribution of data and have a fast learning mechanism (Kirkos et al., 2007: 185).

4. Empirical Results

4.1. Descriptive Univariate Results

In Table 1, we present the descriptive statistics and univariate tests of the variables. Columns (A) and (B) of Table 1 present mean and standard deviation values of the firms with unqualified and qualified audit opinions. Column (C) presents the results of parametric t-test statistics. Table 1 indicates that fims with qualified opinions have lower liquidity ratios (X_1, X_2, X_3, X_4), lower equity to total liabilities (X_5, X6), lower productivity of operations (X_7, X_8) and lower profitabillity (X_9, X_{10}, X_{11}, X_{12}). Thus, the univariate results show that there is a difference between firms with unqualified and qualified opinions for financial ratios that used in the study.

4.2. Multivariate Results

In the multivariate analysis, we assessed firstly the assumptions of discriminant analysis that include multivariate normality, homogeneity of variance-covariance matrices and non-multicollinearity. After Tabachnic and Fidell (1996) study, in literature, discriminant analysis is found to be relatively robust to violation of multivariate normality if the violation is not caused by outliers, and robust to the violation of homogeneity of variance-covariance matrices if the sample sizes are large or equal across groups. Deakin (1976) indicated that assumptions of normality for financial accounting ratios would not be defensible. He stated that normality can be achieved in certain cases by transforming the data although there are no rules which transformation would be apprropriate in a given situation. Tam ve Kiang (1992: 928) argued that transformations to approximate normal distributions complicated the interpretation of the transformed variables. Therefore, we implemented no transformations to the variables and multivariate normality and equal covariance assumptions are not presented separately. Additionally, we found that there was no multicollinearity problem between the independent variables by examining the correlation matrix, VIF (Variance Inflation Factor) and Tolerance values. In this study, in order to develop a model to predict qualified audit opinions in listed companies in Istanbul Stock Exchange of Turkey, we used the SPSS Statistics version 18 to analyze the data by using discriminant (stepwise) and logistic methods. The findings of these multivariate analyses are reported in the following.

4.2.1. Discriminant Results

We firstly examined the Eigenvalue and Wilks' Lambda statistics to identify the importance of discriminant functions were used in the analysis and presented in panel A and B of Table 2. In panel A of Table 2, eigenvalues is calculated as 1.064. This value indicate that dependent variable is explained by discriminant function and the function differenciates effectively the firms with unqualified and qualified opinions. Additionally, the square value of canonical correlation in Panel A of Table 2 indicates that the discriminant function explaines 51,6% of variance of dependent variable. Panel B of Table 2 shows the Wilks' Lambda statistics. In Panel B, the value of Wilks' Lambda is 0.484 and statistically significant at 1% level. This finding states that disciminant function have differenciate the groups statistically significant value and 48.4 of total variance can't be explained by discriminant function. In Table 3, we present standardized canonical discriminant function coefficients. The standardized coefficients of the variables indicate that the most efficient variable for separating the firms with unqualified and qualified opinions is X_{10} (retained earnings to total assets) and the other efficient variables are X_6 (equity to total liabilities), X_{12} (net income to equity), X_7 (net sales to total assets) and X_3 (Working capital to total assets), respectively.

Table 4 shows the unstandardized canonical discriminant function coefficients. We used the coefficients in Table 4 and developed a discriminant model for identifying the audit opinions of firms, as follows:

$$Z_i = - 1.390 + 0.940\ X_3 + 0.314\ X_6 + 1.165\ X_7 + 0.590\ X_{10} + 1.881\ X_{12}$$

Where:

Z_i : Discriminant scores,
X_3 : working capital to total assets,
X_6 : equity to total liabilities,
X_7 : net sales to total assets,
X_{10} : retained earnings to total assets,
X_{12} : net income to equity.

In Table 5, we present the classification results of discriminant analysis to review the success of accuracy prediction of the model. The classification results in Table 5 shows that the discriminant model classifies correctly 92.7% of the unqualified opinion and 81.8% of the qualified opinions. Additionally, total classification success of the model is 87.3%.

4.2.2. Logistic Regression Results

Logistic regression analysis is used alternative to discriminant analysis due to have less assumptions. In the study, forward-stepwise method of binary logistic regression analysis is also used to predict the qualified audit opinions and to compare the accuracy prediction of the both models. Table 6 presents the Hosmer and Lemeshow statistical test results. The results indicate that goodness of fit for the logit model.

We obtained the following logistic regression model by using the logistic coefficients in Table 7.

$$\ln\left(\frac{P}{1-P}\right) = 2.401 - 5.451\ X_3 - 1.001\ X_6 - 7.587\ X_{10}$$

Where:

X_3 : working capital to total assets,
X_6 : equity to total liabilities,
X_{10}: retained earnings to total assets.

In Table 7, X_{10} (retained earnings to total assets), X_6 (equity to total liabilities) and X_3 (working capital to total assets) are the statistically effective variables for predicting the qualified audit opinions, respectively. The negative signs of the logistic coefficients in Table 7 indicate that the probability of qualified opinons decreases when there is one unit increase in the working capital to total assets, retained earnings to total assets and equity to total liabilities. This result indicates that the higher the level of liquidity (X_3), profitability (X_{10}) and financial leverage (X_6), more likely to be higher the favorable audit opinion. Additionally, Nagelkerke R^2 statistics in step 3 indicate that independent variables explaine %78.4 of dependent variable. The classification results of binary logistic analysis in Table 8 show that the logit model classifies correctly 98.2% of the unqualified opinion and 87.3% of the qualified opinions. Additionally, total classification success of the model is 92.7%. The classification results in Table 8 indicate that the success of correct prediction of logit model (92.7%) is higher than sucess of the discrmininat model (87.3%).

4.2.3. C5.0 Decision Tree Results

In the study, we used C5.0 algorithm of decision tree algorithm in addition to discriminant and logit analyses to identify the qualified audit opinions by using the Clementine 12.0 programme that is a data mining tool. In C5.0 algorithm of decision tree, we compose the model by splitting the data set into the training and testing data set. Thus, we used 80% of the data set for training and 20% of the data for testing. In figure 1, the decision tree obtained by the C5.0 algorithm is given. In the taining set, aproximately 43 of the total 84 companies received unqualified audit opinion and 41 of the companies received qualified opinions. The root of the decision tree in Figure 1 indicate that X_{10} (retained earnings to total assets) variable have the strongest impact on the determination of audit opinions. In addition, X_5 (total liabilities to total assets), X_{12} (net income to equity), X_9 (net income to total assets) are found to be other important variables that have an greatest effect on audit opinions, respectively. The result is consistent with the discriminant and logistic results of this study. We identified following 7 rules that predict audit opinions based on the decision tree obtained by C5.0 algorithm in Figure 1. Tablo 10 shows the classification value obtained by C5.0 decision tree. The classification results of the decision tree model show that the model classifies correctly 96.4% of the unqualified opinion and 100% of the qualified opinions. Total classification success of the model is 98.2%. The classification results indicate that the success of correct prediction of C5.0 algorithm of decision tree (98.2%) is higher than the success of discriminant (87.3%) and logit models (92.7%).

5. Conclusions

In this research, we analyzed audit opinions in an emerging market of Turkey where institutional environment is different from the developed countries. We developed discriminant, logit and C5.0 Decision Tree models and identified qualified audit opinions by using financial ratios in Istanbul Stock Exchange.

The univariate results show that fims with qualified opinions have lower liquidity ratios, lower equity to total liabilities, lower productivity of operations and lower profitabillity (like the univariate results of the Laitinen and Laitinen, 1998). The result of discriminant analysis showed that X_{10} (retained earnings to total assets) variable is the most efficient variable for separating the firms with unqualified and qualified opinions. Other significant variables in the analysis are found to be X_6 (equity to total liabilities), X_{12} (net income to equity), X_7 (net sales to total assets) and X_3 (Working capital to total assets), respectively. The result is consistent with univariate results of this study. In addition, the discriminant model classified correctly 92.7% of the unqualified opinion and 81.8% of the qualified opinions. Total classification success of the model is 87.3%. The logistic regression model showed that X_{10} (retained earnings to total assets), X_6 (equity to total liabilities) and X_3 (working capital to total assets) are the statistically significant variables for predicting the qualified audit opinions, respectively. This result indicates that firms with have a higher level of liquidity, profitability and financial leverage, more likely to be lower the probability of qualified audit opinions. Additionally, the accurate prediction of the model is 98.2% for the unqualified opinion and 87.3% for the qualified opinions. Additionally, total classification success of the model calculated as 92.7%.

The C5.0 decision tree results indicate that X_{10} (retained earnings to total assets) variable have the efficient variable for predicting audit opinions. Other important variables have an greatest effect on audit opinions are found to be X_5 (total liabilities to total assets), X_{12} (net income to equity), X_9 (net income to total assets), respectively. The result is consistent with the univariate results of the study. The classification results of the decision tree model show that the model classifies correctly 96.4% of the unqualified opinion and 100% of the qualified opinions. Total classification success of the model is 98.2%. The results of the study demonstrated that The C5.0 decision tree has the highest accurate prediction of audit opinions compared to the multivariate discriminant and logit models. The results also showed that retained earnings to total asset is the most efficient variable on audit opinions by all of the three models. The possible limitation of this study is that some other variables that have an effect on audit qualification (such as prior audit opinion, audit firm size, company size) are not incorporated in the analysis to limit the scope of the study. Therefore, these variables may be used for the most accurate estimation of the audit opinions in the future research. In addition, other classification techniques of the data mining can be used in the future researches to increase the comparability of the results between emerging and developed markets.

References

Altman, E. I. (1968). Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy, Journal of Finance, 4, 589-609.

Beaver, W. (1966). Financial ratios as predictors of failures. Journal of Accounting Research, 4, 71-102.

Burns, Robert B., and Burns, Richard A. (2009). Business Research Methods and Statistics Using SPSS, Sage Publications Ltd.

Deakin, E. B. (1976). Distributions of financial accounting ratios: Some empirical evidence. Accounting Review, 51(1), 90-96.

Dopuch, N., Holthausen, R., and Leftwish, R. (1987). Predicting audit qualifications with financial and market variables. The Accounting Review, 62(3), 431-454.

Doumpos, M., Gaganis, Ch., and Pasiouras, F. (2005). Explaining qualifications in audit reports using a SVM methodology. Intelligent Systems In Accounting, Finance and Management, 13, 197-215.

Gaganis, C., Pasiouras, F., and Doumpos, M. (2007a). Probabilistic neural networks for the identification of qualified audit opinions. Expert Systems with Applications, 32(1), 114-124.

Gaganis, C., Pasiouras, F., and Doumpos, M. (2007b). A comparision of nearest neighbours, discriminant and logit models for auditing decisions. Intelligent Systems in Accounting, Finance and Management, 15, 23-40.

Hopwood, W., McKeow, J. and Mutchler, J. (1989). A test of the incremental explanatory power of opinions qualified for consistency and uncertainity. The Accounting Review, LXIV (1), 28-47.

Keasey, K. and Watson, R. (1987). Non-financial symptoms and the prediction of small company failure: a test of Argenti's hypotheses. Journal of Business Finance and Accounting, 14(3), 335-354.

Keasey, K., Watson, R., and Wynarzcyk, P. (1988). The small company audit qualification: a preliminary investigation. Accounting and Business Research, 18(72), 323-333.

Kirkos, E., Spathis, C., and Manolopoulos, Y. (2007). Data mining techniques fort he detection of fraudulent financial statements. Expert Systems with Applications, 32(4), 995-1003.

Koh, H.C. (1991). Model predictions and auditor assessments of going concern status. Accounting and Business Research, 21(84), 331-338.

Laitinen, E. K., and Laitinen, T. (1998). Qualified audit reports in Finland: Evidence from large companies. European Accounting Review, 7(4), 639-653.

Landsman, W., Nelson, K., and Rountree, B. (2009). Auditor Switches in the Pre-and Post-Enron Eras: Risk or realigment?. The Accounting Review, 84(2), 531-558.

Pasiouras, F., Gaganis, Ch., Zopounidis, C. (2006). Multicriteria decision support methodologies for auditing decisions: The case of qualified audit reports in the UK. European Journal of Operational Research, 180, 1317-1330.

Spathis, Ch. (2003). Audit qualification, firm litigaion, and financial information: An empirical analaysis in Greece. International Journal of Auditing, 7(1), 71-85.

Spathis, Ch., Doumpos, M., and Zopounidis, C. (2003). Using client performance measures to identify pre-engagement factors associated with qualified audit reports in Greece. The International Journal of Accounting, 38(3), 267-284.

Saif, S. M., Sarikhani, M., and Ebrahimi, F. (2012). Finding rules for audit opinions prediction through data mining methods. European Online Journal of Natual and Social Sciences, 1(2), 28-36.

Tabachnic, B. G., and Fidel, L. S. (1996). Using multivariate statistics, Harper Collins College Publisher, Third Edition, Northridge.

Tam, K. Y., and Kiang, M. Y. (1992). Managerial applications of neural networks: The case of bank failure predictions. Management Science, 38, 926-947.

Valipour, H., Salehi, F., and Bahrami, M. (2013). Predicting audit reports using Meta-Heuristic Algorithms. Journal of Distribution Science, 11(6), 13-19.

Zaki, M. J., and Wagner, M. Jr. (2014). Data Mining and Analysis - Fundamental Concepts and Algorithms, Cambridge University Press, Newyork, USA.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. Journal of Accounting Research, 22, 59-82.

Table 1. Descriptive Univariate Statistics

	(A) Unqualified Opinion (N=55)		(B) Qualified Opinions (N=55)		(C) Mean Difference
Variables	Mean	Std. Dev.	Mean	Std. Dev.	
X_1	2.598	1.504	1.243	1.235	1.355**
X_2	0.100	0.200	-0.002	0.183	0.102**
X_3	0.276	0.204	-0.125	0.418	0.401**
X_4	0.340	0.161	0.266	0.157	0.074*
X_5	0.331	0.152	0.650	0.388	-0.319**
X_6	2.790	1.917	1.125	1.310	1.665**
X_7	0.825	0.280	0.615	0.370	0.210**
X_8	1.326	0.616	0.701	2.252	0.624*
X_9	0.068	0.079	-0.045	0.125	0.113**
X_{10}	0.207	0.178	-0.996	1.447	1.203**
X_{11}	0.083	0.096	0.007	0.126	0.076**
X_{12}	0.104	0.111	0.002	0.333	0.101*

** Significant at 1% level

* Significant at 5% level.

Table 2. Eigenvalues and Wilks' Lambda Statistics

Panal A. Eigenvalues Statistics				
Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.064	100.0	100.0	0.718
Panel B. Wilks' Lambda statistics				
Test of Function	Wilks' Lambda	Chi-Square	df	Sig.
1	0.484	76.473	5	0.000

Table 3. Standardized Canonical Discriminant Function Coefficients

Variables	Fonksiyon 1
X_3 (working capital to total assets)	0.309
X_6 (equity to total liabilities)	0.515
X_7 (net sales to total assets)	0.382
X_{10} (retained earnings to total assets)	0.608
X_{12} (net income to equity)	0.468

Table 4. Canonical Discriminant Function Coefficients

Variables	Fonksiyon 1
X_3 (working capital to total assets)	0.940
X_6 (equity to total liabilities)	0.314
X_7 (net sales to total assets)	1.165
X_{10} (retained earnings to total assets)	0.590
X_{12} (net income to equity)	1.881
(Constant)	-1.390

Table 5. Classification Results[a]

Audit opinion			Predicted Group Membership		Total
			Unqualified	Qualified	
Original	Count	Unqualified	51	4	55
		Qualified	10	45	55
	%	Unqualified	92.7	7.3	100.0
		Qualified	18.2	81.8	100.0

a. 87.3% of original grouped cases correctly classified.

Table 6. Hosmer and Lemeshow Test

Step	Chi-Square	df	Sig.
1	20.199	8	0.010
2	24.805	8	0.002
3	9.033	8	0.339

Table 7. Logistic Regression Results

		□	S.E	Wald	df	Sig.	Exp (B)
Step 1[a]	X_3	-6.353	1.444	19.350	1	0.000	0.002
Constant		0.758	0.306	6.142	1	0.013	2.135
Step 2[b]	X_3	-6.414	2.533	6.411	1	0.011	0.002
X_{10}		-6.698	2.126	9.923	1	0.002	0.001
Constant		0.799	0.483	2.729	1	0.099	2.222
Step 3[c]	X_3	-5.451	2.782	3.838	1	0.050	0.004
X_6		-1.001	0.397	6.356	1	0.012	0.368
X_{10}		-7.587	2.581	8.642	1	0.003	0.001
Constant		2.401	0.841	8.153	1	0.004	11.039
		0.588					
Cox & Snell R^2		0.784					
Nagelkerke R^2							

a. Variable entered on step 1: X_3 (working capital to total assets)
b. Variable entered on step 2: X_{10} (retained earnings to total assets)
c. Variable entered on step 3: X_6 (equity to total liabilities)

Table 8. Classification Table of Logistic Regression[a]

Observed			Predicted		
			AO		Percentage Correct
			Unqualified	Qualified	
Step 1	AO	Unqualified	41	14	74,5
		Qualified	17	38	69,1
	Overall Percentage				71,8
Step 2	AO	Unqualified	51	4	92,7
		Qualified	9	46	83,6
	Overall Percentage				88,2
Step 3	AO	Unqualified	54	1	98,2
		Qualified	7	48	87,3
	Overall Percentage				92,7

a. Cut value 0.500

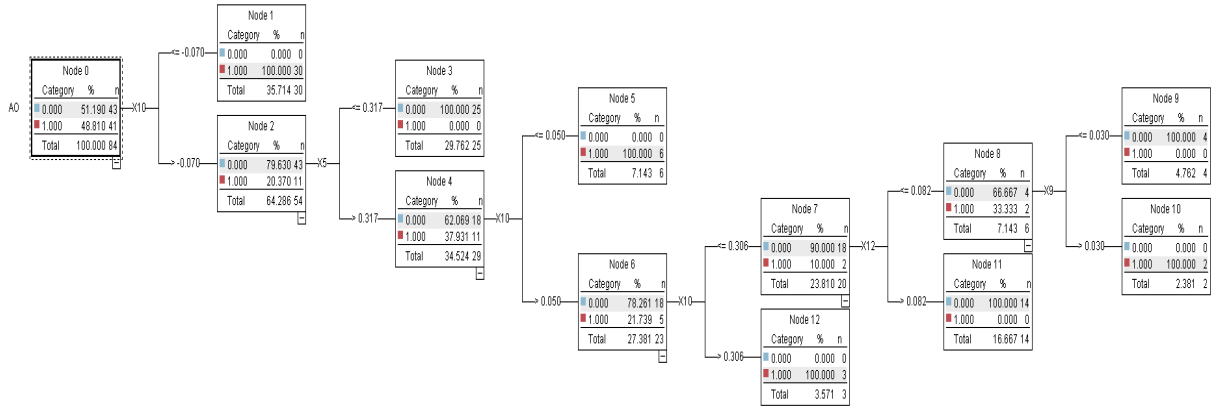

Figure 1. C5.0 Decision Trees for the Audit Opinions

Table 9. The Rules of Audit Opinions Obtained C5.0 Algorithm

	Conditions (IF)	Audit Opinion Type (THEN)
1	X_{10} <= -0.070	Qualified
2	X_{10} > -0.070, X_{10} <= 0.050, X_5> 0.317	
3	X_{10} > 0.306, X_5 > 0.317	
4	X_{10} > -0.070, X_5 > 0.317, X_{12} <= 0.082, X_9 > 0.030	
5	X_{10} > -0.070, X_5 <= 0.317	Unqualified
6	X_{10} > -0.070, X_{10} <= 0.306, X_5> 0.317, X_{12} > 0.082	
7	X_{10} > -0.070, X_{10} <= 0.306, X_5> 0.317, X_{12} <= 0.082, X_{11} <= 0.030	

Table 10. Classification Results of C5.0 Algorithm

Audit Opinion		Predicted			
		Unqualified	Qualified	Total	Percentage Correct
Observed	Unqualified	**53**	2	55	96,4
	Qualified	0	**55**	55	100
	Total	53	57	110	**98,2**