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

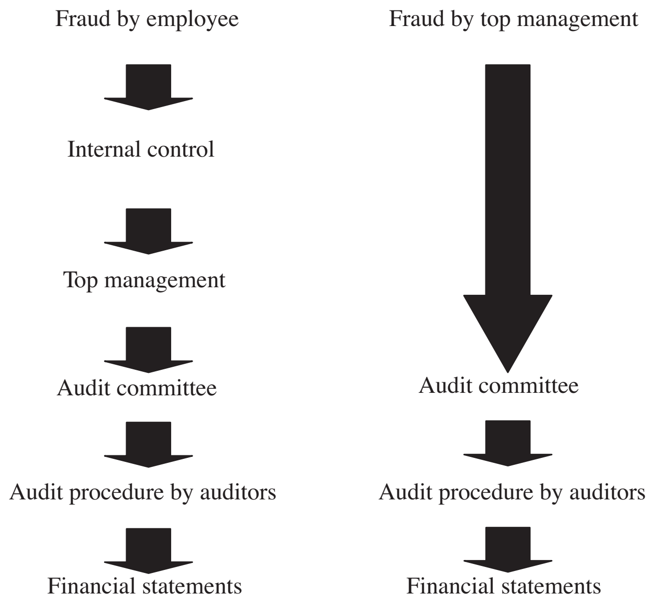

Fig. 1. Barriers between fraud by employee and fraud by top management.

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