Knowledge-Based Systems 24 (2011) 314–321

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

A support vector machine-based model for detecting top management fraud

Ping-Feng Pai [a,b,*], Ming-Fu Hsu [b], Ming-Chieh Wang [b]

[a] Department of Information Management, National Chi Nan University, Taiwan, ROC
[b] Department of International Business Studies, National Chi Nan University, Taiwan, ROC

ARTICLE INFO

Article history:
Received 26 April 2010
Received in revised form 3 October 2010
Accepted 11 October 2010
Available online 13 October 2010

Keywords:
Support vector machine
Classification and regression tree
Feature selection
Top management fraud
Corporate governance

ABSTRACT

Detecting fraudulent financial statements (FFS) is critical in order to protect the global financial market. In recent years, FFS have begun to appear and continue to grow rapidly, which has shocked the confidence of investors and threatened the economics of entire countries. While auditors are the last line of defense to detect FFS, many auditors lack the experience and expertise to deal with the related risks. This study introduces a support vector machine-based fraud warning (SVMFW) model to reduce these risks. The model integrates sequential forward selection (SFS), support vector machine (SVM), and a classification and regression tree (CART). SFS is employed to overcome information overload problems, and the SVM technique is then used to assess the likelihood of FFS. To select the parameters of SVM models, particle swarm optimization (PSO) is applied. Finally, CART is employed to enable auditors to increase substantive testing during their audit procedures by adopting reliable, easy-to-grasp decision rules. The experiment results show that the SVMFW model can reduce unnecessary information, satisfactorily detect FFS, and provide directions for properly allocating audit resources in limited audits. The model is a promising alternative for detecting FFS caused by top management, and it can assist in both taxation and the banking system.

1. Introduction

Fraudulent financial statements (FFS) pose a critical issue in the defense of the global financial market. The word “fraud” denotes an intentional act designed to deceive or mislead another party [4]. It can be classified into two types: employee fraud and top management fraud. Generally, top management fraud involves the deliberation of accounting records, the falsification of transactions, or the misapplication of accounting principles [31]. Fraud by top managers has a devastating effect on a company’s shareholders and employees, and it can ruin a firm’s reputation and credibility [69].

Most FFS is caused by top managers who have authority to override the internal controls and deploy de facto power against audit committees. Some estimates suggest that fraud costs the US business more than $400 billion annually [67]. However, falsified financial statements are not just an American problem. A number of fraud scandals have broken out in Europe; they have shaken investor confidence in the global financial market and multinational trade.

Taiwan, which is located in East Asia, plays an important role in the global supply chain of electronic products; it is also an important capital market among global investors [34]. However, there are serious shortcomings in regard to corporate governance in East Asia, as La Porta et al. [33] have indicated. First, the ownership structure is less dispersed than in Europe and America, which increases the likelihood of fraud, as documented by Beasley [6]; he used logistic regression to analyze 75 fraudulent and 75 non-fraudulent firms. As his study showed, non-fraudulent firms have boards with a significantly higher proportion of outside members than is the case of fraudulent firms. A second shortcoming of many firms in East Asia is that the ultimate controllers (i.e., directors or managers) often maximize their power by means of a pyramid structure and shareholding. They also tend to pledge their shareholdings as loan collateral [34]. When top managers or directors of Taiwanese firms manipulate financial statements, they frequently do so to prevent a decrease in share price. Such managers understand the limitations of an audit and the insufficiency of standard auditing procedures in detecting fraud [52]. Some of the barriers that can prevent fraud from appearing on a firm’s financial statements are presented in Fig. 1.

Several models have been designed to detect fraud in financial statements. Eining et al. [15] found that auditors using expert system discriminated better among situations with varying levels of management fraud risk and made more consistent decisions

* Corresponding author at: Department of Information Management, National Chi Nan University, 470 University Road, Nantou County 54561, Taiwan, ROC. Tel.: +886 492910960x4672; fax: +886 492915205.
E-mail address: paipf@nenu.edu.tw (P.-F. Pai).

0950-7051/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.knosys.2010.10.003

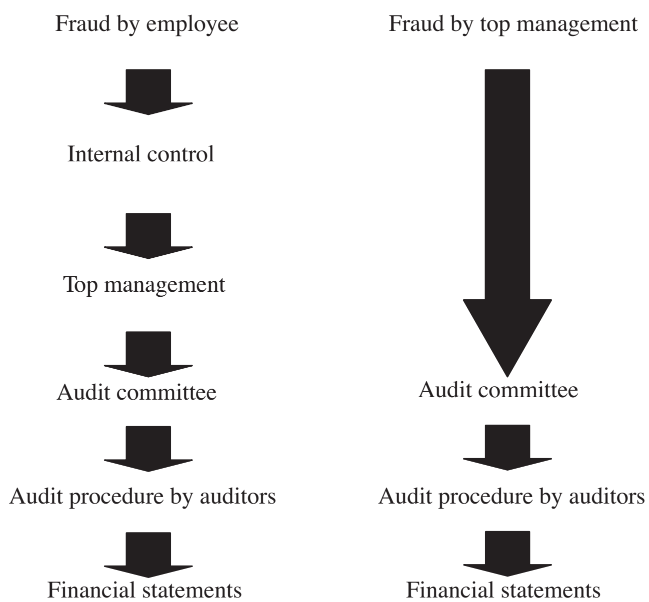

Fig. 1. Barriers between fraud by employee and fraud by top management.

regarding appropriate audit actions. Green and Choi [19] used an artificial neural networks technique to detect fraud, with limited satisfactory results. Fanning and Cogger [16] employed financial ratios and qualitative variables as input vectors to develop a fraud detection model. The proposed model outperformed both discriminant analysis and logistic regression. Spathis et al. [53] explored the effectiveness of an innovative classification methodology in detecting firms that issue falsified financial statements. This approach, which is based on a multi-criteria decision aid and the application of the utilities additives discriminantes classification method, obtained better results than did traditional statistical techniques. Finally, Kirkos et al. [28] compared the usefulness of the decision tree, neural networks, and Bayesian belief network in identifying financial statement fraud. Their results showed that the Bayesian belief network is superior to the other detective models.

As Morton [39] noted, auditors often apply sampling methods to audit; moreover, the probability of auditing is contingent on information obtained regarding the item assumed to be representative of all items in the sample. In this age of information explosion, auditors who perform limited audit procedures find it difficult to distinguish useful information from within the overabundant data. The proposed SVMFW model, which is supported by real example, can assist both internal and external auditors who must allocate limited audit resources to make appropriate decisions. Its application extends to taxation, the banking system, creditors, and regulatory agencies.

The current study presents a support vector machine-based fraud warming (SVMFW) model that integrates sequential forward selection (SFS), a support vector machine (SVM), and a classification and regression tree (CART). The advantage to this model pertains to its ability to minimize audit-related risks by classifying fraudulent financial statements as well as presenting the auditor with a set of comprehensible decision rules. The evidence provided by this study also offers policymakers with the ability to evaluate the policy implications of corporate governance mechanisms as well as to formulate future policies. The remainder of the study is organized as followed: Section 2 introduces the methodologies used in this study. Numerical examples and experimental results are presented in Section 3, and a summary of the findings appears in Section 4.

2. Methodologies

2.1. Sequential forward selection

Sequential forward selection (SFS), a typical heuristic searching scheme, identifies important features from unselected features and places them into a selected feature subset in each iteration [45]. SFS begins with an empty set of features and iteratively selects one feature at a time until no improvement in accuracy can be accomplished. For each iteration, all the features that have not yet been selected are considered for selection, and impacts of features on the evaluation score are recorded. The feature with the best impact is chosen and added into the set of candidate features. The search engine is the K-nearest neighbors (KNN) approach, and the leave-one-out test for estimating recognition rate is shown in Eq. (1).

$$\text{Recognition rate} = \frac{D - d}{D} \tag{1}$$

where D is the total number of training samples and d is the number of miss classification. Only remaining features are considered for the next iteration. The process is repeated until stop condition is triggered. The stop condition is the number of iterations when all features have been processed.

Assume that E is the objective function. Y_k is the feature set that have already been selected from original features set X at kth steps. Then SFS sequentially adds the feature x^+ by the following steps:

Step 1. Begin with the empty set: $Y_0 = \{\varphi\}$.
Step 2. Select the next best feature: $x^+ = \arg\max_{x \in X} \max_{x \in Y_k} [E(Y_k + x)]$.
Step 3. Update: $Y_{k+1} = Y_k + x^+$, $k = k + 1$.
Step 4. Go to step 2 until the stop condition is reached.

2.2. Support vector machine with PSO

The support vector machine (SVM) proposed by Vapnik [63] uses classification techniques based on statistical learning theory [63,64]. SVM produces a binary classifier, the so-called optimal separating hyperplanes, through an extremely non-linear mapping of the input vectors into a high-dimensional feature space. SVM constructs a linear model to estimate a decision function using non-linear class boundaries based on support vectors. If the data are linearly separated, SVM trains linear machines for an optimal hyperplane that separates the data without error and into the maximum distance between the hyperplane and the closest training points. The training points that are closest to the optimal separating hyperplane are called support vectors [29].

There are some advantages to using SVM [50]: (1) there are only two free parameters to be chosen, namely, the upper bound and the kernel parameter; (2) the solution of SVM is unique, optimal, and global since the training of an SVM is done by solving a linearly constrained quadratic problem; (3) SVM is based on the structural risk minimization principle, which means that this type of classifier minimizes the upper bound of the actual risk, whereas other classifiers minimize the empirical risk. Due to the above advantages, a number of studies have been conducted concerning its theory and application. These applications include: handwritten digit recognition [9,12], financial time-series forecasting [37,58,43,38], estimating manufacturing yields [56], face detection using images [41], default prediction [29], medical diagnosis [61], marketing [7], text categorization [25], bankruptcy prediction [22], debris-flow analysis [66], threatening e-mail [3], multi-fault classification problems [59], and business failure forecasting [35].

SVM is a learning machine originally for binary classification problems. Given a training set of instance-label pair (x_i, y_i), $i = 1, \ldots, m$ where $x_i \in R^n$ and $y_i \in \{\pm 1\}$, SVM finds an optimal hyperplane to separate the two classes, so that the margin is maximal. The margin is defined as the distance between the nearest point and the hyperplane. The process is equivalent to solving the following optimization problem:

$$\min \frac{1}{2}(w \cdot w) + C\sum_{i=1}^{m}\zeta_i \tag{2}$$

subject to:

$$\begin{aligned} & y_i((w \cdot x_i) + b) \geqslant 1 - \zeta_i, \quad i = 1, \ldots, m \\ & \zeta_i \geqslant 0, \quad i = 1, \ldots, m \end{aligned} \tag{3}$$

where ζ_i is a slack variable to give a soft classification boundary, w is a weight vector, b is a bias, and C is a constant corresponding to the value w^2.

The solution to dual problem is to maximize the quadratic from (4) under the constraint of (5).

$$\max_{\gamma} \sum_{i=1}^{m}\gamma_i - \frac{1}{2}\sum_{i,j=1}^{m} y_i y_j \gamma_i \gamma_j K(x_i, x_j) \tag{4}$$

subject to:

$$\begin{aligned} & 0 \leqslant \gamma_i \leqslant C, \quad i = 1, \ldots, m \\ & \sum_{i=1}^{z}\gamma_i y_i = 0 \end{aligned} \tag{5}$$

where γ_i are Largrange multipliers from the quadratic programming problem, and there is an γ_i for each vector in training set. Support vectors identify the surface of decision, and correspond to the subset of non-zero γ_i, the vectors can be views as the most critical training vectors. The kernel function $K(x_i, x_j)$ is employed to transform the original input space to a high dimensional space. The distinguishing function can be expressed as Eq. (6):

$$F(x) = sign\left[\left(\sum_{i=1}^{m}\gamma_i y_i K(x, x_i)\right) + b\right]. \tag{6}$$

The kernel function, $K(x,x_i)$, typically has the following alternatives:

$$K(x, x_i) = \exp\{-\|x - x_i\|^2/2\sigma^2\} \quad \text{(RBF kernel)} \tag{7}$$

$$K(x, x_i) = (1 + x \cdot x_i/c)^d \quad \text{(polynomial kernel of degree } d\text{)} \tag{8}$$

$$K(x, x_i) = \tanh(kx \cdot x_i + \theta) \quad \text{(MLP kernel)} \tag{9}$$

In this study, the RBF kernel was used. The selection of two parameters (σ and C) for a SVM model influences the classification accuracy significantly. In this study, the particle swarm optimization (PSO) [27] was used to determine the parameters values of SVM models.

The PSO algorithm is an emerging population-based meta-heuristic that simulates social behavior, such as birds flocking to a promising position, to achieve precise objectives in a multi-dimensional space. A particle is considered as a point in a H-dimensional space and its status is characterized according to its position x_{ih} and velocity v_{ih} [40,13]. The H-dimensional position for the particle i at iteration t can be represented as $x_i^t = \{x_{i1}^t, x_{i2}^t, \ldots, x_{iH}^t\}$ The velocity, which is also a H-dimensional vector, for particle i at iteration t can be described as $v_i^t = \{v_{i1}^t, v_{i2}^t, \ldots, v_{iH}^t\}$. Let $l_i^t = \{l_{i1}^t, l_{i2}^t, \ldots, l_{iH}^t\}$ represent the best solution that particle i has obtained until iteration t, and $l_g^t = \{l_{g1}^t, l_{g2}^t, \ldots, l_{gH}^t\}$ denote the best solution from l_i^t in the population at iteration t. To search for an optimal solution, each particle changes its velocity according to the cognition and social part as follows:

$$V_{ih}^t = V_{ih}^{t-1} + c_1 j_1 (L_{ih}^t - x_{ih}^t) + c_2 j_2 (L_{gh}^t - x_{ih}^t) \quad h = 1, 2, \ldots, H \tag{10}$$

where c_1 indicates the cognitive learning factor, c_2 indicates the social learning factor, and j_1 and j_2 are the random numbers uniformly distributed in $U(0, 1)$. Each particle then moves to a new potential solution based on the following equation:

$$X_{ih}^{t+1} = X_{ih}^t + V_{ih}^t, \quad d = 1, 2, \ldots, D \tag{11}$$

This investigation applies the PSO algorithm to determine SVM models. The procedure is depicted as follows. First, PSO is initialized with a random population of particles and velocities. Then, the fitness of each particle in the population is measured. In this study, the fitness of PSO is defined as the classification accuracy of SVM models. Each particle's velocity is calculated according to Eq. (10). For each particle, the procedure then moves to the next position according to Eq. (11). Finally, if the termination criterion is satisfied, the algorithm is stopped; otherwise it is returned to measure the fitness.

2.3. Classification and regression trees

The knowledge acquisition problem has been identified as a major bottleneck in the expert system development process [65]. Recent empirical research reveals that a classification and regression tree (CART) is significant in helping to extract certain types of knowledge within specific problem domains. CART, a statistical procedure introduced by Breiman et al. [8], is a robust, easy-to-use decision tree tool that automatically shifts through large, complex databases in searching for and isolating significant patterns and relationships [26]. CART adopted a recursive partitioning, a combination of exhaustive searches and intensive testing techniques to identify an informative tree structure in the data, also referred to as a decision tree. The decision tree is then used to generate reliable, easy-to-grasp predictive models in the form of "if-then" rules [26]. Auditors can use these rules to allocate limited audit resources.

2.4. The support vector machine based fraud warning model (SVMFW)

The issues of data preprocessing [66,35], parameter determination [59], and rule generation [66] influence the performance of SVM models a lot. Thus, this study developed a SVMFW model integrating unique strength of data preprocessing, parameter determination, and rule generation together into a SVM model to detect top management fraud. The sequential forward selection technique and particle swarm optimization were used for data preprocessing and parameter determination respectively. Furthermore, training data with the smallest classification error obtained from the cross-validation (CV) procedure of SVM models were employed for CART to generate rules.

Fig. 2 illustrates the flowchart generated by the SVMFW model for detecting top management fraud. First, the FFS data is processed by the SFS method to select critical attributes, which are then scaled. Second, the scaled data are divided into two set: training and testing data sets. The training data set is used only to form a rule-base derived from the SVMFW model, while the testing data set is used to examine the classification performance of a well-trained model. The PSO method is used to determine sequentially the SVM parameters; thus, each succeeding iteration produces a smaller classification error. The parameter search procedure is undertaken until the stop criterion of PSO is reached. The two parameters resulting in the smallest training error are then employed to generate a finalized SVM model for classification. Rather than the expression of complicated mathematical functions, it is easier for auditors to realize the relation as well as the strength between features and FFS intuitionally from a set of decision rules. The investigation applied the CV training data set with the most accurate testing result to derive decision rules. Finally, "if-then"

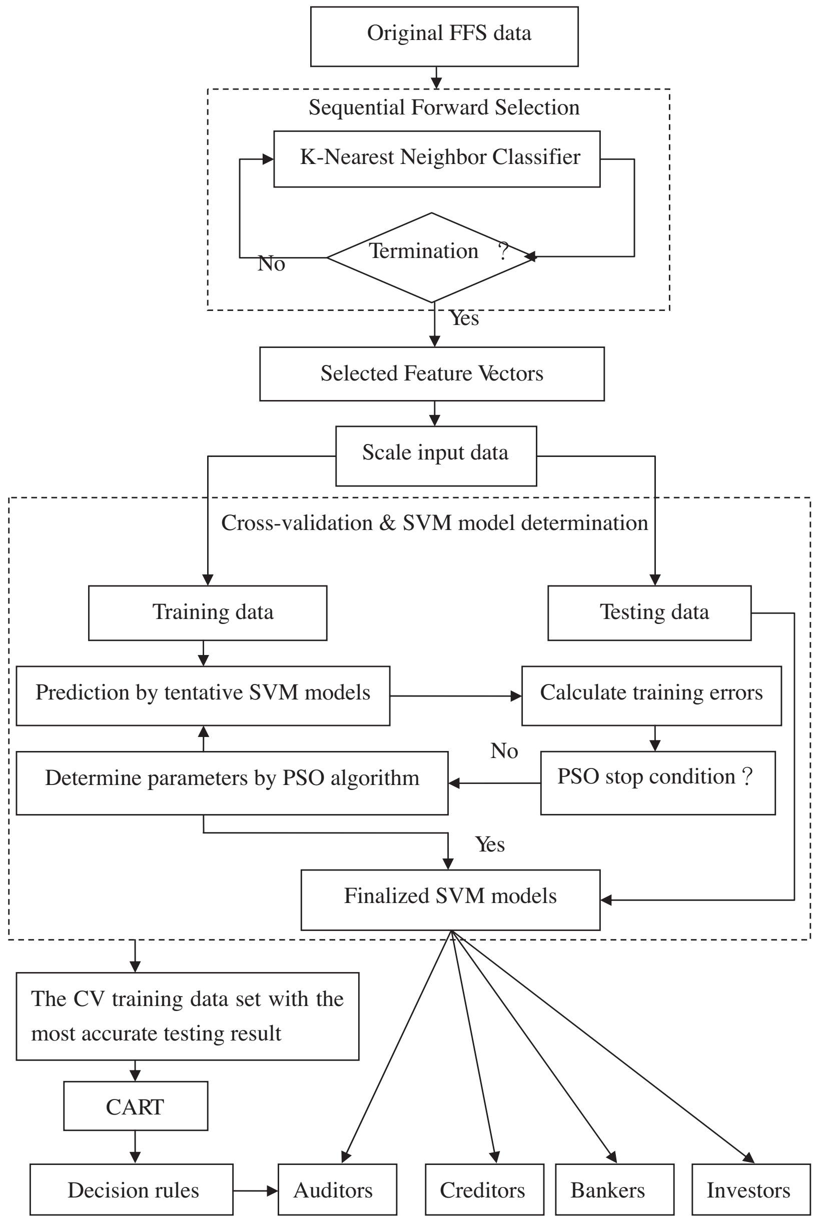

Fig. 2. Flowchart of the SVMFW model for fraud detection.

rules are derived from CART which can help auditors to allocate limited audit resources.

3. A numerical example and simulation results

The data in this study contain 75 listed firms in Taiwan's stock market. These firms include 25 FFS firms and 50 non-FFS firms. Published indication or proof of involvement in issuing FFS was found for the 25 FFS firms. The classification of a financial statement as fraudulent is based on the Security and Futures Investor Protection Center of Taiwan (SFI) [49] and the Financial Supervisory Commission of Taiwan (FSC) [18] during the 1999–2005 reporting period. All the financial and corporate governance features used in the sample are extracted from formal financial statements, such as balance sheets and income statements. The financial and corporate governance features are derived from the database of Taiwan Economic Journal (TEJ) and the website of Taiwan Securities and Futures Institute (SFI). Table 1 lists 18 features of financial statements used in this study. The 18 features consist of 16 financial features and two corporate governance features, which are based on the work of Spathis [52], Spathis et al. [53], Feroz et al. [17], Chen [10], Stice [55], Fanning and Cogger [16], Stalebrink and Sacco [54], Kotsiantis et al. [32], Kinney and McDaniel [30], Rezaee [46]. Furthermore, feature selection is essential in designing a classifier. Three other feature selection methods, namely factor analysis (FA) [51], stepwise regression (SWR) [50], and *t*-test [23], were employed to deal with the same financial data. Table 2 shows that selected features by four feature selection methods. The iteration number and recognition rate by SFS were shown in Fig. 3. Table 3 illustrates the performance of the four feature selection methods fed into SVM models. Table 4 summarizes Type I and Type II errors for four feature selection methods with SVM models. To sum up, the SFS approach outperformed the other three feature selection methods when the SVM technique is used as a classifier in this study. Additionally, the grid search (GS), genetic algorithms (GA), and simulated annealing algorithms (SA) were used to deal with the same data for comparing with results

Table 1
Illustration of features used in the study.

Category	Independent features	FFS		Non-FFS	
		Mean	Standard deviation	Mean	Standard deviation
Profitability features	X7:NP/TA (Net Profit to Total Assets)	−0.211	0.333	0.465	0.956
	X8:EBIT (Earnings before Interest and Tax)	−1466997	1996178	401219	930094
	X13:NP/S (Net Profit to Sales)	1.123	0.692	2.322	2.594
Leverage features	X1:LOTD (Logarithm of Total Debt)	6.594	0.412	6.336	0.448
	X4:TD/TE (Total Debt to Total Equity)	10.808	48.661	0.856	0.875
	X16:TD/TA (Total Debt to Total Assets)	1.774	0.687	3.1	2.805
	X11:LTD/TA (Long-term Debt to Total Assets)	0.185	0.151	0.295	0.894
Liquidity features	X6:QA/CL (Quick Assets to Current Liability)	0.492	0.409	0.837	0.773
	X10:WC/TA (Working Capital to Total Assets)	0.145	0.144	0.126	0.17
	X14:CA/CL (Current Assets to Current Liability)	0.116	0.096	0.129	0.133
Efficiency features	X2:AR/S (Account Receivable to Sales)	0.242	0.149	0.21	1.367
	X3:NI/FA (Net Income to Fixed Assets)	−1.28	2.366	0.223	0.529
	X5:I/TA (Inventory to Total Assets)	0.116	0.096	0.129	0.133
	X9:AR/S (Account Receivable to Sales)	−0.049	0.329	0.154	0.192
	X12:I/S (Inventory to Sales)	−0.428	0.513	0.117	0.264
	X15:S/TA (Sales to Total Assets)	0.763	0.708	0.036	0.701
Corporate governance features	X17:DO (Director Ownership)	14.084	8.586	35.005	20.914
	X18:PSD (Pledged Shares of Directors)	42.642	33.396	5.23	12.246

Table 2
The selected features by four feature selection methods.

Methods	Selected features
SFS	X3, X7, X8, X12, X16, X18
SWR	X8, X17, X18
FA	X2, X7, X9, X12, X13
t-test	X3, X7, X8, X10, X13, X17, X18

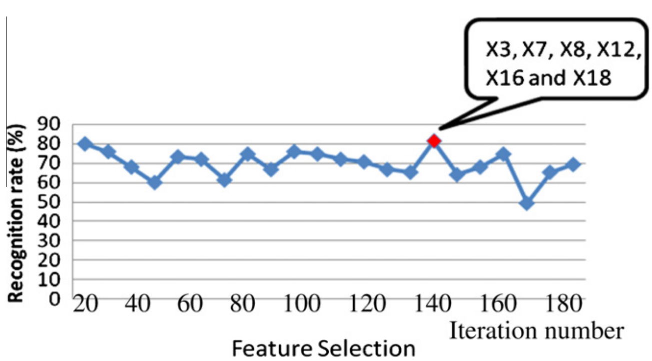

Fig. 3. Selected features by sequential forward selection.

Table 3
Test accuracies of SVM models with four feature selection methods.

	Accuracies obtained by four feature selection methods (%)			
	SFS	SWR	FA	*t*-test
CV-1	93.33	73.33	80	80
CV-2	80	73.33	86.67	80
CV-3	93.33	80	80	86.67
CV-4	93.33	73.33	80	80
CV-5	93.33	80	73.33	86.67
Average accuracy	92	76	80	82.67

Table 4
Type I and Type II errors of SVM models with four feature selection methods.

Feature selection methods	Error rates (%)	
	Type I	Type II
SFS	5	16
SWR	22	28
FA	18	24
t-test	16	20

obtained by PSO in selecting SVM parameters. The initial values of σ for fivefold cross validation of the PSO algorithm are 0.65, 0.23, 0.19, 0.84, and 0.61, respectively. These values are random numbers between 0 and 1. Table 5 shows classification performances of four approaches in determining SVM parameters. It can be concluded that the PSO algorithm was superior to the other three approaches in terms of average testing accuracy. Sequentially, five other classifiers, namely linear discriminant analysis (LDA) [1,2], logistic regression (LR) [14,44,42,24], C4.5 classification tree (C4.5) [5,28,57], multi-layer perception (MLP) neural networks [70,60], RBF networks [47,21] were tested in this study. Table 6 indicates that the SVM model performed the best among six classifiers in terms of testing accuracy for both data sets.

In this study, the selected features by SFS consist of financial items (X3 for ''Net Income to Fixed Assets'', X7 for ''Net Profit to Total Assets'', X8 for ''Earnings before Interest and Tax'', X12 for ''Inventory to Sales'', and X16 for ''Total Debt to Total Assets'') and corporate governance items (X18 for ''Pledged Shares of Directors''). Debt structure is of interest because it shifts the risk from equity owners and managers to debt owners [52]. Chow and Rice [11] reported that the potential for wealth transfer from debt holders to managers increases as leverage increases. Management might manipulate financial statements to meet certain debt covenants, which means that higher levels of debt may enhance the probability of FFS. The feature of ''Total Debt to Total Assets'' can be used to evaluate the debt structure.

While debt structure might not be ignored, certain other items, including sales, account receivables, allowance for doubtful accounts, and inventory are more likely to be manipulated by management, since they are subjective in nature and therefore more difficult to audit [68,20,48,36]. Specifically, accounts receivables and inventory depend on subjective judgment in estimating uncollected accounts and obsolete inventory [53]. Management may

Table 5
Classification performances of four methods in determining SVM parameters.

Methods	Cross-validation	Values of SVM parameters		Testing accuracies (%)	Average testing accuracies (%)
		C	σ		
Grid	CV-1	8	0.0781	86.67	80
	CV-2	32	0.0781	80	
	CV-3	8	0.125	73.33	
	CV-4	32768	0.0195	80	
	CV-5	512	0.0126	80	
GA	CV-1	5.1222	1.4761	80	84
	CV-2	9.3353	5.3568	86.67	
	CV-3	8.9834	4.7898	80	
	CV-4	4.0860	5.7674	86.67	
	CV-5	7.8006	9.3548	86.67	
SA	CV-1	4.44234	7.04061	80	86.67
	CV-2	5.76101	1.77803	86.67	
	CV-3	2.255	7.25708	86.67	
	CV-4	8.8286	4.31921	93.33	
	CV-5	0.610982	3.22787	86.67	
PSO	CV-1	3.15	0.0238	93.33	92
	CV-2	1.2217	0.0075	80	
	CV-3	5.286	0.0233	93.33	
	CV-4	7.6676	0.018	93.33	
	CV-5	3.615	0.0449	93.33	

Table 7
Type I and Type II errors with two data sets.

Models	Error rates (%)	
	Type I	Type II
SFSLR with data set one	10	22
SFSLR with data set two	12	56
SFSLDA with data set one	10	20
SFSLDA with data set two	12	44
SFSC4.5 with data set one	18	16
SFSC4.5 with data set two	18	32
SFSMLR with data set one	14	20
SFSMLR with data set two	12	32
SFSRBFNN with data set one	18	16
SFSRBFNN with data set two	18	32
SVMFW with data set one	5	16
SVMFW with data set two	10	36

manipulate financial statements to utilize the characteristics of these subjective items, thereby affecting the following features: ‘‘Net Income to Fixed Assets’’, ‘‘Earnings before Interest and Tax’’, and ‘‘Inventory to Sales.’’

Previous research has examined other financial red flags, including ‘‘Net Profit to Total Assets’’, for their ability to predict FFS [36]. Lin [34] declared that the ownership structure in East Asia is less dispersed than it is in Europe and America, and that controllers often maximize their control by means of a pyramid structure and cross shareholdings. The author also reported that it is common for a director or top manager of Taiwanese firms to pledge their shareholdings as loan collateral. If the share price falls, directors or top managers may abuse their authority to manipulate financial statements. It follows that pledged shares can be an indication of FFS.

This investigation further examined the corporate governance’s impact on detecting FFS. Six models are used to deal with the financial statement data collected from the Taiwan Economic Journal [62]. The data are divided into two sets: the data with the feature, ‘‘X18 for Pledged Shares of Directors’’ (data set one), and the data without this feature (data set two). To obviate the over-fitting problem, a 5-fold cross-validation procedure is conducted whereby training data are divided into subsets of equal size. Each subset is tested sequentially by adopting the classifier trained on the remaining four subsets. An instance of the whole of the training set is predicted once; the cross-validation accuracy is the percentage of data that are correctly classified. Table 6 shows the experimental results of the six models embedded SFS with and without the feature, ‘‘X18 for Pledged Shares of Directors.’’ The results indicate that this feature is useful in improving the testing accuracy of the detective model, and that the SVMFW model incorporating this feature outperforms the other models with different data sets.

Another indicator used in this study, for assessing the performance of the SVMFW model, is the misclassification rate. A Type I error means that a non-FFS firm has been misclassified as a FFS one. A Type II error means that a FFS firm has been misclassified as a non-FFS one. Type I and Type II errors have different cost. Type I errors may lead to additional investigation. Type II errors may result in unacceptable audits that damage reputation and lead to huge economic losses. Table 7 depicts both types of errors obtained from the six models with the two data sets. Given this finding, it is further concluded that the SVMFW model with data set one can derive the lowest Type I and Type II error rates.

To assist auditors in allocating audit resources, the CART approach was used to derive ‘‘if-then’’ rules from the CV training data set with the best testing result in terms of testing accuracy. Table 8 lists rules derived from CART method, the accuracy, and two types of errors of CART. Fig. 4 shows the tree structure of these rules. It can be observed that the feature of ‘‘Pledged Share of Directors’’

Table 6
Testing accuracies of six classifiers with SFS for two data sets.

Classifiers	Data sets	Testing accuracies (%)					
		CV-1	CV-2	CV-3	CV-4	CV-5	Average testing accuracies
LR	Data set one	86.67	66.67	93.33	73.33	73.33	78.67
	Data set two	66.67	80	66.67	80	73.33	73.33
LDA	Data set one	93.33	80	86.67	73.33	73.33	81.33
	Data set two	73.33	73.33	86.67	80	66.67	76
C4.5	Data set one	73.33	80	86.67	93.33	86.67	84
	Data set two	66.67	73.33	86.67	73.33	80	76
MLP	Data set one	73.33	86.67	80	86.67	86.67	82.67
	Data set two	73.33	80	80	80	86.67	80
RBF	Data set one	86.67	80	80	86.67	80	82.67
	Data set two	80	73.33	80	80	73.33	77.33
SVM	Data set one	93.33	86.67	93.33	93.33	93.33	92
	Data set two	86.67	73.33	80	93.33	73.33	81.33

Table 8
Decision rules derived from CART.

(1) If "X18 for Pledged Shares of Directors " ⩾ 44.405, then "FFS"
(2) If "X18 for Pledged Shares of Directors" < 44.405 and "X7 for Net Profit to Total Assets" < −0.3229, then "FFS"
(3) If "X18 for Pledged Shares of Directors" < 44.405, "X7 for Net Profit to Total Assets" ⩾ −0.3229 and "X3 for Net Income to Fixed Assets" ⩾ 0.0497, then "non-FFS"
(4) If "X18 for Pledged Shares of Directors" < 44.405, "X7 for Net Profit to Total Assets" ⩾ −0.3229, "X3 for Net Income to Fixed Assets" < 0.0497 and "X8 for Earnings before Interest and Tax" < −42,220, then "non-FFS"
(5) If "X18 for Pledged Shares of Directors" < 44.405, "X7 for Net Profit to Total Assets" ⩾ −0.3229, "X3 for Net Income to Fixed Assets" < 0.0497, "X8 for Earnings before Interest and Tax" ⩾ −42,220, and "X16 for Total Debt to Total Assets" ⩾ 1.48 then, "FFS"
(6) If "X18 for Pledged Shares of Directors" < 44.405, "X7 for Net Profit to Total Assets" ⩾ −0.3229, "X3 for Net Income to Fixed Assets" < 0.0497, "X8 for Earnings before Interest and Tax" ⩾ −42220, and "X16 for Total Debt to Total Assets" < 1.48 then, "non-FFS"
Accuracy: 86.67%; Type I error: 10%; Type II error: 20%

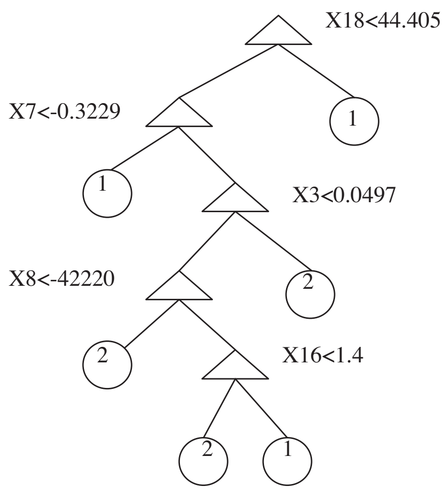

Fig. 4. The tree structure of rules obtained by the CART.

is the first split point. This means that shares pledged by directors are critical in prediction FFS by top management, a finding that corroborates the study of Lin [34]. Clearly, auditors need to give particular consideration to this red flag in their audit procedures. The selected feature "X12 for Inventory to Sales" was pruned in the CART decision tree. Inventory to sales" was pruned in the CART decision tree. That implies the feature "X12 for Inventory to Sales" is less important compared with the other selected features. The auditors thus can allocate limited audit resources according to the importance of features indicated in the tree structure of rules obtained by the CART. The importance of feature is decreasing from the top of the tree (X18) to the bottom of the tree (X16).

4. Conclusion

The detection of FFS is an important and challenging issue that has been rigorously investigated in recent years, as the number of management fraud cases has increased. Many different kinds of technology have been introduced to deal with the audit-related risk, and the attempt to improve on these models continues. The current investigation presents a SVMFW model for analyzing financial statement data. The empirical results indicate that the proposed model is an effective and efficient alternative in detecting top management fraud. Rather than presenting complicated mathematical functions, the SVMFW model provides a set of comprehensible decision rules for auditors, who must allocate limited audit resource. In the future, other input features such as audit committees, boards with a significant proportion of outside members, CPA tenure and CEO duality might included in the SVMFW model to further enhance its effectiveness in analyzing financial statement data.

Acknowledgement

The authors would like to thank the National Science Council of the Republic of China, Taiwan for financially supporting this research under Contract No. 96-2628-E-260-001-MY3 & 99-2221-E-260-006.

References

[1] E.L. Altman, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, J. Finance 23 (3) (1968) 589–609.
[2] E.L. Altman, I. Edward, R. Haldeman, P. Narayanan, A new model to identify bankruptcy risk of corporations, J. Bank. Finance 1 (1977) 29–54.
[3] S. Appavu, R. Rajaram, M. Muthupandian, G. Athiappan, K.S. Kashmeera, Data mining based intelligent analysis of threatening e-mail, Knowl.-based Syst. 22 (5) (2009) 392–393.
[4] A. Arens, J. Loebbecke, Auditing: An Integrated Approach, Englewood Cliffs, NJ, Prentice Hall, 1996.
[5] B. Baesens, T.V. Gesyel, S. Viaene, M. Steoanova, J. Suykens, J. Vanthienen, Benchmarking state-of-the-art classification algorithms for credit scoring, J. Oper. Res. Soc. 54 (6) (2003) 627–635.
[6] M.S. Beasley, An empirical analysis of the relation between the board of director composition and financial statement fraud, Account. Rev. 71 (4) (1996) 443–466.
[7] S. Ben-David, M. Lindenbaum, Learning distributions by their density levels: a paradigm for learning without a teacher, J. Comput. Syst. Sci. 55 (1) (1997) 171–182.
[8] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classification and Regression Trees, Wadsworth International Group, Belmont, California, 1984.
[9] C.J.C. Burges, B. Schokopf, Improving the accuracy and speed of support vector machines, Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA, 1997. pp. 475–481.
[10] G. Chen, Positive research on the financial statement fraud factors of listed companies in China, J. Mod. Account. Audit. 2 (6) (2006) 25–34.
[11] C. Chow, S. Rice, Qualified audit opinions and share prices: an investigation, Auditing: J. Pract. Theory 1 (1982) 35–53.
[12] C. Cortes, V.N. Vapnik, Support vector networks, Mach. Learn. (20) (1995) 273–297.
[13] C.A.C. Coello, G.T. Pulido, M.S. Lechuga, Handling multiple objectives with particle swarm optimization, IEEE Trans. Evolut. Comput. (8) (2004) 256–279.
[14] A.I. Dimitras, S.H. Zanakis, C. Zopounidis, A survey of business failure with an emphasis on prediction methods and industrial applications, Eur. J. Oper. Res. 90 (3) (1996) 487–513.
[15] M.M. Eining, D.R. Jones, J.K. Loebbecke, Reliance on decision aids: an examination of auditors' assessment of management fraud, Auditing: J. Pract. Theory 16 (2) (1997) 1–19.
[16] K. Fanning, K. Cogger, Neural network detection of management fraud using published financial data, Int. J. Intell. Syst. Account. 7 (1) (1998) 21–24.
[17] E. Feroz, K. Park, V. Pastena, The financial and market effects of the SECs accounting and auditing enforcement releases, J. Account Res. 29 (1991) 107–142.
[18] Financial Supervisory Commission, Executive Yuan, ROC. <http://www.fscey.gov.tw/Layout/main_ch/index.aspx?frame=1>.
[19] B.P. Green, J.H. Choi, Assessing the risk of management fraud through neural network technology, Audit: J. Pract. Theory 16 (1) (1997) 14–28.
[20] B. Green, Identifying Management Irregularities Through Preliminary Analytical Procedures, Unpublished Doctoral Dissertation, Kent State University, Kent, OH, 1991.
[21] P.A. Gutiérrez, M.J. Segovia-Vargas, S. Salcedo-Sanz, C. Hervás-Martínez, A. Sanchis, J.A. Portilla-Figueras, F. Fernández-Navarro, Hybridizing logistic regression with product unit and RBF networks for accurate detection and prediction of banking crises, Int. J. Manage. Sci. Omega 38 (5) (2010) 333–344.
[22] W. Haärdle, R. Moro, D. Schaäfer, Predicting Corporate Bankruptcy with Support Vector Machines, Humboldt University and the German Institute for Economic Research, 2003.

[23] Z. Huang, H. Chen, C.J. Hsu, W.H. Chen, S. Wu, Credit rating analysis with support vector machines and neural networks: a market comparative study, Decision Support Syst. 37 (4) (2004) 543–558.
[24] Z.S. Hua, B. Zhang, A hybrid support vector machines and logistic regression approach for forecasting intermittent demand of spare parts, Appl. Math. Comput. 181 (2) (2006) 1035–1048.
[25] T. Joachims, Learning to Classify Text Using Support Vector Machines, Springer: Kluwer Academic Publishers, 2002.
[26] D. Karthik-Chandra, V. Ravi, I. Bose, Failure prediction of dotcom companies using hybrid intelligent techniques, Expert Syst. Appl. 36 (3) (2009) 4830–4837.
[27] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proceedings of IEEE Conference on Neural Network, vol. 4, 1995, pp. 1942–1948.
[28] E. Kirkos, C. Spathis, Y. Manolopoulos, Data mining techniques for the detection of fraudulent financial statements, Expert Syst. Appl. 32 (4) (2007) 995–1003.
[29] H.S. Kim, S.Y. Sohn, Support vector machines for default prediction of SMEs based on technology credit, Eur. J. Oper. Res. 201 (3) (2010) 838–846.
[30] W. Kinney, L. McDaniel, Characteristics of firms correcting previously reported quarterly earnings, J. Account. Econ. 11 (1) (1989) 71–93.
[31] C.A. Knapp, M.C. Knapp, The effects of experience and explicit fraud risk assessment in detecting fraud with analytical procedures, Account. Org. Soc. 26 (1) (2001) 25–37.
[32] S. Kotsiantis, E. Koumanakos, D. Tzelepis, V. Tampakas, Forecasting fraudulent financial statements using data mining, Int. J. Comput. Intell. 3 (2) (2006) 104–110.
[33] R. La Porta, F. Lopez-de-Silances, A. Shleifer, Corporate ownership around the world, J. Finance 54 (2) (1999) 471–517.
[34] T.H. Lin, A cross model study of corporate financial distress prediction in Taiwan: multiple discriminant analysis, logit, probit and neural networks models, Neurocomputing 72 (16) (2009) 3507–3516.
[35] F. Lin, C.C. Yeh, M.Y. Lee, The use of hybrid manifold learning and support vector machines in the prediction of business failure, Knowl.-based Syst. (2010) 95–101.
[36] L. Loebbecke, M. Eining, J. Willingham, Auditor's experience with material irregularities: frequency, nature and detectability, Auditing: J. Pract. Theory 9 (1) (1989) 1–28.
[37] C.J. Lu, T.S. Lee, C.C. Chiu, Financial time series forecasting using independent component analysis and support vector regression, Decis. Support Syst. 47 (2) (2009) 115–125.
[38] S. Mukherjee, E. Osuna, F. Girosi, Nonlinear prediction of chaotic time series using support vector machines, in: Proceedings of the IEEE Workshop on Neural Networks for Signal Processing, 1997, pp. 511–520.
[39] S. Morton, Strategic auditing for fraud, Account. Rev. 68 (4) (1993) 825–839.
[40] A. Noorul Haq, K. Karthikeyan, K. Sivakumar, R. Saravanan, Particle swarm optimization algorithm for optimal machining allocation of clutch assembly, Int. J. Adv. Manuf. Technol. (27) (2006) 865–869.
[41] E. Osuna, R. Freund, F. Girosi, Training support vector machines: an application to face detection, in: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 1997, pp. 130–136.
[42] J. Ohlson, Financial ratios and the probabilistic prediction of bankruptcy, J. Account. Res. 18 (1) (1980) 109–131.
[43] P.F. Pai, C.S. Lin, A hybrid ARIMA and support vector machines model in stock price forecasting, Int. J. Manage. Sci. Omega 33 (6) (2005) 497–505.
[44] C. Pantalone, M.B. Platt, Predicting commercial bank failure since deregulation, New England Econ. Rev. (1987) 37–47.
[45] P. Pudil, J. Novovicova, J. Kittler, Floating search methods in feature selection, Pattern Recognit. 15 (11) (1994) 1119–1125.
[46] Z. Rezaee, Corporate governance role in financial reporting, Res. Account. Regul. 17 (2004) 107–149.
[47] H. Sarimveis, P. Doganis, A. Alexandridis, A classification technique based on radial basis function neural networks, Adv. Eng. Software 37 (4) (2006) 218–221.
[48] H. Schilit, Financial Shenanigans: How to Detect Accounting Gimmicks and Fraud in Financial Reports, McGraw-Hill, New York, 1993.
[49] Securities and Futures Investors Protection Center of Taiwan. <http://www.sfipc.org.tw/english/main.asp>.
[50] K.S. Shin, T.S. Lee, H.J. Kim, An application of support vector machines in bankruptcy prediction model, Expert Syst. Appl. 28 (1) (2005) 127–135.
[51] K.S. Shin, Y.J. Lee, A genetic algorithm application in bankruptcy prediction modeling, Expert Syst. Appl. 23 (3) (2002) 321–328.
[52] C. Spathis, Detecting false financial statements using published data: some evidence from Greece, Manag. Audit. J. 17 (4) (2002) 179–191.
[53] C. Spathis, M. Doumpos, C. Zopounidis, Detecting falsified financial statements: a comparative study using multi-criteria analysis and multivariate statistical techniques, Eur. Account. Rev. 11 (3) (2002) 509–535.
[54] O.J. Stalebrink, J.F. Sacco, Rationalization of financial statement fraud in government: an Austrian perspective, Crit. Perspect. Account. 18 (4) (2007) 489–507.
[55] J. Stice, Using financial and market information to identify pre-engagement market factors associated with lawsuits against auditors, Account. Rev. 66 (3) (1991) 516–533.
[56] D. Stoneking, Improving the manufacturability of electronic designs, IEEE Spectrum 36 (6) (1999) 70–76.
[57] J. Sun, H. Li, Data mining method for listed companies' financial distress prediction, Knowl.-based Syst. 21 (2008) 1–5.
[58] F.E.H. Tay, L. Cao, Application of support vector machines in financial time series forecasting, Int. J. Manage. Sci. Omega 29 (4) (2001) 309–317.
[59] X. Tang, L. Zhuang, J. Cai, C. Li, Multi-fault classification based on support vector machine trained by chaos particle swarm optimization, Knowl.-based Syst. 23 (5) (2010) 486–490.
[60] K.Y. Tam, M.Y. Kiang, Managerial applications of neural networks: the case of bank failure predictions, Manage. Sci. 38 (7) (1992) 926–947.
[61] L. Tarassenko, P. Hayton, N. Cerneaz, M. Brady, Novelty detection for the identification of masses in mammograms, in: Proceedings Fourth IEE International Conference on Artificial Neural Networks, Cambridge, 1995, pp. 442–447.
[62] Taiwan Economic Journal Co., Ltd. <http://www.tej.com.tw/twsite/> .
[63] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.
[64] V.N. Vapnik, Statistical Learning Theory, Springer, New York, 1998.
[65] W.P. Wagner, J. Otto, Q.B. Chung, Knowledge acquisition for expert systems in accounting and financial problem domains, Knowl.-based Syst. 15 (8) (2002) 439–447.
[66] S. Wan, T.C. Lei, A knowledge-based decision support system to analyze the debris-flow problems at Chen-Yu-Lan River, Taiwan, Knowl.-Based Syst. 22 (8) (2009) 580–588.
[67] T.J. Wells, Occupational Fraud and Abuse, Obsidian Publishing, Austin, 1997.
[68] A. Wright, R. Ashton, Identifying audit adjustments with attention-direction procedures, Account. Rev. 64 (4) (1989) 79–83.
[69] S.A. Zahra, R.L. Priem, A.A. Rasheed, Understanding the causes and effects of top management fraud, Organ. Dynam. 36 (2) (2007) 122–139.
[70] G. Zhang, M.Y. Hu, B.E. Patuwo, D.C. Indro, Artificial neural networks in bankruptcy prediction: general framework and cross-validation analysis, Eur. J. Oper. Res. 116 (1999) 16–32.