USING DATA MINING TECHNOLOGIES TO CONSTRUCT LITIGATION WARNING MODEL

Tzong-Huei Lin
National Kaohsiung University of Applied Sciences, Taiwan, R.O.C., tzonglin@kuas.edu.tw
Yu-Hsin Lu
Feng Chia University, Taiwan, R.O.C., luyh@fcu.edu.tw
Yu-Cheng Lin
National Kaohsiung University of Applied Sciences, Taiwan, R.O.C., linyc715@gmai.com
Yun-Hsuan Wang
Lancaster University, United Kingdom., y.wang55@lancaster.ac.uk

Abstract

This study used data mining technologies to construct a litigation warning model for auditors to assess risk when they evaluate whether accept or terminate an appointment, thus improve audit quality and prevent losses due to litigation. Feature selection technology was used to extract six important influencing factors among the many variables affecting auditors' litigation risk. Then a decision tree was used to establish a litigation warning model and a decision table for auditors' reference. The results indicated that the eight outcomes provided by the decision table could effectively distinguish the level of litigation risk with an accuracy rate of 92.708%

Keywords: Litigation risk, Data mining Technologies, Feature selection, Classifier technique

1. INTRODUCTION

During the 2001 Enron case, one of the Big 5 accounting firms at the time, Arthur Andersen, was sued for obstruction of justice by the US Department of Justice for destroying evidence; afterwards, Arthur Andersen was dissolved. According to deep pocket theory, when a company is charged with fraud, creditors and investors often pursue litigation against well-paid accountants, despite a lack of audit failure, in order to obtain more compensation

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

for losses [1]. Fraud cases not only directly put auditors at risk from litigation, but also come with large legal costs and can cause substantial harm to reputations. Lee and Gu[2] provided that increasing auditors' legal liability forces auditors to increase their independence in order to reduce the litigation risk from clients. The heavy burdens brought on by lawsuits have forced accounting firms to assess the possible litigation risk when accepting new clients. For example, after the Enron and Procomp Informatics cases, Deloitte Taiwan established a Reputation and Risk Department to assess clients' industry status and level of risk; KPMG Taiwan established an independent assessment team and a client risk screening team to determine whether new clients should be accepted [3]. When the risk is too high, auditors may refuse appointment or request termination. Due to the recent increase in the number of lawsuits, auditors can no longer turn a blind eye to the litigation risk. However, this decision making process is extremely complex as underestimating the risk may lead to future litigation and damage reputations. Therefore, investigating the factors influencing litigation against auditors and providing auditors with a litigation warning model is important.

Many studies indicated that data mining technologies, such as decision trees and artificial neural networks, are superior to traditional statistical methods in constructing detection models. This study used them to construct litigation warning model for auditors to assess risk in order to aid in evaluating whether auditors should accept or terminate appointments. The procedures included a literature review and the collection of factors influencing litigation against auditors. Data mining technologies were then used for feature selection to extract key factors [4]. Next, data mining technologies constructed the litigation warning model and decision table. Finally, the performances of warning models generated using data mining technologies and statistical methods were compared to serve as a reference for auditors when evaluating the litigation risk.

This study provides the following contributions: (1) In contrast with prior research, this study both reviewed relevant literature and collected all the factors influencing auditor litigation risk. The sampleused in this study consisted of companies listed on the Taiwan stock exchange. The legal liability of Taiwanese auditors was weaker than that of auditors in developed countries. However, the complexity of litigation taken against auditors is increasing. Therefore, the factors extracted during feature selection can be seen as the key factors influencing the litigation risk for auditors in Taiwan and serve as a reference for other developing countries. (2) In performance of models, the warning model constructed by data mining technologies in this study was superior to logistic regression and discriminant analysis. This result demonstrated the value of applying data mining in other relevant areas. (3) The difference between prior related literature and this paper was that this study did not only aim to improve model accuracy, but also to construct an understandable classification model founded on rules for auditors to use when evaluating risk in order to help auditors

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

screen potentially risky clients and decide which appointments to take, thus improving audit quality and lowering the litigation risk.

The following sections of the study include a literature review, methods, and the empirical results and analysis. The final section describes our conclusions.

2. LITERATURE REVIEW

Affecting Factors of Auditors' Litigation Risk

Assessing auditors' litigation risk is a complex procedure with many affecting factors. According to Arens, Elder [5], audit risk is the risk of an auditor issuing inaccurate audit opinion when the financial statement may be materially misstated; i.e., the risk of audit failure. Audit risk is affected by inherent risk, control risk, and detection risk. Therefore, this study explored the affecting factors of litigation against auditors from prior literatures in accordance with inherent risk, control risk, and detection risk in audit risk, and legal environment.

Inherent risk is the possibility of material misstatement in financial statements before considering the effectiveness of client's related internal controls. Changes in the industry or complex transactions may influence the inherent risk of a company. Therefore, when an auditor assesses inherent risk, the company's organizational situation, nature of services, type of transactions, fraud-related assets and liabilities, and socio-economic environment are often carefully considered. The size of the company [6-8], bankruptcy or major losses suffered by the client [9], companies with poor finances [10-12], are measurement variables. Summers and Sweeney [13] found that some financial indicators have strong explanatory power when predicting fraud in financial statements and were able to predict the possibility of auditors facing future litigation. In addition, the type of industry is also a critical factor determining litigation; the high technology and financial industries have high rates of litigation [7, 14, 15]. Therefore, when auditors assess litigation risk, they must also consider the client's industry.

Control risk is the possibility that financial statements have material misstatements that cannot be prevented or detected by client's internal controls. The governance level of client has the obligation to monitor the management to ensure the establishment and maintenance of internal control, reliable financial statements, efficient and effective operation, and compliance with relevant ordinances. The actualization of corporate governance monitors business performance, prevents conflicts of interest, and ensures compliance with various ordinances. Therefore, many prior literatures indicated significant correlations between the quality of corporate governance and the prevention of fraud which, in turn, influences the litigation risk against auditors [16]. Otherwise, Schwartz and Soo [17] and Knechel, Naiker [18] found that the market has a significant negative response to replacing auditors or

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

selecting smaller accounting firms as a replacement late in the fiscal year. Signaling theory implies that the viewpoints of the auditor and client under investigation differ more greatly when the time at which the auditor is replaced is later in the fiscal year. Selecting smaller accounting firms implies dismal financial prospects or pessimistic expectations [19], and increases the litigation risk against auditors.

Additionally, Pierre and Anderson [9] concluded that a large number of litigations against auditors is associated with insufficient disclosure of information of clients; this includes the type of accounts receivable, mortgages, subsequent events, and off-balance sheet debts. Therefore, any illegal activity committed by the clients that was punished by the competent authority increases the probability of litigation against auditors. Restatement of financial statements to correct material misstatements isequivalent to announcing to investors, "We lied." This assists investors in providing evidence, increasing the possibility of class action or private lawsuits [20].

Detection risk is the risk that the audit evidence collected is unable to detect materiality misstatements. This risk is correlated to the auditor's or accounting firm's audit quality and characteristics. DeAngelo [21] showed that the quality of audit services is defined to be the market-assessed joint probability that a given auditor will both (a) discover a breach in the client's accounting system, and (b) report the breach. The discovery of a breach is dependent on the professional competency of the auditor and the reporting of the breach is contingent upon auditor independence. With regards to professional competency, as larger accounting firms invest in special skills more than smaller accounting firms, they are more capable of finding materiality misstatements in financial statements. With a large number of clients, the quasi-rent from a certain client only accounts for a small percentage of the total quasi-rents for a firm; therefore, the degree of economic dependence on a given client is relatively low. As such, it is unlikely that independence will be compromised and more likely that the materiality misstatements in the financial statements of the client will be reported. Because larger accounting firms have the professional competency and objective independence to provide more accurate audit opinions, they are less often the subject of litigation [22-24]. Recently, many accounting firms have focused on developing industry specialism in order to achieve economics of scale and broaden market opportunities. Many literatures have used industry specialist to measure the audit quality on exploring the litigation risk against auditors [25-28].

With regards to auditor independence, Reynolds and Francis [29] stated that when the quasi-rent from one client accounts for a higher proportion of the quasi-rents from other clients, auditors are more likely to make compromises for the inaccuracies in the statements of more important clients; this has a negative influence on auditor independence and

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

increases the litigation risk against auditor. Arguments have always been presented in favor of and opposing auditor rotation to enhance independence. Those in favor of rotation claim that long-term appointment leads to closer ties between the auditor and client, which causes audits to become merely a formality; enforced rotation will help prevent pressure from management and increase auditor independence [30]. Those opposing rotation claim that auditing work involves a learning effect. Johnson, Khurana [31] proposed that enforcing auditor rotations may lower auditing efficiency and earnings quality. In terms of the auditor tenure, DeFond and Subramanyam [32] indicated that client's earnings management is more serious in the later period of auditor tenure, then the litigation risk against the auditor and the client's probability of a financial crisis increase. However, Lee and Lin [33]found that longer auditor tenure suppress earnings management by client. Due to going concerns assumption, companies may do audit opinion-shopping. In order to maintain clients, auditors tend to be unwilling to issue qualified opinion; thus, financial statements doctored by company management are provided to users for decision making [34]. When certified auditors fail to fulfill their intended role, they are often seen as conspirators in financial fraud by the inspection authority. Auditors should issue non-standard unqualified opinion audit reports to familiarize investors with the company's actual financial status; however, auditors do the opposite, causing harm to the readers of the financial reports. Palmrose [35] pointed out that the issuance of modified reports1 can help auditors counter accusations in court. The issuance of auditing reports other than unqualified opinions can moderately lower the litigation risk against auditors.

In 2004, Taiwan saw a quick succession of fraud cases involving Procomp Informatics, Infodisc Technology, and Summit Computer Technology which led to the Financial Supervisory Commission (FSC) of Taiwan issuing warnings about or canceling the certification of a number of auditors. The FSC made major revisions to the Certified Public Accountant Act, extending auditors' ever-increasing civil and criminal liabilities. A verdict was reached in the China Rebar case at the end of 2008; the court issued criminal sentences to the two auditors involved. Lin and Lin [36] warned the accounting field that future cases similar to the China Rebar case will further increase the criminal liability of auditors.

Use of Data Mining Technologies in the Formation of the Warning Model

Studies on predicting or detecting models developed rapidly after the first use of univariate discriminant analysis by Beaver [37], in which sample matching was used to predict financial crises in sample US companies. For example, Altman [38] used multiple discriminant analysis to construct a bankruptcy detection model; the results selected 22 financial ratios and developed the Z-score model often used in later studies [39]. Ohlson [40] used logistic regression to develop a bankruptcy detection model using nine financial ratios. Subsequent studies on crisis prediction models not only included non-financial variables in the model,

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

but also tried different statistical methods and tools to create warning models. However, traditional statistical methods are severely limited. For example, the variables must comply with normal distribution, which is impossible for certain variables. Thus, data mining technologies have gradually been introduced into crisis prediction models.

Frawley, Piatetsky-Shapiro [41] stated that data mining technologies find potential data hidden in previously unknown valuable information. Simply put, data mining efficiently searches databases for useful knowledge and principles by finding patterns and relationships. Recently, studies have used data mining technologies to construct prediction or detection models and have found that the prediction outcomes are superior to those of traditional statistical methods [42 -44]. Odom and Sharda [45] and Coats and Fant [46] utilized the Altman [38] Z-score along with a neural network and multivariate analysis, respectively, to create models. The results indicated that neural networks were superior to multivariate analysis. Chaveesuk, Srivaree-Ratana [47] used a back-propagation neural network, radial basis function networks, learning vector quantization, multiple regression, and logistic regression to create models; the results indicated that back-propagation neural networks were the superior method. Min and Lee [44] used a support vector machine, logistic regression, and multiple discriminant analysis to create a detection model and found that the support vector machine was optimal. In summary, past research has found that the warning model prediction accuracy of data mining technologies is superior to that of traditional statistical methods.

3. METHODS

Sample and Source of Data

The sample for this study consisted of 64 companies listed on the Taiwan stock exchange who had litigation taken against their auditors; these companies were chosen based on the 2002-2013 "Summary of Indictments and Sentences for Major Securities Crimes" issued by the Securities and Futures Bureau of the FSC and the class action litigation announced by the Securities and Futures Investors Protection Center. Sample matching was conducted in accordance with the methods used in Beaver [37]; non-litigated companies within the same time period, similar industries, and with similar asset scales acted as the control sample. A 1:2 [46] sample matching method comprised of 64 litigated companies and 128 non-litigated companies was used in this study. The study period was the year in which litigation was taken against the auditor; the financial variables used the obtainable financial data for the year of the litigation; corporate governance variables also utilized data from the year of the litigation. The sources of the data were the Taiwan Economic Journal (TEJ) database, the Market Observation Post System, and the financial reports and prospectuses of companies listed on the Taiwan stock exchange.

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

Variable Definition and Measurement

This study examined the factors influencing litigation taken against auditors from the standpoint of auditors. Influencing factors for audit risk were divided into inherent risk, control risk, and detection risk, and influencing factors for legal environment were also considered. After referencing relevant literature, 39 representative variables were selected (Table 1). The dependent variable was litigation against auditors, where 1 indicated litigation and 0 indicated no litigation.

Table 1 Study variables

	Variable	Method of measurement
	Influencing factors of audit risk- Inherent risk	
1	Asset scale	Natural logarithm of total assets
2	Period listed on stock exchange	Years listed on stock exchange
3	Credit risk index	Credit risk index listed in the TEJ database
4	Debt ratio	Total debt / total assets
5	Accounts receivable ratio	Accounts receivable / total assets
6	Accounts receivable turnover rate	Annual sales / average accounts receivable and negotiable instruments
7	Inventory ratio	Inventory / total assets
8	Inventory turnover rate	Annual sales / average inventory
9	Inventory growth rate	(Average inventory for current period / average inventory for previous period) - 1
10	Sales growth rate	(Sales revenue for current period / sales revenue for previous period) - 1
11	Operating profit ratio	Operating profit ratio / sales revenue
12	Operating profit growth rate	(Operating profit for current period / operating profit for previous period) - 1
13	Return on operating assets	Operating profit for the past four quarters / average total assets
14	Annual return on stock	Market-adjusted annual return on stock
15	Stock price fluctuation	Fixed asset model beta coefficient
16	Annual loss	1 for annual net income after taxes that signified loss; 0 if not
17	Heavy losses	1 for shareholder equity less than authorized capital stock; 0 if not
18	Working capital	Working cash flow / total debt
19	Insufficient working cash flow	1 for negative working cash flow; 0 for positive
20	Type of industry	1 for companies in the electronics industry; 0 if not
	Influencing factors of audit risk- Control risk	
21	Chairperson is an executive director	1 for if the chairperson is an executive director; 0 if not
22	Change of chairperson	Number of times a new chairperson was named within the past three years
23	Ratio of stock owned by directors	Number of stocks owned by directors / number of stocks sold to outsiders at year-end
24	Ratio of independent board members	Number of independent board members / total number of board members
25	Ratio of stock owned by legal entities	Number of stocks owned by legal entities / total number of stocks sold to outsiders at year-end
26	Change of auditor	Number of times a new auditor was appointed within the past three years
27	Illegal behavior	1 for any illegal activity committed by the company or auditor that was punished by the competent authority; 0 for none.
28	Restatement of financial reports	Number of times financial restatements were made
	Influencing factors of audit risk-Detection risk	
29	Large accounting firm	1 for companies audited by one of the four largest accounting firms; 0 if not
30	Industry expert_CPA	Number of clients listed on the stock exchange in an industry / total number of companies listed on the stock exchange in that industry

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

31	Client importance_CPA	log(revenue of company under investigation) / log(revenues of all CPA clients listed on the stock market)
32	Client importance_firm	log(revenue of company under investigation) / log(revenues of all accounting firm clients listed on the stock market)
33	Non-audit fees	1 for if the company's non-audit service fees account for over 1/4 of all fees paid to the accounting firm; 0 if not
34	Auditor's signature lag	Number of days between balance sheet date and audit report date
35	Firm appointment	Number of years the company has been audited by the same accounting firm
36	Auditor appointment	Number of years the company has been audited by the same certified accountant
37	Auditor industry appointment	Number of years the auditor has audited within the industry to which the company belongs
38	Audit opinion type	1 for if the company received an unqualified opinion the previous year; 0 if not
Influencing factors of litigation		
39	China Rebar case	1 for if the year of audit was after the China Rebar case (including 2007); 0 if not

Constructing the Litigation Warning Model by Using Data Mining Technologies

After gathering and organizing the influencing factors of litigation against auditors reported in prior literature, feature selection extract critical factors which were then categorized to create the litigation warning model and decision table in order to help auditors determine if they should accept or terminate an appointment. The detailed research procedure is shown in Figure 1.

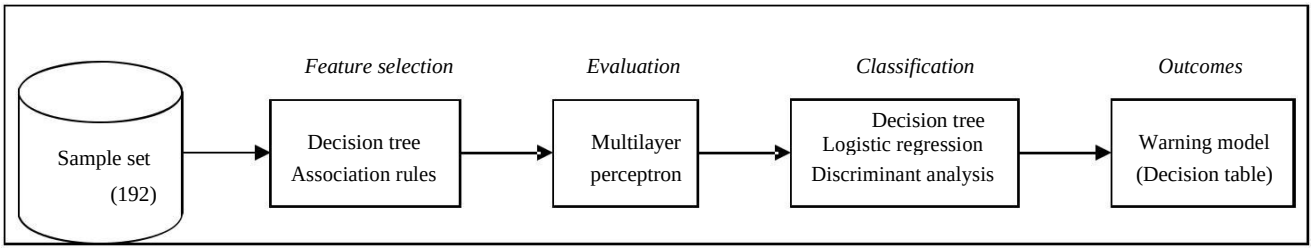


Figure 1 Model construction process

Feature selection

Improving prediction accuracy is an important part of creating an effective prediction or detection model. However, not all features and attributes within the data are associated with prediction accuracy. Prediction accuracy will decrease if insignificant features and attributes are included in classification and prediction. Methods to improve performance effectively remove clutter and irrelevant attributes. Traditional statistical methods mainly use dimension reduction or the extraction of variables; univariate methods primarily use partial correlation analysis and means testing; multivariate methods use principal component analysis and factor analysis. In data mining technologies, these methods are called feature selection processes; commonly used methods are decision trees and association rules [48, 49]. The objective of this study was to use data mining technologies to create a model; thus, decision trees and association rules were used in the feature selection process.

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

The screened variables were then input into a multilayer perceptron (MLP) algorithm and the resulting prediction abilities were used to assess the performance of decision trees and association rules in feature selection. The MLP algorithm used a back-propagation neural network leaning algorithm. Smith and Gupta [50] and Olafsson, Li [51] pointed out that neural networks are capable of pattern classification and pattern recognition. Neural networks are widely applied in many business solutions and the most commonly used algorithms in neural networks are MLP algorithms. Three steps were taken in the evaluation of feature selection performance in this study. First, the data set including all variables was entered in the MLP algorithm as the basis for the evaluation. Second, the key feature variables selected from the decision tree and association rules were separately entered in the MLP algorithm for comparison. Third, the performances of each model were compared with the benchmark; the comparison included prediction accuracy, type I and type II error rates, and the factor extraction rate. The feature selection process is shown in Figure 2.

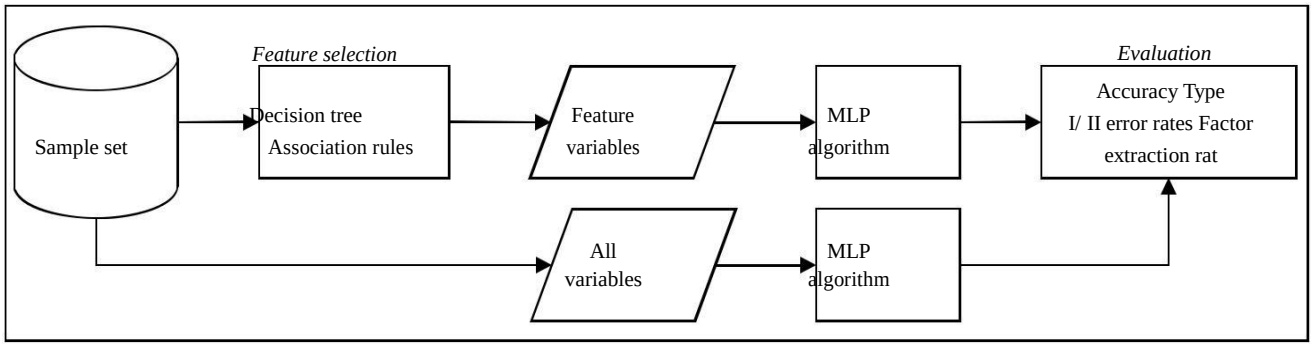


Figure 2 Feature selection process

When the benchmark and comparison models were created using the MLP algorithm, different parameters were set to prevent overtraining and obtain optimized models. Different numbers of hidden layers (8, 12, 16, 24, and 32) and learning cycles (50, 100, 200, 300, and 500) were paired to attain 25 prediction accuracies and type I and type II error results for each of the three models. The mean value for the 25 performance evaluations were used to compare prediction results. In addition, cross-validation was used for each model; 10-fold cross- validation2 of the accuracy and error rate from the original training sample set can prevent variations and deviations in the results due to different samples [52].

Classification

Classification is one of the most important techniques in data mining and is used to categorize the data to be processed according to attributes. The classification model helps illustrate the features of each category and can be used to detect data classifications. A classification technique commonly used in past research, the decision tree, was chosen for this study. The reason decision trees were used in this study was to provide decision makers with useful information for their reference when making decisions. A decision table was used

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

to present and analyze decision situations. This method was compared with the traditional statistical methods of logistic regression and discriminant analysis.

Model performance evaluation

For feature selection performance evaluations, the prediction accuracies, type I and type II errorrates, and factor extraction rates of the two models created using feature selection tools were compared with the benchmark model made without feature selection. Analysis of variance (ANOVA) was also used to determine whether the differences in performance between the three models were significant.

For classification technique performance evaluations, the prediction accuracies of the early warning models created using the decision tree, logistic regression, and discriminant analysis were compared. The evaluation methods were the same as those for the feature selection processes; 10-fold cross-validation was used to verify the prediction accuracies of the original training sample set, and the prediction accuracies and type I and type II error rates were the basis for model assessment. The area under the receiver operating characteristic (ROC) curve (AUC) 3 was also used to determine the predictive accuracy of the models.

4. EMPIRICAL RESULTS AND ANALYSIS

Warning Model Analysis

Feature selection results

An MLP algorithm was used to evaluate the feature screening performance of the two feature selection tools and a model that did not use feature selection. The results in Table 2 show that as the benchmark model screened for variables without a feature selection tool, the average training and testing times were the longest; moreover, the large number of dependent variables caused interference which lead to the poorest prediction accuracy (74.167%), type I error rate (19.375%), and type II error rate (38.750%) among the three models. The 10 variables chosen using the association rules effectively reduced the average training and testing times, but this model had poor prediction accuracy (71.292%), type I error rate (18.094%), and type II error rate (49.938%). The 6 variables chosen using the decision tree effectively reduced the average training and testing times and had the best prediction accuracy (85.625%), type I error rate (7.250%), and type II error rate (28.625%) among the three models. Table 3 shows that the performances of the decision tree, association rules, and benchmark models were significantly different. The decision tree model had the best performance, followed by the association rules and benchmark models. Tables 2 and 3 show that feature selection reduced both the training and testing times as well as interference from an excessive number of variables, improving the model accuracy, efficiency, and effectiveness [48, 49]. Moreover, the accuracy rate of 85.625% confirms that the six factors

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

extracted using the decision tree can be seen as key factors for determining whether auditors may face litigation according to the companies listed on the Taiwan stock exchange. In summary of the above, the model which used the decision tree for feature selection had the best performance. The key factors extracted were credit risk index4, stock price fluctuation, client importance_firm, accounts receivable ratio, auditor's signature lag, and auditor appointment.

Table 2 Feature selection tool performance assessment

Model	Number of features	Rate of extraction (%)	Accuracy (%)	Type I error rate (%)	Type II error rate (%)
Comparison model - Decision tree	6	17.949	85.625	7.250	28.625
Comparison model - Association rules	10	25.641	71.292	18.094	49.938
Benchmark model - All variables	39	100.000	74.167	19.375	38.750

Table 3 Comparison of feature selection tool performances

	Comparison model - Decision tree		Comparison model - Association rules		Benchmark model - All variables		Main effect	Comparison
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation		
Accuracy (%)	85.625	1.145	71.292	2.913	74.167	1.591	349.837***	Decision tree >Benchmark model >Association rules
Type I error rate (%)	7.250	1.998	18.094	3.724	19.375	1.444	166.834***	Decision tree <Association rules Benchmark model
Type II error rate (%)	28.625	1.844	49.938	2.571	38.750	4.134	314.533***	Decision tree <Benchmark model <Association rules

Note: Main effects in this table are F values; *p<0.05; **p<0.01; ***p<0.001.

Classification results

After analysis of the feature selection results, the six extracted features were used to create an early warning model for litigation against auditors and a decision table using a decision tree. These were then compared to logistic regression and discriminant analysis methods to provide auditors with a referencewhen selecting clients. The performance assessment results are shown in Table 4. Inspection of the AUC values for the three models revealed that all AUC were greater than 0.5, indicating that all three models have predictive value. The decision tree diagnostic model developed in this study had an accuracy of 92.708% and the type I and type II error rates were 1.563% and 18.750%, respectively. The performance of this model was superior to those created using logistic regression and discriminant analysis.

Table 4 Classification performance assessment

Model	Accuracy (%)	Type I error	Type II error	AUC value

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

		rate (%)	rate (%)	
Data mining model - Decision tree	92.708	1.563	18.750	0.915
Statistical model - Logistic regression	86.979	8.594	21.875	0.895
Statistical model - Discriminant analysis	83.333	14.844	20.313	0.824

Warning Model Decision Information

This study used a decision tree to create an early warning model for litigation against auditors. In addition to having better accuracy than logistic regression and discriminant analysis, the decision tree also provides a decision table that illustrates the relationships and rules between important variables and the risk of litigation. The reference created from the data mining results can be used as a supplementary tool when auditors evaluate litigation risk in order to reduce risk probability and prevent damage to the reputations of both the auditor and the accounting firm.

Table 5 shows the eight key rules for decision makers to assess the litigation risk. Rule 1 in the decision table indicates that when the credit risk index is 95, auditors have a high risk of facing litigation. Thus, when a company under investigation faces worsening operations and is labeled a high credit risk, a financial crisis may lead to a lawsuit from the investors against the management and auditors for compensation [9, 53, 54]. However, Rule 8 indicates that when the credit risk index is less than or equal to 6, the company's financials are more stable and credit risk and the risk of financial crisis are low; thus, auditors have a low risk of facing litigation.

Rules 2-7 in the decision table consider multiple variables in determining whether a client is high or low risk for litigation. Rule 2 indicates that when the credit risk index is 7 or 8 and the fluctuation in stock prices is less than 0.37, the company has a stable performance in the securities market as the stock prices have not caused large fluctuations; therefore, auditors have a low risk of facing litigation. Rule 3 indicates that when the credit risk index is 7 or 8 and the fluctuation in stock prices is greater than 0.37, if the client importance to the firm is greater than 0.52 (the client's fees account for no less than 52% of the firm's entire income), then auditors have a high risk of facing litigation. The main reason for this is that auditors may rely on economic factors and lose their independence [29]. Rules 4-7 indicate that when the credit risk index is 7 or 8, the fluctuation in stock prices is greater than 0.37, and the client importance to the firm is less than 0.52, then the risk of facing litigation depends on the accounts receivable, auditor’s signature lag, and auditor appointment. Rule 5 indicates that under the conditions above, if the accounts receivable rate is greater than 16.18 (i.e., the

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

accounts receivable account for over 16.18% of the total assets), and the auditor's signature lag is over 84.5 days, then auditors have a high risk of facing litigation. Rule 7 indicates that in the conditions above, if the auditor's signature lag is less than 84.5 days and the auditor has been appointed for less than 5 years, then auditors also have a high risk of facing litigation. In addition, Rules 4 and 6 indicate that auditors have a low risk of facing litigation. The main discrepancy between these rules is in the influences of auditor appointment and auditor's signature lag; when the client has a higher ratio of accounts receivable, if the accountant has been newly appointed and requires more time to complete the audit, this indicates that the auditor has little knowledge of the company, which increases the risk of litigation [9].

Table 5 Auditors' litigation risk decision table

	if						***then***	
Rule	Credit risk index	Stock price fluctuation	Client importance_firm	Accounts receivable ratio	Auditor's signature lag (days)	Auditor appointment	**High risk of litigation**	**Low risk of litigation**
1	9	–	–	–	–	–	✓	
2	7≤8	≤0.37	–	–	–	–		✓
3			>0.52	–	–	–	✓	
4		>0.37	≤0.52	≤16.18	–	–		✓
5				>16.18	>84.5	–	✓	
6					≤84.5	>5		✓
7						≤5	✓	
8	≤6	–	–	–	–	–		✓

5. CONCLUSION

After the recent series of fraud cases, investor protection mechanisms have urged the competent authority to revise the legal system in order to increase the legal liability of auditors and their financial statements, thus preventing fraud and litigation due to audit failure and improving audit quality and the degree of confidence in financial statements. Changes in the legal environment and laws have made it more difficult for auditors to assess whether to accept clients and to avoid indemnification and a damaged reputation due to litigation. The warning model and decision table for litigation against auditors developed in this study can serve as a reference for auditors. Prior literature regarding litigation risk against auditors was reviewed to collect affecting factors for both audit risk and the legal environment. Feature selection was then used to extract critical variables, after which, categorization techniques constructed a representative and convenient warning model and decision table for auditors' reference.

Two procedures were used to construct the warning model. First, 39 influencing variables

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

were collected from prior literature and a decision tree was used to select six key factors: credit risk index, stock price fluctuation, client importance_firm, accounts receivable ratio, auditor's signature lag, and auditor appointment. Compared to the extracted variables using other selection tools, the variables chosen using the decision tree had higher accuracy and lower type I and type II error rates; therefore, these variables can be seen as key influencing factors for auditors' litigation risk. Second, the six chosen factors were used in a decision tree to construct a litigation warning model. The results showed that the accuracy rate was 92.708% and the type I and type II error rates were below 10%; eight categorization rules were extracted which were compiled in a decision table.

This study contributed a new concept regarding the use of data mining technology to determine the influencing factors of auditors' litigation risk. This study reviewed relevant literature and collected all influencing factors. Then, key affecting factors were extracted using feature selection methods, effectively increasing the prediction accuracy of the model. The precision of the diagnostic model established using data mining technologies in this study was higher than that for models created using logistic regression and discriminant analysis, demonstrating the value of applying data mining in other relevant areas.

Because the sample used in this study consisted of companies listed on the Taiwan stock exchange, the legal liability of Taiwanese auditors was weaker than that of auditors in developed countries. However, the complexity of litigation against auditors is increasing. The factors extracted during feature selection can be seen as the key factors influencing the litigation risk for auditors in Taiwan and serve as a reference for other developing countries. Additionally, the objective of this study was to create a warning model for litigation against auditors. The difference between this and other studies predicting litigation risk against auditors was that this study did not only aim to improve model accuracy, but also to create an understandable classification model founded on rules for auditors to use when evaluating risk in order to help auditors screen potentially risky clients, thus improving audit quality and lowering the litigation risk.

Finally, while this study collected all influencing factors for auditors' litigation risk from previous literature, some other factors may have been left out or measured using different methods. Therefore, future studies can include other variables to construct a more complete and accurate warning model.

Reference

[1] Carcello, J.V. and Z.-V. Palmrose, Auditor Litigation and Modified Reporting on Bankrupt Clients. Journal of Accounting Research, 1994. 32: p. 1-30.

[2] Lee, C.-W.J. and Z. Gu, Low Balling, Legal Liability and Auditor Independence. The

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

Accounting Review, 1998. 73(4): p. 533-555.

[3] Liu, C.-W., T.-C. Wang, and S.-T. Lai, Litigation Risk and Large Audit Firms' Acceptable Level of Clients' Financial Risk. NTU Management Review, 2009. 20(1): p. 1-40.

[4] Guyon, I. and A. Elisseeff, An Introduction to Variable and Feature Selection. Journal of Machine Learning Research 2003. 3: p. 1157-1182.

[5] Arens, A.A., R.J. Elder, and M.S. Beasley, Auditing and Assurance Services. 15th ed. ed. 2014: Pearson.

[6] Palmrose, Z.-V. and S. Scholz, The Circumstances and Legal Consequences of Non-GAAP Reporting: Evidence from Restatements. Contemporary Accounting Research, 2004. 21(1): p. 139-180.

[7] Schmidt, J.J., Perceived Auditor Independence and Audit Litigation: The Role of Nonaudit Services Fees. The Accounting Review, 2012. 87(3): p. 1033-1065.

[8] Kaplan, S.E. and D.D. Williams, Do Going Concern Audit Reports Protect Auditors from Litigation? A Simultaneous Equations Approach. The Accounting Review, 2013. 88(1): p. 199-232.

[9] Pierre, K.S. and J.A. Anderson, An Analysis of the Factors Associated with Lawsuits against Public Accountants. The Accounting Review, 1984. 59(2): p. 242-263.

[10] Kreutzfeldt, R.W. and W.A. Wallace, Error Characteristics in Audit Populations: Their Profile and Relationship to Environmental Factors. Auditing: A Journal of Practice & Theory, 1986. 6(1): p. 20-43.

[11] Kinney, W.R. and L.S. McDaniel, Characteristics of Firms Correcting Previously Reported Quarterly Earnings. Journal of Accounting and Economics, 1989. 11(1): p. 71-93.

[12] McKeown, J.C., J.F. Mutchler, and W. Hopwood, Towards an Explanation of Auditor Failure to Modify the Audit Opinions of Bankrupt Companies. Auditing: A Journal of Practice & Theory, 1991. 10: p. 1-13.

[13] Summers, S.L. and J.T. Sweeney, Fraudulently Misstated Financial Statements and Insider Trading: An Empirical Analysis. The Accounting Review, 1998. 73(1): p. 131-146.

[14] Bonner, S.E., Z.-V. Palmrose, and S.M. Young, Fraud Type and Auditor Litigation: An Analysis of SEC Accounting and Auditing Enforcement Releases. The Accounting Review, 1998. 73(4): p. 503-532.

[15] Cahan, S.F. and W. Zhang, After Enron: Auditor Conservatism and Ex-Andersen Clients. The Accounting Review, 2006. 81(1): p. 49-82.

[16] Beasley, M.S., An Empirical Analysis of the Relation between the Board of Director Composition and Financial Statement Fraud. The Accounting Review, 1996. 71(4): p. 443-465.

[17] Schwartz, K.B. and B.S. Soo, The Association Between Auditor Changes and Reporting Lags. Contemporary Accounting Research, 1996. 13(1): p. 353-370.

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

[18] Knechel, W.R., V. Naiker, and G. Pacheco, Does Auditor Industry Specialization Matter? Evidence from Market Reaction to Auditor Switches. Auditing: A Journal of Practice & Theory, 2007. 26(1): p. 19-45.

[19] Titman, S. and B. Trueman, Information Quality and the Valuation of New Issues. Journal of Accounting and Economics, 1986. 8(2): p. 159-172.

[20] Lev, B., S. Ryan, and M. Wu, Rewriting Earnings History. Review of Accounting Studies, 2008. 13(4): p. 419-451.

[21] DeAngelo, L.E., Auditor Size and Audit Quality. Journal of Accounting and Economics, 1981. 3(3): p. 183-199.

[22] Palmrose, Z.-V., 1987 Competitive Manuscript Co-Winner: An Analysis of Auditor Litigation and Audit Service Quality. The Accounting Review, 1988. 63(1): p. 55-73.

[23] Krishnan, J. and J. Krishnan, Litigation Risk and Auditor Resignations. The Accounting Review, 1997. 72(4): p. 539-560.

[24] Kim, J.-B., R. Chung, and M. Firth, Auditor Conservatism, Asymmetric Monitoring, and Earnings Management. Contemporary Accounting Research, 2003. 20(2): p. 323-359.

[25] Craswell, A.T., J.R. Francis, and S.L. Taylor, Auditor Brand Name Reputations and Industry Specializations.Journal of Accounting and Economics, 1995. 20(3): p. 297-322.

[26] Solomon, I., M.D. Shields, and O.R. Whittington, What Do Industry-Specialist Auditors Know? Journal of Accounting Research, 1999. 37(1): p. 191-208.

[27] Beasley, M.S. and K.R. Petroni, Board Independence and Audit-Firm Type. Auditing: A Journal of Practice & Theory, 2001. 20(1): p. 97-114.

[28] Gramling, A.A. and D.N. Stone, Audit Firm Industry Expertise: A Review and Synthesis of the ArchivalLiterature. Journal of Accounting Literature, 2001. 20: p. 1-29.

[29] Reynolds, J.K. and J.R. Francis, Does Size Matter? The Influence of Large Clients on Office-Level Auditor Reporting Decisions. Journal of Accounting and Economics, 2000. 30(3): p. 375-400.

[30] Chi, W.-C., H.-C. Yu, and S.-T. Chiu, Mandatory Rotation and Auditor Independence-An Analysis of Auditor's Reputation Effect. Taiwan Accounting Review, 2004. 5(1): p. 71-104.

[31] Johnson, E., I.K. Khurana, and J.K. Reynolds, Audit-Firm Tenure and the Quality of Financial Reports.Contemporary Accounting Research, 2002. 19(4): p. 637-660.

[32] DeFond, M.L. and K.R. Subramanyam, Auditor Changes and Discretionary Accruals. Journal of Accounting and Economics, 1998. 25(1): p. 35-67.

[33] Lee, J.-Z. and H.-F. Lin, The Relations between Auditor Tenure and Abnormal Accruals. Management Review, 2005. 24(4): p. 103-126.

[34] Liu, C.-W. and T.-C. Wang, Going Concern Opinions: before and after SAS No.33. Journal of management, 2005. 22(4): p. 525-549.

[35] Palmrose, Z.-V., Who Got Sued? Journal of Accountancy 1997. 183(3): p. 55-73.

[36] Lin, C.-C. and H.-L. Lin, Auditor's Liability for Financial Statement Fraud in Taiwan

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

from Li-Bar Verdict: An Empirical Analysis. National Taiwan University Law Journal, 2010. 39(9): p. 223-288.

[37] Beaver, W.H., Financial Ratios As Predictors of Failure. Journal of Accounting Research, 1966. 4: p. 71-111.

[38] Altman, E.I., Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. The Journal of Finance, 1968. 23(4): p. 589-609.

[39] Kirkos, E., C. Spathis, and Y. Manolopoulos, Data Mining Techniques for the Detection of Fraudulent Financial Statements. Expert Systems with Applications, 2007. 32(4): p. 995-1003.

[40] Ohlson, J.A., Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research, 1980. 18(1): p. 109-131.

[41] Frawley, W.J., G. Piatetsky-Shapiro, and C.J. Matheus, Knowledge Discovery in Databases: An Overview. AI Magazine, 1992. 13(3): p. 57-70.

[42] Varetto, F., Genetic Algorithms Applications in the Analysis of Insolvency Risk. Journal of Banking & Finance, 1998. 22(10–11): p. 1421-1439.

[43] Cristianini, N. and J. Shawe-Taylor, Support Vector Machines. 2000: Cambridge: Cambridge University Press.

[44] Min, J.H. and Y.-C. Lee, Bankruptcy Prediction Using Support Vector Machine with Optimal Choice of Kernel Function Parameters. Expert Systems with Applications, 2005. 28(4): p. 603-614.

[45] Odom, M.D. and R. Sharda. A Neural Network Model for Bankruptcy Prediction. in 1990 IJCNN International Joint Conference. 1990. San Diego, CA, USA.

[46] Coats, P.K. and L.F. Fant, Recognizing Financial Distress Patterns Using a Neural Network Tool. Financial Management, 1993. 22(3): p. 142-155.

[47] Chaveesuk, R., C. Srivaree-Ratana, and A.E. Smith, Alternative Neural Network Approaches to Corporate Bond Rating. Journal of Engineering Valuation and Cost Analysis, 1999. 2(2): p. 117-131.

[48] Questier, F., et al., The Use of CART and Multivariate Regression Trees for Supervised and Unsupervised Feature Selection. Chemometrics and Intelligent Laboratory Systems, 2005. 76(1): p. 45-54.

[49] Sugumaran, V., V. Muralidharan, and K.I. Ramachandran, Feature Selection Using Decision Tree and Classification through Proximal Support Vector Machine For Fault Diagnostics of Roller Bearing. Mechanical Systems and Signal Processing, 2007. 21(2): p. 930-942.

[50] Smith, K.A. and J.N.D. Gupta, Neural Networks in Business: Techniques and Applications for the Operations Researcher. Computers & Operations Research, 2000. 27(11–12): p. 1023-1044.

[51] Olafsson, S., X. Li, and S. Wu, Operations Research and Data Mining. European Journal of Operational Research, 2008. 187(3): p. 1429-1448.

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang

[52] Tam, K.Y. and M.Y. Kiang, Managerial Applications of Neural Networks: The Case of Bank Failure Predictions. Management Science, 1992. 38(7): p. 926-947.
[53] Palmrose, Z.-V., Litigation and Independent Auditors: The Role of Business Failures and Management Fraud.Auditing: A Journal of Practice & Theory, 1987. 6(2): p. 90-103.
[54] Lys, T. and R.L. Watts, Lawsuits against Auditors. Journal of Accounting Research, 1994. 32: p. 65-93.

©copyright by Tzong-Huei Lin, Yu-Hsin Lu, Yu-Cheng Lin, Yun-Hsuan Wang