An Application Of A Data Mining Technique For Assessing The Risk Of Management Fraud

Ashutosh Deshmukh, (E-mail: avd1@psu.edu), Penn State, Erie
Lakshminarayan Talluru, (E-mail: lxt6@psu.edu), Penn State, Erie

Abstract

Data mining techniques identify relationships, patterns, trends, and predictive information from large and complex databases. This study demonstrates the use of a data mining technique to assess the risk of management fraud. We use a data mining tool to analyze the management fraud data, presence or absence of red flags in fraud and no fraud cases, collected by a Big Six firm. The ensuing results compare favorably with the statistical and neural network results obtained by the other studies. The study illustrates the ease of using data mining techniques by demonstrating the rapid development of models, querying capabilities, and ease of encoding statistical models in audit decision making.

Introduction

The responsibility of an independent auditor concerning management fraud remains a controversial issue. The newly issued Statement on Auditing Standard (SAS) No. 82 "Consideration of Fraud in a Financial Statement Audit," which supersedes SAS No. 53, continues to require the auditor to plan and perform the audit to provide a reasonable assurance that the financial statements are free of management fraud. The standard focuses on the risk of material misstatements due to fraud and provides risk factors that the auditor should consider during the audit.

Readers with comments or questions are encouraged to contact the authors via e-mail.

Traditionally auditors have utilized personal, business, and economic risk factors (referred to as red flags) in assessing the risk of management fraud and planning the audit. Red flags are potential symptoms existing within a company's business environment that would indicate a higher risk of an intentional misstatement of financial statements (Price Waterhouse, 1985). SAS Nos. 6, 16, 17, 53, and 82 provide discussion of red flags. SAS No. 82 cites several examples of red flags such as excessive interest by management to maintain or increase the entity's stock price or earnings trend through the use of unusually aggressive accounting practices, high turnover of management, counsel, or board members, and overly complex organizational structures involving numerous or unusual legal

entities, managerial lines of authority, or contractual lines of authority without apparent business purpose.

The purpose of this study is to demonstrate the use of a data mining technique to assess the risk of management fraud. Data mining techniques discover existing relationships, patterns, trends, and predictive information from large and complex databases. Data mining is data driven, not user driven or verification driven (Mena, 1996); data mining tools discover previously unknown relationships by working bottom-up on the data set. We use a data mining technique to analyze the management fraud data collected by a Big Six firm. This database contains information about the presence or absence of red flags in fraud and no fraud cases. The results of data mining are then compared with the traditional statistical and neural network results obtained by the other studies. The results compare favorably with these methods. The data mining tool is very user friendly, has query capabilities to assess the risk of management fraud for a given audit engagement, and the resulting statistical models can be easily incorporated as stand alone decision aids. These characteristics make data mining tools efficient and effective in practical audit applications.

The rest of the paper is as follows. First, empirical research based on red flags will be reviewed. Secondly, the data mining technique that utilizes statistical networks and its application to the analysis of management fraud data will be discussed. Finally, conclusions of the research, the costs and benefits, and merits of utilizing data mining approach to aid audit decision making will be discussed.

Literature Review

Loebbecke and Willingham (1988) proposed a causal model to assess the likelihood of management fraud. The model was specified as follows:

$$P(MI) = f(C, M, A)$$

where; P (MI) is the probability of management fraud (or material irregularities), C is the degree to which conditions are such that a material management fraud can be committed, M is the degree to which person or persons in positions of authority and responsibility in the entity have a reason or motivation to commit management fraud, and A is the degree to which the person or persons in the positions of authority and responsibility in the entity have an attitude or set of ethical values such that they would allow themselves (or even seek) to commit management fraud, and where, if C, or M, or A = 0, then $P(MI) = 0$. Thus, this model posits that if one of the requirements C, M, or A is not present, then the likelihood of management fraud is zero.

Loebbecke and Willingham categorized the red flags mentioned in SAS No. 53 according to the model's three components. Then an analysis of 71 known cases of management fraud as reported in the Securities and Exchange Commission (SEC) Accounting and Auditing Releases (AAER) was performed. The frequency of the red flags as mentioned in SAS No. 53 was noted and some of the red flags present in the fraud cases but not mentioned in SAS No. 53 were identified and added to the model. Then, based on the frequency analysis, all the prominent red flags were labeled as "primary indicators," and the other red flags were labeled as "secondary indicators." This modified model was then reapplied to 51 AAER fraud cases having a comprehensive presentation that permitted a meaningful analysis. The model was found to be a good fit and the authors concluded that the continued development of this model was merited.

Loebbecke, Eining, and Willingham (1989) extended the above research by conducting a survey of KPMG Peat Marwick partners who had encountered management fraud. The survey sought detailed information about management fraud incidents in terms of frequency of occurrence, nature, and impact on the client's financial statements. The survey resulted in an addition of new red flags, and deletion, reclassification, and recombination of some of the red

flags in the model. The authors then presented a revised fraud assessment model that incorporated red flags, classified as “primary” and “secondary” indicators for the model’s three components, as identified by the survey of the KPMG Peat Marwick partners and AAER cases. The resultant red flags are presented in Table 1.

Bell et al. (1993) expanded the above database by collecting data from engagements where no fraud was discovered. The questionnaire used by Bell et al. in collecting this data is shown in Table 2.

Bell et al. (1993), Fanning et al. (1995), Hansen et al. (1996), and Bell et al. (1996) analyzed this enhanced database. These studies used

logit, neural networks, generalized qualitative response model, and logit, respectively. We briefly discuss the results of these studies.

Bell et al. (1993) primarily used a cascaded logit model estimation approach. The authors also tested the stand-alone and combined discriminatory power of red flags. The individual red flags were tested by 2X2 contingency tables and the related χ^2 for independence. The results indicated that 8 out of 21 factors in SAS No. 53 and 14 of the 25 additional factors identified by the Loebbecke et al. were significant in identifying fraud firms. However, in this approach red flags that are correlated can be identified as significant, but these red flags may not be incrementally significant. This approach also ig-

Table 1: Risk Assessment Model Applicable to Management Fraud

Adapted from Loebbecke, Eining, and Willingham (1989).

$$P(MI) = f(C, M, A)$$

where, **C** is the degree to which conditions are such that a material fraud can be committed, **M** is the degree to which person(s) in the positions of authority and responsibility in the entity have a reason or motivation to commit management fraud, and **A** is the degree to which person(s) in the positions of authority and responsibility in the entity have attitude or set of ethical values such that they would allow themselves (or even seek) to commit management fraud.

Primary indicators of C are: <ul style="list-style-type: none"> - Dominated decisions - Major transactions - Related party - Weak internal control - Difficult-to-audit transactions Secondary indicators of C are: <ul style="list-style-type: none"> - Significant judgments - High management turnover - Decentralized organization - Assets subject to misappropriation - New client - Rapid growth - Inexperienced management - Conflict of interest 	Primary indicators of M are: <ul style="list-style-type: none"> - Industry decline - Inadequate profits - Emphasis on earnings projections - Significant contractual commitments Secondary indicators of M are: <ul style="list-style-type: none"> - Rapid growth - Rapid industry change - Sensitive operating results - Incentive compensation - Adverse legal circumstances - Significant portion of management’s wealth - Management’s job threatened 	Primary indicators of A are: <ul style="list-style-type: none"> - Dishonest management - Emphasis of earnings projections - Personality anomalies - Prior year irregularities - Lies or evasiveness - Aggressive attitude toward financial reporting Secondary indicators of A are: <ul style="list-style-type: none"> - Weak internal control - Conflict of interest - Poor reputation - Frequent disputes with auditor - Undue pressure on auditor - Disrespectful attitude
--	--	---

Table 2: Questionnaire Used for Management Fraud Evaluation

- Q1. Are management operating and financial decisions dominated by a single person or a few persons acting in concert?
- Q2. Does management place undue emphasis on meeting earning projections or other quantitative targets?
- Q3. Have managers recently entered into collusion with outsiders?
- Q4. Does your experience with management indicate a degree of dishonesty?
- Q5. Does management display a propensity to take undue risks?
- Q6. Does management display significant disrespect for regulatory bodies?
- Q7. Have managers lied to the auditors or been overly evasive in responses to audit inquiries, or have they shown some other indications of dishonesty?
- Q8. Do client personnel display significant resentment of authority?
- Q9. Is management's attitude toward financial reporting unduly aggressive?
- Q10. Do key managers exhibit strong personality anomalies?
- Q11. Is this a new client?
- Q12. Is there a need to cover up an illegal act?
- Q13. Does the client have a weak control environment?
- Q14. Are there frequent and significant difficult-to-audit transactions or balances?
- Q15. Is the client a public company?
- Q16. Is a significant amount of judgment involved in determining the total of account balance or class of transactions?
- Q17. Have there been instances of material management fraud in prior years?
- Q18. Is the client's organization decentralized without adequate monitoring?
- Q19. Is the client in a period of rapid growth?
- Q20. Does the client have solvency problems?
- Q21. Does a conflict of interest exist involving the client entity and/or its personnel?
- Q22. Do accounting personnel exhibit inexperience or laxity in performing their duties?
- Q23. Is the client confronted with adverse legal circumstances?
- Q24. Is the client's profitability relative to its industry inadequate or inconsistent?

Additional Questions

- 1. P2. Has the client entered into one or a few specific transactions that have a material effect on the financial statements?
- 2. P3. Has the client entered into a significant transaction(s) with one or more related parties?
- 3. P6. Is management and/or key accounting personnel turnover high?
- 4. P11. Is management inexperienced?
- 5. P17/37. Is the client currently or was the client recently involved in a purchase, sale, or merger transaction with another company?
- 6. P18. Has the company recently entered into a significant number of acquisition transactions?
- 7. P21. Is the direction of change in the client's industry declining with many business failures?
- 8. P22. Is the client's industry in a state of rapid change?
- 9. P24. Are the client's operating results highly sensitive to economic factors (inflation, interest rates, unemployment, etc.)?
- 10. P26. Does a substantial portion management compensation depend on meeting quantified targets?
- 11. P29. Are there adverse conditions in the client's industry or external environment?
- 12. P31. Is the client subject to significant contractual commitments?
- 13. P33. Is a significant portion of management's personal wealth in the form of holdings in the client entity?
- 14. P36. Does management perceive their job is threatened by poor performance?
- 15. P39. Does management exhibit undue concern with the need to maintain or improve the reputation/image of the entity?
- 16. P41. Do managers appear to engage in an inappropriate lifestyle, to have personal financial difficulties, or to live beyond their means?
- 17. P42. Is management's reputation in the business community poor?
- 18. P47. Has management engaged in frequent disputes with the auditors, particularly about aggressive application of accounting principles that increase earnings?
- 19. P48. Does management place undue pressure on the auditors, particularly through the fee structure or the imposition of unreasonable deadlines?
- 20. P49. Has the client engaged in opinion shopping?
- 21. P50. Do managers display a hostile attitude toward the auditors?
- 22. P51. Are key managers considered highly unreasonable?
- 23. P52. Do key managers display a significant lack of strength of character?

nores the possible interactions among red flags.

Bell et al. also tested the conceptual model developed by Loebbecke et al. by applying different statistical techniques. First, the authors evaluated the model components (C, M, and A) using primary risk factors and then secondary risk factors. Here also bivariate tests were conducted using 2X2 contingency tables. The results indicated that the relations between primary components and occurrence of fraud were significant for all components. In case of secondary risk factors the "Motivation" factor was not significant. However, in the examination of primary and secondary factors, "Attitude" had the strongest relationship. Second, when all the risk factors were combined all the factors were significant. Third, if only significant red flags as identified earlier were used then all the factors show even stronger statistical significance. However, in these models the individual factors could not be allocated different weights and the output was dichotomous.

Finally, the authors used cascaded logit models to overcome the above mentioned limitations. The authors used only significant red flags in the logit models. Three separate logit models were estimated for each model component and fourth logit model was estimated to combine output of the earlier three models. The "Attitude risk model" had the strongest relationship. The predictive ability of the model was very good. Bell et al. (1996) summarized the above study and again using a logit analysis created a decision aid for the auditor to assess the risk of management fraud. However, results of this study should be interpreted cautiously since the authors did not include "Additional Questions" identified in Table 2 in their analysis.

Fanning et al. (1995) used neural networks to analyze the same data. These researchers applied two types of self-generating artificial neural networks, generalized adaptive neural network architecture and adaptive logic network architecture. The authors discuss design issues related with neural networks, time required to manipulate data and process it, and the various

cut-off methods used to calculate results. The analysis yielded a more parsimonious set of red flags that may be useful in detecting management fraud as compared to Bell et al. study. The significance of condition, motivation, or attitude is not assessed; however, the study concentrates on individual red flags. The authors note that many red flags underlie the attitude factor and as such the results are consistent with the Bell et al. study. The statistical results are somewhat superior to the Bell et al. (1993) study.

Hansen et al. (1996) also analyzed the same data set using generalized qualitative response model (EGB2). The EGB2 model subsumes probit and logit models as special cases. EGB2 model also allows inclusion of asymmetric costs of classification errors. For example, costs of type I error (the audit indicates management fraud given no management fraud) are generally less than costs of type II error (the audit indicates no management fraud given management fraud). Hansen et al. analyzed fraud data under differing assumptions of symmetric and asymmetric costs. The authors, similar to Fanning et al., did not assess the significance of condition, motivation, or attitude but looked at individual red flags. The statistical results are not superior to logit or neural network models. But EGB2 allows inclusion of costs of type I and type II errors, which the authors contend is an advantage in an audit situation.

The set of red flags identified as significant in each study is different and sometimes conflicting. However, many significant red flags are common to all studies. The significant red flags discovered in each study are summarized in Table 3. The statistical results for hit rates of each study are also different and these are summarized in Table 4.

The statistical results of various studies indicate that the data set is complex and depending on the method of analysis, provides different red flags as significant in assessing the risk of management fraud. The statistical results of the analyses should be interpreted cautiously. Different studies have different training and holdout

Table 3: Comparison of Results: Significant Red Flags Found in Earlier Studies

Bell, Szykowsky, and Willingham (1993)	Fanning, Cogger, and Srivastava (1995)	Hansen, McDonald, Messier, and Bell (1996)	Bell, Carcello, and Willingham (1996)
<ol style="list-style-type: none"> 1. Dominated decisions 2. Weak Internal Control 3. Difficult-to-audit transactions 4. Contentious accounting issues 5. Decentralized organization 6. Rapid growth 7. Conflict of interest 8. Ownership (public vs. private) 9. Inadequate Profitability 10. Undue emphasis on earnings 11. Adverse legal circumstance 12. Solvency problems 13. Rapid growth 14. Collusion with outsiders 15. Need to cover up an illegal act 16. Management dishonesty 17. Undue emphasis on earnings 18. Management risk taking attitudes 19. Strong personality anomalies 20. Misstatements in prior periods 21. Management evasiveness 22. Aggressive attitudes 23. Disrespect for regulatory authorities 	<ol style="list-style-type: none"> 1. Does management place undue emphasis on meeting earnings projections or other quantitative targets? 2. Does management display significant disrespect for regulatory authorities? 3. Do client personnel display significant resentment of authority? 4. Do key managers exhibit strong personality anomalies? 5. Does the client have a weak internal control environment? 6. Are there frequent and significant difficult-to-audit transactions or balances? 7. Is the client confronted with adverse legal circumstances? 8. Is the direction of change in the client's industry declining with many business failures? 9. Does the substantial portion of management compensation depend on meeting quantified targets? 10. Is a significant portion of management's personal wealth in the form of holdings in the client entity? 11. Are key managers considered highly unreasonable? 	<ol style="list-style-type: none"> 1. Is there any need to cover up an illegal act? 2. Does your experience with management indicate a degree of dishonesty? 3. Is the client's organization decentralized without adequate monitoring? 4. Are there significant and difficult-to-audit transactions or balances? 5. Is a significant amount of judgment involved in determining the total of an account balance or class of transactions? 6. Does the client have solvency problems? 7. Does management display a propensity to take undue risks? 8. Is the client a public company? 9. Do client personnel display significant resentment of authority? 10. Does management display significant resentment of regulatory bodies? 11. Is this a new client? 12. Does a conflict of interest exist involving the client entity and/or its personnel? 13. Do key managers exhibit strong personality anomalies? 14. Have managers recently entered into collusion with outsiders? 15. Do accounting personnel exhibit inexperience or laxity in performing their duties? 16. Is management attitude toward financial reporting unduly aggressive? 17. Does management place undue emphasis on meeting earnings projections or other quantitative targets? 18. Is the client confronted with adverse legal circumstances? 	<ol style="list-style-type: none"> 1. Weak internal control environment 2. Rapid company growth 3. Inadequate or inconsistent relative profitability 4. Management places undue emphasis on meeting earnings projections 5. Management lied to the auditors or was overly evasive 6. The ownership status (public vs. private) 7. Interaction term between a weak internal control environment and an aggressive attitude toward financial reporting

Table 4: Comparison of Best Hit Rates for Holdout Samples

	Fraud	No Fraud	Total
Bell et al. (1993) Logit - Cut off -0.4	70%	91%	87%
Fanning et al. (1995) Neural networks - Cut off - 0.4	75%	91%	89%
Hansen et al. (1996) EGB2			
Symmetric costs	63%	95.5%	89.5%
Asymmetric costs (Costs of type II error 10 times higher than costs of type I error)	88.6%	84.5%	85.3%

samples and cut-off points. For example, Hansen et al. use a statistical method which repeatedly uses (jackknife) the same data set for estimating the model. Bell et al. and Fanning et al. use a 0.40 (if probability of management fraud is greater than 0.40 then the risk of management fraud is high) cut-off point for their best results. The training and holdout sample size used in this study are different than Bell et al. and Fanning et al. studies. To overcome these methodological differences we report only the *best* results obtained under each method.

Data Mining Techniques

Fayyad et al. (1996) advocate the term "knowledge discovery in databases" to refer to overall process of discovering useful knowledge from data. Data mining then becomes a particular step in the overall process - applying statistical and mathematical algorithms to discover existing patterns in the data. Data mining is a growing field that is driven primarily by the database explosion and availability of powerful computers. Data mining provides for the identification of relationships, patterns, trends, and predictive information that exists in large and complex databases. Data mining tools use various types of processing algorithms. The primary four methods are associations, clustering, classification, and sequential patterns (Bischoff and Alexander, 1997). These four methods can be further grouped; for example, decision trees, rule discovery, neural networks, rough sets, etc

can achieve classification. The Internet (<http://www.kdnuggets.com/siftware.html>) lists numerous companies producing software (software to sift through databases) using the above approaches.

The management fraud database analyzed in this paper is fairly large and complex. The database consists of 47 red flags or questions and 382 companies. The red flags are independent variables and the dependent variable is existence/nonexistence of fraud. All variables are binary: one indicates the existence of red flag (management fraud) and zero indicates the non-existence of red flag (management fraud). The complexity of this database and the absence of a priori hypotheses mean that data mining can be utilized in analyzing this database.

The data mining tool used for the analysis of this data is ModelQuest (MQ) Expert. MQ Expert is one data mining tool among a family of PC based data mining tools (MQ Miner, MQ Prospector, and MQ Enterprise) developed by AbTech Corporation. MQ Expert is a higher end product capable of handling 100 variables and 32,000 examples. System requirements for MQ Expert are relatively modest. The minimal configuration is an IBM PC or compatible (486 or above, a math-coprocessor is necessary), MS Windows or NT, 8MB of RAM, MS-Windows compatible mouse, and a CD-ROM drive. MQ Expert requires at least 20MB space on a hard drive. As the data sets to be analyzed become

larger the hard disk space required also becomes larger. The software will of course run faster on a higher performance computer. MQ tools have been applied to a variety of problems in military and industry. The illustrative applications of MQ tools are direct marketing, credit card fraud detection, stock market prediction, risk analysis, fault detection, and analysis of clinical data.

The processing algorithms, utilized by MQ Expert, are based on concepts drawn from statistics, neural networks, and expert systems. By integrating these concepts AbTech has developed proprietary statistical networks, which are at the heart of the data processing algorithms in MQ Expert. Statistical networks are similar to neural networks, however, the nodes in the statistical network represent mathematical functions. Each node accepts a set of inputs and using mathematical functions calculates the output. There are seven different types of nodes (or elements) such as singles, doublets, triplets, linear elements, normalizers, and unitizers, which are differentiated by the type of algebraic form of mathematical functions represented by that node.

Singles, doubles, and triplets have one, two, and three input variables respectively, these three elements are represented by third degree polynomial equations and the doublets and triplets have cross-terms, allowing for interaction among the node input variables. Linear elements linearly weigh the input variables for the first layer, normalizers transform all input variables using mean-sigma normalization, and unitizers transform the model outputs for training the model. The output of the elements along with the original input variables is fed into subsequent layers. The training of the model continues from layer to layer until the model ceases to improve based on the predicted squared error criterion (ModelQuest Expert Manual, 1996).

The MQ Expert manual claims that statistical networks are effective because mathematical functions can compactly capture numeric knowledge and that networks can deal with complex problems by subdividing them into smaller problems. The manual also states that a system

of one million production rules can be easily represented by a function of only six variables.

MQ Expert trains inductively by using a set of examples. MQ Expert handles numeric parameters such as expert judgments, probabilities, fuzzy values, prices, costs, sensor readings, failure rates, etc. Models are formulated and evaluated automatically by the program and the resulting network is implemented as a layered network of feed-forward functional elements. The functional element coefficients, number of network elements, types of node, and connectivity are developed automatically from the data. A variety of statistical evaluations are available with the trained model, which we found very useful especially when comparing MQ Expert with traditional statistical methods and neural networks. Additionally, MQ Expert makes no independence and distribution assumptions.

Building a model with MQ Expert then involves four steps: importing the data, training the model, evaluating the model performance, and implementing the model. MQ Expert has an import utility, which allows data to be imported from an Excel spreadsheet. MQ Expert splits the data into two groups: 75% for training and 25% for evaluation (holdout sample) purposes. The model was thus trained on 287 observations and 95 observations were used for evaluation purposes.

Once you split the data, you can train the model using the Train tab. The initial parameters such as complexity penalty multiplier, number of layers, and layer size are at default values. The term complexity penalty multiplier (CPM) needs some explanation. The CPM monitors the trade-off between complexity of the model and accuracy of the model. The higher value of CPM will encourage formation of simpler networks that are easily generalizable. The lower value will result in complex networks that may use many input variables and may overfit the data. MQ Expert also provides a utility that can be used to calculate an optimum value of CPM, which was used in the analysis. This utility saves a great deal of time as compared to neural

networks where derivation of an optimal network structure requires trial and error approach. The training and model building which uses statistical networks is automatic. When the training is going on the screen is split in three sections as shown in Figure 1.

The statistical network created by MQ Expert considers ten red flags as important, which is the most parsimonious set of red flags identified by any method. These red flags are: (1) management's undue emphasis on meeting earnings projections, (2) managers lied to audi-

Figure 1: The Statistical Network Training

The left side window shows the current best network found by the MQ Expert, the right side top window shows the hypothesized network (that will be compared with the current best network), at the bottom of the right window statistics describing the current best network are available. The training process is extremely fast; creating a statistical network takes approximately two minutes.

The trained network can be viewed by using a View tab, which is shown in Figure 2. The relationship between the inputs and output is shown in the trained network, if you click at node you can view relevant statistics or the equation for that part of the polynomial. The view of the statistical network created by MQ Expert based on the data is shown in Figure 2.

tors or been overly evasive in response to audit inquiries or have shown some other indications of dishonesty, (3) client personnel display significant resentment of authority, (4) weak internal control environment, (5) frequent and significant difficult to audit transactions or balances, (6) public vs. private ownership, (7) period of rapid growth, (8) client's profitability relative to its industry inadequate or inconsistent, (9) client involved in a purchase, sale, or merger transaction with another company, and (10) management exhibits undue concern with the need to maintain or improve the reputation/image of the entity. MQ Expert considers these red flags as important. However, considering these red flags in isolation will not enable us to assess the risk of management fraud. The relationships and interactions among these red flags need to be cap-

Figure 2: The Statistical Network Created by MQ Expert

tured.

MQ Expert provides a way to represent a relationship of these variables with the output (DEP - existence or nonexistence of management fraud) as shown in Figure 2. The statistical network shows the mathematical relationships among the significant red flags. As you can see, all the elements in the statistical network are triplets, and the network consists of the ten significant red flags mentioned earlier.

The network then can be evaluated against the evaluation data. Specify the data source (the holdout sample, which is created earlier) and click the Apply button. The trained network is applied to a new model. The results show two windows: one provides evaluation statistics and the other window provides a histogram of error. The evaluation statistic window shows descriptive statistics for evaluation and training data, maximum absolute error, average absolute error, average squared error, and R^2 . The MQ Expert manual states that average absolute error and maximum absolute error are important in determining whether you have a good

model. If the average absolute error is low compared to the output maximum and minimum then you have a good model. The maximum absolute error indicates the worst performance of the model. The summary statistics for training set and the histogram of that statistics is shown in Figure 3. The summary statistics for evaluation set and the histogram for that statistics are shown in Figure 4.

Since the nature of the data set is binary it is difficult to interpret the statistics generated by MQ Expert. The average absolute error for the training sample is 0.0827 and for the evaluation sample is 0.1223. The average absolute error is low compared to output maximum (one) but no such conclusion can be drawn about output minimum (zero). The R^2 for the training sample is 0.7442 and for the evaluation sample is 0.5967, which indicates fairly strong explanatory power of the statistical network.

In the histogram the horizontal axis is the number of observations and vertical axis is the range of errors. Each error range is labeled to the left of the colon and to the right of the colon

Figure 3: The Summary Statistics for the Training Set

Figure 4: The Summary Statistics for the Evaluation Set

is the number of observations that fall within this error range. If the errors are normally distributed then statistical network is a good fit for the data. Figures 3 and 4 both indicate approximate normal distribution for the errors.

The ultimate test of a model is actual prediction performance. The performance of the statistical network can be summarized in Table 5.

Table 5: Results using MQ Expert

	Holdout Sample		Training Sample	
	Hit Rate	Miss Rate	Hit Rate	Miss Rate
Fraud	72.63%	27.27%	80.36%	19.64%
No Fraud	98.63%	1.37%	98.69%	1.31%
Total	92.70%	7.30%	95.13%	4.87%

Probability Ranges

Holdout Sample					
	1-0.75	0.75-0.50	0.50-0.25	0.25-0	Total
Fraud	16	0	1	5	22
No Fraud	1	0	5	67	73
Total					95
Training Sample					
	1-0.75	0.75-0.50	0.50-0.25	0.25-0	Total
Fraud	37	7	4	7	55
No Fraud	0	3	10	219	232
Total					287

The cut-off point is 0.50, that is, if MQ Expert shows a probability of 0.50 or more then the case is considered a fraud case and if MQ Expert shows a probability of less than 0.50 then the case is considered a no fraud case. The hit rate for evaluation sample for fraud (no fraud) cases was 72.63% (98.63%) and miss rate was 27.27% (1.37%). The hit rate for training sample for fraud (no fraud) cases was 80.36%

(98.69%) and miss rate was 19.64% (1.31%). We further analyze the probability shown by MQ Expert in four ranges, 1 to 0.75, 0.75 to 0.5, 0.5 to 0.25, and 0.25 to 0. Table 5 indicates that in case of an evaluation sample all the correctly identified fraud cases have a probability greater than 0.75. For no fraud cases only one case is incorrectly identified and 67 (of the 72) correctly identified cases have a probability of 0.25 or less. In case of the training sample 37 (of the 44) correctly identified fraud cases have a probability of 0.75 or greater. For no fraud cases only three cases are incorrectly identified and 219 (of the 229) cases have a probability of 0.25 or less. These results are fairly stable and do not significantly change according to a cut-off point. For example, if the cut-off point is 0.40 the results will not change at all, if the cut-off point is 0.30 there is a change of couple of percentage points. However, best results are obtained under a cut-off point of 0.50 (or 0.40) and these are reported.

These results favorably compare to the neural network and statistical methods. The hit rate of MQ Expert in case of fraud for a holdout or evaluation sample is 72.63%, which is better than logit and EGB2 (symmetric costs) but slightly worse than neural networks. However, hit rate of MQ Expert in case of no fraud cases (and an overall hit rate) is better than all other methods. EGB2 has a better hit rate under asymmetric costs but at the cost of a low hit rate for no fraud cases. MQ Expert obtains similar results if the cut-off point is lowered to 0.10. These results indicate that the statistical network created by MQ Expert can be deployed in the field.

Another important feature of MQ Expert is query facility. The query screen is shown in Figure 5. As can be seen the red flags used to construct the statistical network are listed. Also,

red flags in this case need not be measured as zero or one. The input allows any numerical values between zero and one and thus can accommodate categorical and fuzzy inputs, which is an improvement over statistical techniques or neural networks. The red flags cannot be always measured as one or zero. For example, internal control can be weak, moderate, strong, or very strong. If the auditor wishes to measure red flags on a categorical or continuous scale then MQ Expert can accommodate such measures. The input values shown indicate the relative significance of the red flags according to MQ Expert. For example, weak internal control (Q 13) has a value of 0.25, which means MQ Expert considers weak internal control as the most important red flag. Assessing the risk of management fraud using a query facility is very easy. One can input the values of ten red flags (as measured in the audit engagement) then the probability for management fraud is displayed in the top right hand corner.

This query facility can be used to assess the risk of management fraud for any given en-

gagement. Virtually all CPA firms, as a part of the routine audit, collect data on red flags. The data on specific red flags can be fed into the network and the risk of management fraud can be assessed. The statistical network created by MQ Expert can be automatically encoded in a C language and can be embedded in any program or decision aid. This facility makes the statistical network portable and independent of the software (MQ Expert) that originated it.

Summary and Conclusions

Management fraud is a major problem for the auditing profession and has contributed to spectacular monetary judgments against public accounting firms. The research on red flags has shown a great promise in assessing the risk of management fraud. This paper uses a data mining tool to explore a fairly complex management fraud database. The database used in this paper is compiled by a Big Six accounting firm and has been analyzed by statistical techniques and neural networks. The results produced by MQ Expert compare very favorably with other techniques.

Figure 5: The Query Facility for the MQ Expert

Kept Inputs	Input Value
Q2	0.11
Q7	0.09
Q8	0.02
Q13	0.25
Q14	0.13
Q15	0.22
Q19	0.20
Q24	0.16
P1737	0.24
P39	0.08

Encode StatNet Model into ANSI 'C' Source Code

However, statistical techniques and neural networks are costly and time consuming to use. Required expertise is scarce in many small firms. The new generation of data mining tools employs powerful statistical techniques and yet is user friendly. Such tools can produce decision aids that are portable, scaleable (additional data can be included), and cost-effective.

During the process writing this paper and conducting our research, the authors devoted time and energy to learning about data mining and using MQ Expert. There are number of benefits in using data mining software. These are: 1) the data mining tools are relatively inexpensive, 2) the learning curve is less steep compared to statistical or neural network methods, 3) the training and usage costs are relatively low, 4) this approach can be used by small practitioners, and 5) It aids in structured decision making. On the other hand the costs associated with this approach are: 1) a basic understanding of data mining tools is required, 2) this approach can take more time than traditional checklists, and 3) the output of data mining tool is only as good as the input. The data mining tool does not supplant but supplements an auditor's judgment.

References

1. American Institute of Certified Public Accountants (AICPA) (1975). "Statement on Auditing Standards No. 6, Related Party Transactions." NY, NY, AICPA.
2. American Institute of Certified Public Accountants (AICPA) (1977). "Statement on Auditing Standards No. 16, The Independent Auditor's Responsibility for the Detection of Errors or Irregularities." NY, NY, AICPA.
3. American Institute of Certified Public Accountants (AICPA) (1977). "Statement on Auditing Standards No. 17, Illegal Acts." NY, NY, AICPA.
4. American Institute of Certified Public Accountants (AICPA) (1989). "Statement on Auditing Standards No. 53, The Auditor's Responsibility to Detect and Report Errors and Irregularities." NY, NY, AICPA.
5. American Institute of Certified Public Accountants (AICPA) (1989). "Statement on Auditing Standards No. 82, Consideration of Fraud in a Financial Statement Audit." NY, NY, AICPA.
6. Bell, T., S. Szykowny, and J. Willingham. (1993). "Development of a Decision Aid for Assessing the Likelihood of Fraudulent Financial Reporting." Working paper, KPMG Peat Marwick.
7. Bell, T., J. Carcello, and J. Willingham. (1996). "Development of a Decision Aid for Assessing the Likelihood of Fraudulent Financial Reporting." Working paper, KPMG Peat Marwick.
8. Bischoff, J. and T. Alexander. (1997). "Data Warehouse: Practical Advice from the Experts." Prentice Hall, Upper Saddle River, NJ.
9. Fanning, K., K. Cogger, and R. Srivastava. (1995). "Detection of Management Fraud: A Neural Network Approach." *Intelligent Systems in Accounting, Finance, and Management* 4: pp. 113-26.
10. Fayyad, U., G. Piatetsky-Shapiro, and P. Smyth. (1996). "The KDD Process for Extracting Useful Knowledge from Volumes of Data." *Communications of the ACM* 39:11, pp. 27-34.
11. Hansen J., J. McDonald, W. Messier, T. Bell. (1997). "A Generalized Qualitative-Response model and the Analysis of Management Fraud." *Management Science* 42:7, pp. 1022-32.
12. Loebbecke, J., and J. Willingham. (1988). "Review of SEC Accounting and Auditing Enforcement Releases." Working paper, KPMG Peat Marwick.
13. Loebbecke, J., M. Eining, and J. Willingham. (1989) "Auditor's Experience with Material Irregularities: Frequency, Nature, and Detectability." *Auditing: A Journal of Practice & Theory* 9, pp. 1-28.
14. Mena, J. (1996). "Automatic Data Mining." *PC AI*, Nov/Dec, pp. 16-20.
15. *ModelQuest Expert User's Manual Version 1.0*. (1996). AbTech Corporation.
16. Price Waterhouse. (1985). *Challenges and Opportunity for the Accounting Profession*.

Strengthening the Public's Confidence.
Price Waterhouse.