陳育仁 (2022)，「企業資產掏空偵測」，*資訊管理學報*，第二十九卷，
第一期，頁 75-101。

企業資產掏空偵測

陳育仁*
國立高雄科技大學 會計資訊系

摘要

對於投資大眾來說，最關注其投資的企業是否發生舞弊之現象；而企業最常見的舞弊手法包括財報不實、資產掏空與內線交易等，其中又以資產掏空所造成的後果最為嚴重，可能導致整個企業停擺，使得投資者血本無歸。因此，如何有效的偵測企業掏空與否，已成為舞弊審計重要的課題之一。

本研究主要考量企業財務結構、償債能力、經營能力、獲利能力、現金流量與成長力等財務性指標以及公司治理方面之股權結構、董監事組成、關係人交易與管理型態等非財務性指標，並經由主成分分析與逐步迴歸進行資產掏空偵測指標之建立，再整合支援向量機與改良式基因演算法進行企業資產掏空之偵測，以提供投資大眾選擇目標企業投資時之決策參考，進而降低投資者之投資風險。針對上述目的，本研究主要研究項目包括: (i)資產掏空偵測特徵指標之建立，(ii)資產掏空偵測方法之發展以及(iii)資產掏空方法之驗證與評估。

關鍵詞：資產掏空、舞弊偵測、主成分分析、逐步迴歸、支援向量機、改良式基因演算法

---

* 本文通訊作者。電子郵件信箱：yjchen@nkust.edu.tw
2021/08/17 投稿；2021/11/29 修訂；2021/12/15 接受

Chen, Y.J. (2022). Fraud detection for corporate asset misappropriation, *Journal of Information Management*, 29(1), 75-101.

Fraud Detection for Corporate Asset Misappropriation

Yuh-Jen Chen*
Department of Accounting and Information Systems, National Kaohsiung University of Science and Technology

Abstract

Management fraud is a major concern among investors. Misstated financial statements, asset misappropriation, and insider trading are common forms of fraud by enterprises. Among these, asset misappropriation presents the highest risk for substantial losses as it may result in corporate shutdown and the loss of the life savings of investors. Therefore, identifying methods to effectively detect bad management at the earliest and prevent corporate asset misappropriation is critical in the study of fraud audits.

This study considers several financial structure indicators - solvency, operating capacity, profitability, cash flow, and growth ability. It also examines several non-financial indicators - shareholding structure, board composition, related party transactions, and management style in corporate governance. The feature indicators for asset misappropriation detection are first established through principal component analysis and stepwise regression. Support vector machine (SVM) and queen genetic algorithm (QGA) are then combined to effectively detect corporate asset misappropriation, providing a reference to investors and creditors for investment decision-making and thereby reducing their investment risks. This objective is achieved by (i) establishing feature indicators for asset misappropriation detection, (ii) developing an asset misappropriation detection method, and (iii) demonstrating and evaluating the proposed method.

Keywords: Asset misappropriation, fraud detection, principal component analysis, stepwise regression, support vector machine, queen genetic algorithm

---

* Corresponding author. Email: yjchen@nkust.edu.tw

2021/08/17 received; 2021/11/29 revised; 2021/12/15 accepted

1. INTRODUCTION

In recent years, fraudulent events such as misstated financial statements, asset misappropriation, and insider trading have become frequent among enterprises. Among these, asset misappropriation presents the highest risk for the most substantial losses for investors and can even result in corporate shutdown (Akman et al. 2020; Zager et al. 2016; Chen et al. 2009; Goel 2013; Lehman 2015; Olmo et al. 2011; Park & Lee 2010; Zhou & Kapoor 2011). Therefore, effectively detecting corporate asset misappropriation is a key while examining fraud audits.

Recently, corporate asset misappropriation has been studied extensively. For instance, Cecchini et al. (2010) developed a methodology for automatically analyzing text to aid in discriminating firms that encounter catastrophic financial events. The dictionaries created from Management Discussion and Analysis Sections (MD&A) of 10-Ks discriminate fraudulent from non-fraudulent firms 75% of the time and bankrupt from nonbankrupt firms 80% of the time. Mustafa & Ben Youssef (2010) investigated the relationship between the financial expertise of the audit committee and the incidence of asset misappropriation in publicly held American companies. The results support the notion that an independent audit committee member is only effective in reducing the occurrence of asset misappropriation in publicly held companies if they are also a financial expert. Kassem (2014) considered the audit literature, the perceptions of the most crucial red flags of asset misappropriation of the Egyptian auditors, and their experience regarding the most effective fraud risk response to propose a framework for external auditors that might help accurately assess and respond to factors of fraud risk that arise from asset misappropriation. Nia & Said (2015) provided a more in-depth view on the reasons for asset misappropriation by Iranian bank employees. The exploratory study adopted quantitative methods to determine why bank employees commit fraud at work. Although with some modifications, their measurement of asset misappropriation was adopted from Gates & Sullivan (2013), Achilles (2006), and Sameer Taher (2003). The survey questionnaires covered several situations, including using bank assets and facilities for personal purposes, the intention to borrow company assets to address temporary personal problems, and abuse of their positions and authority. Karim et al. (2015) assessed the possibility of asset misappropriation among Royal Malaysian Police. They designed a questionnaire consisting of two sections. The first section required respondents to fill out demographic information on their gender, age, marital status, job position, average monthly salary, education level, and the number of years of service in the department. In the second section, the respondents were asked to provide their opinion on their own normal practices regarding using office assets and facilities; only two out of 186 respondents agreed that they have committed this offense. Zager et al. (2016) examined the roles and the responsibilities of the key stakeholders of financial reporting in the prevention and detection of fraud. They further examined the methods used and types of transactions most vulnerable to fraudulent financial reporting. Hajek & Henriques (2017) examined whether an improved financial fraud detection system could be developed by combining specific features derived from financial information and managerial comments in corporate annual reports. Both intelligent feature selection and classification using a wide range of machine learning methods were employed. Thi Thu & Manh Dung (2018) examined the effects of internal control systems on asset misappropriation in Vietnamese firms. Based on questionnaires collected from internal auditors, accountants, and department managers in Vietnamese firms, the study assessed the influence of the five internal control components of the Committee of Sponsoring Organization of the Treadway Commission (COSO) on the likelihood of asset misappropriation in the relevant firms. Kazemian et al. (2019) examined the influence of the elements of fraud diamond on asset misappropriation within the Iranian banking industry. Primary data were collected through 191 survey questionnaires administered among employees of the top three Iranian banks. Koomson et al. (2020) investigated the prevalence of asset misappropriation at the workplace and examined the dominant factors that influence asset misappropriation among individuals at the workplace. Based on the theory of fraud (the stimulus/pressure, capability, opportunity, rationalization and ego (S.C.O.R.E. model)), the study examined the effect of pressure, rationalization, capability, opportunity/strength of internal control systems, and ego on asset misappropriation at the workplace while controlling for the effect of ethical values.

Additionally, several studies asserted that culture is the beliefs of the members about how the organization should behave and operate. The concept of culture is thus sometimes interchangeably used to explain organizational fraud behavior. For example, Button & Gee (2013) proposed the Fraud, Resilience and Culture Model fusing individual fraudsters, organizational structures, and cultural factors together to mitigate the fraud problem. According to Button (2008), an organization always confronts a security system with a malefactor, who is a person that commits illegal activities. However, the individuals are not in a vacuum. Furthermore, multiple factors have effects on them. In this context, organizational and industrial cultures can be highlighted in influencing whether a person indulges in fraud. Schwartz (2013) argued that the following three essential elements must be present if illegal or unethical activities within or on behalf of the corporation are to be minimized sustaining an ethical corporate culture: (1) core ethical values, (2) ethics programs, and (3) ethical leaderships.

Previous relevant studies focused primarily on questionnaire-based explorations of the risk factors of asset misappropriation in firms to strengthen their internal controls. However, no studies have focused on detecting fraud within a firm using information technology and systems, especially asset misappropriation. Moreover, most of these related studies use classical statistical methods, which means that if the feature selection method is not robust enough, it easily leads to insignificant selected features and even collinearity problems. Therefore, corporate asset misappropriation is often not effectively detected. Investment risks are hence increased.

Based on the above reasons, this study considers the financial indicators of financial structure, solvency, operating capacity, profitability, cash flow, and growth ability. The study also examines the non-financial indicators of shareholding structure, board composition, related party transactions, and management style in corporate governance as feature indicators for the detection of asset misappropriation. The indicators are first established through principal component analysis and stepwise regression. The artificial intelligence methods - support vector machine (SVM) and queen genetic algorithm (QGA) are then combined to effectively detect corporate asset misappropriation, providing a reference for investors and creditors in making decisions regarding investment, thereby reducing their investment risks. The objective of this study is achieved by (i) establishing feature indicators for detecting asset misappropriation, (ii) developing an asset misappropriation detection method, and (iii) demonstrating and evaluating the proposed method.

The rest of this paper proceeds as follows. Section 2 presents the establishment of the feature indicators for the detection of asset misappropriation. Section 3 provides the development process of the method for detecting asset misappropriation. Section 4 demonstrates the effectiveness of the proposed method. Section 5 presents the conclusion and suggestions for future research.

2. ESTABLISHMENT OF FEATURE INDICATORS FOR THE DETECTION OF ASSET MISAPPROPRIATION

To enhance the effectiveness of the detection of corporate asset misappropriation, the feature indicators for asset misappropriation detection are established as indicator variables. These variables comprise both financial and non-financial indicators. Financial indicators are mainly financial ratios computed from financial statements, whereas non-financial indicators are mainly related to corporate governance. In this study, the sets of financial and non-financial indicators required for this research are first collected through exploring the studies of financial indicators of financial statements and non-financial indicators of corporate governance. Subsequently, principal component analysis and stepwise regression are used to select financial and non-financial indicators for establishing feature indicators that are used to detect corporate asset misappropriation, as described in the subsequent sections.

2.1 Financial Indicator Collection

Financial ratios are calculated using two meaningful financial numbers and are used to analyze and interpret financial statements. Financial ratios can be used to forecast the future operating revenues of a company, determine whether a company has too much debt, and examine whether assets are being used properly, managers have misappropriated assets, or the equity of shareholders has been improperly reduced. Therefore, financial ratios are often used as a reference for investment. For instant, Kumar & Ravi (2007) argued that financial ratios are used to analyze and potentially forecast bankruptcy as a means of identifying the characteristics (in terms of financial ratios) of strongly- or poorly-performing firms and estimating their potential values. Ravisankar et al. (2011) used liquidity, profitability, safety, and efficiency as indicators for detecting financial statement fraud to judge the merits of a company based on the debt-paying ability, return on investment, and asset management of the company. Pai et al. (2011) argued that the debt structure is meaningful because debt shifts the risk from equity owners and managers to debt owners. Other items that can be easily and subjectively managed by top management sales, accounts receivable, bad debt reserves, and inventory can also affect other indicators of financial performance, including net income to fixed assets, earnings before interest and tax, and inventory to sales. These financial indicators can also be used by senior managements to detect fraud. Chen (2011a) selected 33 financial ratios as financial indicators to forecast financial distress and categorized them into five major types: earning ability, financial structure ability, management efficiency ability, management performance, and debt-repaying ability. Tinoco & Wilson (2013) examined total funds from operations to total liabilities (TFOTL), total liabilities to total assets (TLTA), no credit interval (NOCREDINT), and interest coverage (COVERAGE) as indicators of financial distress and bankruptcy to determine the financial risk and optimal debt financing of a company given a possible corporate crisis. From the financial perspective of a company, Zhang et al. (2013) indicated that regulatory authorities require companies to publicly disclose their financial state so that 26 indicators can be included. Zhang used these indicators, including the debt-paying ability, share index, quality of earnings, profitability, operating capacity, capital structure and so on, to construct a financial crisis early-warning model. These financial indicators are the sources of financial feature indicator establishment for detecting corporate asset misappropriation of this study, as summarized in Table 1.

Table 1: The Collection of Financial Indicators

Authors	Aspects	Ratios	Aspects	Ratios
Ravisankar et al. (2011)	Liquidity	• current ratio • quick ratio	Profitability	• gross profit margin • net profit margin • return on assets • return on equity
	Safety	• debt to equity • EBIT/interest • cash flow to current maturity of long-term debt	Efficiency	• accounts receivable turnover • accounts payable turnover • sales to total assets
Pai et al. (2011)	Profitability	• net profit to total assets • earnings before interest and tax • net profit to sales	Liquidity	• quick assets to current liability • working capital to total assets • current assets to current liability
	Leverage	• logarithm of total debt • total debt to total equity • total debt to total assets • long-term debt to total assets	Efficiency	• account receivable to sales • net income to fixed assets • inventory to total assets • inventory to sales • sales to total assets
Chen (2011a)	Earning	• pretax margin • return on total assets • return on equity • earnings per share • gross margin ratios	Management Efficiency	• turnover rate of inventory • turnover rate of account receivable • turnover rate of fixed assets • turnover rate of total assets • turnover rate of equity • turnover rate of working capital ratios
	Financial Structure	• debt to assets • times interest earned	Debt-Repaying	• current ratio • acid test ratio

		• book value per share • financial leverage ratio • debt to equity, short term & long term debt to book value ratio • fixed assets to total assets ratio • gross margin to total assets ratio • inventory to total assets ratio • inventory to sales ratio • investment ratio • current assets to total assets ratios		• cash ratio • cash flow ratio • cash flow to long term debt • cash flow to total debt • cash flow to short term ratio • long term debt ratio
	Management Performance	• pretax margin growth ratio • gross margin growth ratio • sales growth ratio		
Zhang et al. (2013)	Debt-paying	• liquidity ratio • quick ratio • equity ratio • EBIT/liabilities • equity/total liabilities ratio • net cash flow/total liabilities ratio • number of times interest earned	Profitability	• gross profit margin • net profit margin • return on equity • return on asset
	Share Index	• earnings per share • diluted earnings per share • funds from operations per share • net asset value per share	Operating Capacity	• inventory turnover ratio • accounts receivable turnover • current asset turnover • fixed assets turnover • total capital turnover

	Capital Structure	• debt to assets ratio • equity multiplier • current assets/total assets • non-current assets/total assets • current liabilities/total assets	
Tinoco & Wilson (2013)		• operations to total liabilities • total liabilities to total assets • the no credit interval • interest coverage	

2.2 Non-financial Indicator Collection

Corporate governance is a mechanism for supervising business managers to protect the rights and interests of shareholders and other stakeholders, strengthen the functions of the board of directors, and enhance the transparency of corporate information. The soundness of corporate governance affects the development of the enterprise. The failure of corporate governance to function can result in criminal manipulation. Indicators related to corporate governance, therefore, are typically of substantial concern to shareholders, particularly external investors, creditors, and other stakeholders. As such, they are frequently used by investors as a reference for investment. The relevant literature includes the following studies. Claessens et al. (2002) showed that the intention of controlling shareholders to plunder corporate value at the expense of minority stockholders is stronger when corporate control and the rights to share in corporate profits deviate. Chen & Du (2009) proposed a model of predicting financial distress that integrates artificial neural networks (ANN) and data mining (DM) techniques. In the analysis of the risk of financial distress, two corporate features were used to enhance the accuracy of the assessment of the probability of financial distress. Brazel et al. (2009) argued that weak corporate governance might lead to fraudulent financial statements. They applied two corporate governance features to assess fraud risk. Lin et al. (2009) examined 21 attributes including two corporate governance features to construct a hybrid failure prediction (HFP) model to assess the probability of business failure. Hsiao et al. (2010) used three board of directors-related features to build a financial distress prediction model and assess the likelihood of financial distress. Pai et al. (2011) adopted director ownership and director share pledging to detect senior management fraud. Chen (2011b) adopted 50 financial variables, including two corporate governance features and numerous macroeconomic ratios to extract suitable variables for assessing potential financial distress. Platt & Platt (2012) revealed that the number of independent board directors is positively related to the financial health of a company. Parng & Fu (2011) adopted eight corporate governance characteristics to detect firms with going-concern risks. Connelly, Limpaphayom & Nagarajan (2012) revealed that the use of a pyramidal structure negatively affects firm value after controlling for relevant factors. Kim & Upneja (2014) examined various financial ratios, board holding ratio, stock price trends, earnings per share (EPS), management practice, and changes in GDP to differentiate financially distressed restaurants from financially non-distressed restaurants. Based on these studies of corporate governance, these relevant indicators are collected as non-financial feature indicators for detecting corporate asset misappropriation, as summarized in Table 2.

Table 2: The Collection of Non-Financial Indicators

Authors	Indicators	Authors	Indicators
Claessens et al. (2002)	• control rights • cash-flow rights	Chen & Du (2009) Chen (2011b)	• the proportion of collateralized shares by the board of directors • the insider holding ratio
Brazel et al. (2009)	• the CEO is also the Chairman of the Board • percentage of insiders on the board	Lin et al. (2009)	• manager-director • director and supervisor shareholding
Hsiao et al. (2010)	• the ratio of the stock ownership of directors • the pledged shares of directors • the ratio of independent directors • the chairman of the board who concurrently acts as managing director	Pai et al. (2011)	• director ownership • pledged shares of directors
Platt & Platt (2012)	• the number of independent directors on corporate boards	Parng & Fu (2011)	• BDSIZE • INDEPENDENT • DUAL • PLE • FAMILY • SEATCON • CASHCON • BIG4
Connelly et al. (2012)	• ratio of cash flow rights to voting rights	Kim & Upneja (2014)	• board holding ratio

2.3 Feature Indicator Selection

Principal component analysis (Jolliffe 2002; Li & Zhang 2011; Tsai 2009), a statistical method, examines the correlation between variables and further linearly combines the variables. Finally, eigenvalues and eigenvectors are used to filter the types with varimax. The appropriateness of the principal component analysis is confirmed by KMO (The Kaiser-Meyer-Olkin) (Kaiser 1974; Bartlett 1951) tests. Stepwise regression (Prost et al. 2008; Ssegane et al. 2012) is a method for determining the optimal combination of independent variables. Based on the predictive variable interpretation, the effect of each variable is examined to select the predictive variables to be used in the model.

As such, 90 financial indicators listed in Table 1 are first investigated using KMO (Eq. (1)) and Bartlett tests (Eq. (2)). The 8 financial feature indicators for asset misappropriation detection from 62 financial variables are then filtered using principal component analysis (Eq. (3)), stepwise regression (Eq. (4)), and the *F* test (Eq. (5)). Subsequently, the 4 non-financial feature indicators for the detection of asset misappropriation are selected from 19 non-financial variables listed in Table 2 through stepwise regression (Eq. (4)) and the F test (Eq. (5)). The *F* test uses the *F* probability distribution to determine whether the slope $\beta_0, \beta_1, \beta_2, \cdots, \beta_k$ in the regression equation of the sample data is equal to 0 to verify the significance of the regression relationship. If the *F* value is significant, the variables in the regression can explain the dependent variables. Table 3 presents the selected financial and non-financial feature indicators for the model of the detection of corporate asset misappropriation used in this study.

$$KMO = \frac{\sum_i \sum_{j(i \neq j)} \gamma_{ij}^2}{\sum_i \sum_{j(i \neq j)} \gamma_{ij}^2 + \sum_i \sum_{j(i \neq j)} s_{ij}^2}$$

(1)

where γ_{ij} is the simple correlation coefficient between variables *i* and *j*;

s_{ij}^2 is the partial correlation coefficient between variables *i* and *j*;

H_0: correlation matrix is unit matrix;

H_1: correlation matrix is not unit matrix;

$$\begin{aligned} X^2 &= -\{n-1-[(2p+5)/6]\} ln|R_{pp}| \\ df &= p[(p-1)/2] \end{aligned} \tag{2}$$

where n is the number of samples;

p is the number of variables;

R_{pp} is the continued product of all eigenvalues or the determinant of correlation matrix;

$$X = \sum_{r=1}^{L} t_r \times \rho_r^T + E \tag{3}$$

where t_r is the score of the r-th principal component, where the score is the projection of data points on the loading direction;

ρ_r is the loading of the r-th principal component;

L is the number of the selected principal component;

$$y_i = \beta_0 + \beta_1 \times x_{1i} + \beta_2 \times x_{2i} + \beta_3 \times x_{3i} + \cdots + \beta_k \times x_{ki} + \varepsilon_i \text{ , } i = 1 \cdots n, k > 0 \tag{4}$$

where y_i denotes whether the i-th company has misappropriation; if so, the value appears 1, otherwise it is 0;

x_{ki} is the k-th indicator of the i-th company;

β_0 is the intercept in multiple regression model;

$\beta_0, \cdots, \beta_k$ are the regression coefficients;

n is the number of companies;

k is the number of all indicators;

$$H_0 : \beta_1 = \beta_2 \cdots = \beta_n = 0$$
$$H_1 : Not\, all\, \beta_1 = 0,\ i = 1, 2, 3, \cdots, n$$

$$F = \frac{MSR}{MSE},\ MSR = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{k},\ MSE = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n - k - 1} \tag{5}$$

where MSE denotes mean squares due to error;

MSR denotes mean squares due to regression;

$y_i - \hat{y}_i$ is the deviation of y_i towards the fitted regression line;

Table 3: The Selected Financial and Non-financial Feature Indicators for Corporate Asset Misappropriation Detection

Financial Feature Indicators	Non-Financial Feature Indicators
• total debt to total assets • working capital to total assets • inventory turnover ratio • earnings per share • fixed assets to total assets ratio • cash flow ratio • cash flow to short term ratio • funds from operations per share	• the ratio of the total pledged shares to the total on-hand shares of the BOD and supervisors • director and supervisor shareholding • the number of independent directors on corporate boards • ratio of the right for cash flow over the right of control

Note: Readers can request the statistical results by the author's email.

3. DEVELOPMENT OF A CORPORATE ASSET MISAPPROPRIATION DETECTION METHOD

SVM is a classification algorithm with high accuracy that generates a separating hyperplane through training data to distinguish different data categories, which maximizes the margins in the different categories' data for attribute classification problems in data mining (Joachims 1998; Tong & Koller 2001). Based on the selected feature indicators discussed in Section 2, the radial basis function kernel-centered SVM (Chen et al. 2019; Chen et al. 2017; Cortes & Vapnik 1995; Harris 2015; Li et al. 2014; Shen et al. 2016; Vafeiadis et al. 2018; Wang et al. 2015; Ramakalyan et al. 2016; Yazdani et al. 2019; Yuan et al. 2010) and QGA (Chen et al. 2019; Chen et al. 2017; Stern et al. 2006) are integrated into a single algorithm for the detection of corporate asset misappropriation. The optimized SVM parameters are determined through QGA to enhance the accuracy of the detection model effectively. Figures 1 and 2 depict the ensemble classifier - Adaboost.M2 model and the algorithm for detecting corporate asset misappropriation, respectively. The related calculations are presented in Eqs. (6), (7), (8), and (9). To mimic evolution, the QGA used the operators of selection, crossover, and mutation along with a fitness function that determined the performance of a candidate solution (Farhan et al. 2015). In the study, the selection mechanism used was rank selection. A common operator, the *Z*-point crossover was applied in one of the variations in this research experiment. Moreover, the mutation operation was utilized on the population after the crossover operator was performed on the previous generation. For each chromosome in the population, each gene was changed with the common mutation rate value, $p_m \in [0.005, 0.01]$. In conclusion, the evolution of the generation was stopped when 90% of the generations took the same fitness (Depczynski et al. 2000) for a typical run.

$$D_{m+1} = M(q_i \times d_i) \tag{6}$$

$$F(d_i) = rank(D_{m+1}) \tag{7}$$

where q_i denotes a queen chromosome randomly selected from the queen cohort;

d_i denotes a chromosome randomly selected from the population;

D_{m+1} denotes the mutation of queen chromosomes and chromosomes;

$F(d_i)$ denotes the individual fitness values of chromosomes based on the rank selection;

$$f(x) = sign(\sum_{i=1}^{n} a_i y_i K(x_i, x_j) + b) \tag{8}$$

$$K(x_i, x_j) = exp\left(-\gamma \|x_i, x_j\|^2\right), \gamma > 0 \tag{9}$$

where $f(x)$ represents the optimal decision function;

a_i represents the Lagrange multiplier;

y_i represents the class index of various indicators;

b represents the bias;

$K(x_i, x_j)$ represents the radial basis function (RBF);

γ represents the parameter of the RBF, and this parameter is defined by the real valued gene;

An SVM is acquired after many iterations. The weight voting for the SVM is performed based on the weighting required to generate the QGA-SVM model. Equation (10) presents the formula for weight voting.

$$H(x) = \arg\max \sum_t (ln \frac{1}{\beta_t}) h_t(x, y) \tag{10}$$

where $H(x)$ denotes the class index of the QGA-SVM model;

$h_t(x, y)$ denotes the class index of the SVM;

β_t denotes the weight of the SVM;

In conclusion, the testing dataset is input into the classification model - QGA-SVM to determine whether the misappropriation of corporate assets exists.

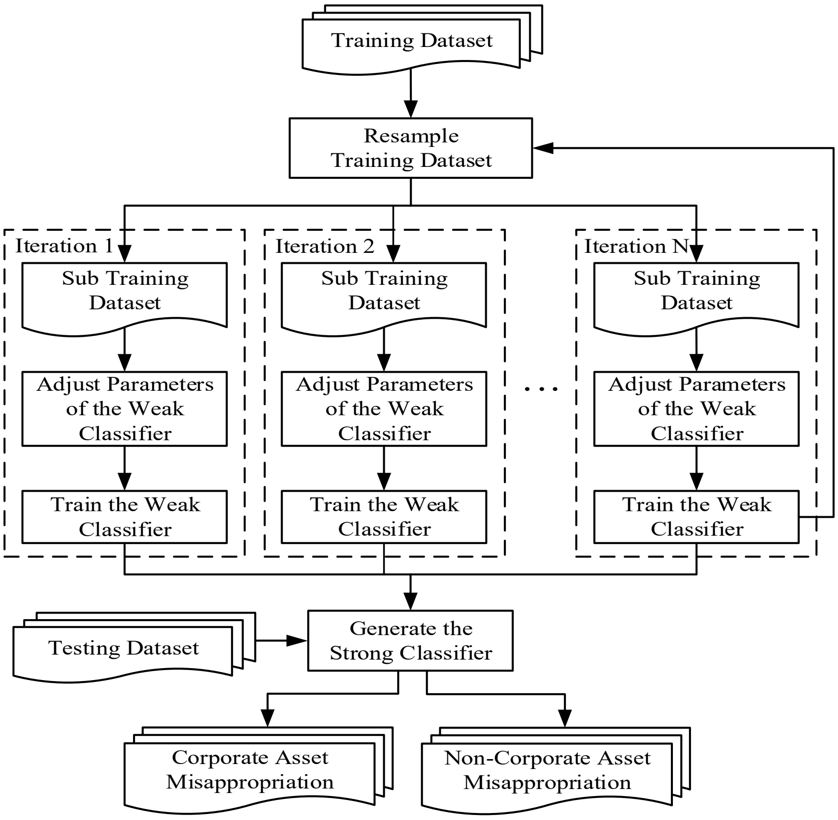

Figure 1: The Ensemble Classifier Model

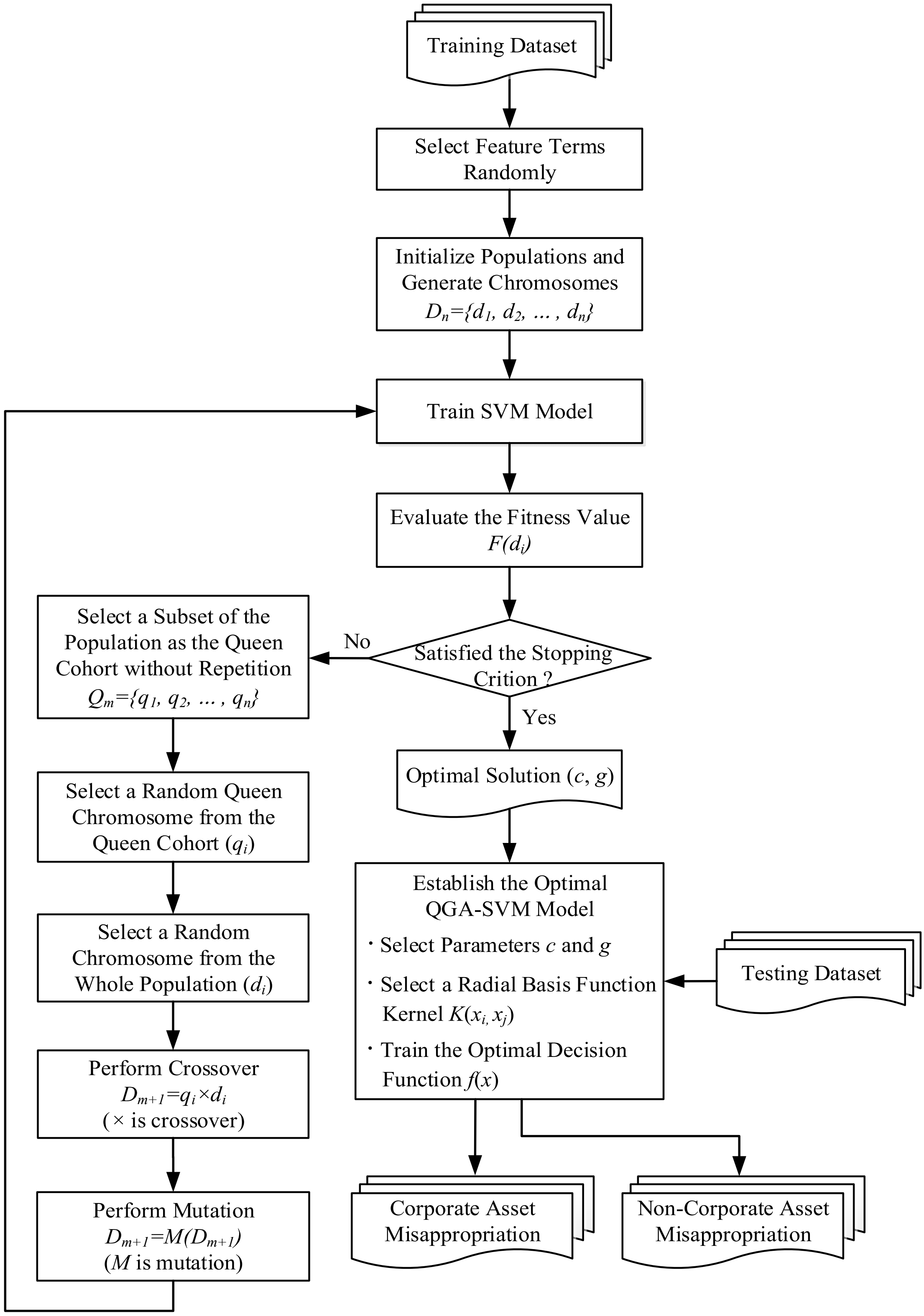

Figure 2: Algorithm for Detecting Corporate Asset Misappropriation

4. DEMONSTRATION AND EVALUATION OF THE PROPOSED ASSET MISAPPROPRIATION DETECTION METHOD

In this section, MATLAB 2014a is first employed to implement the proposed method of detecting corporate asset misappropriation. The financial and corporate governance indicators of listed Taiwanese companies are sampled to verify the feasibility of the proposed method. The detection accuracy is subsequently evaluated by comparing the proposed method with other detection models to ensure its effectiveness.

4.1 Demonstration of the Proposed Method

The listed Taiwanese companies are presented as illustrative examples to demonstrate the feasibility of the proposed corporate asset misappropriation detection method. The detailed steps are as follows.

1. Establishment of Datasets

According to the reports on the listed companies that have experienced asset misappropriation recorded in the Taiwan Economic Journal (TEJ) (http://www.tej.com.tw/twsite/) and a 1:2 sample ratio for asset misappropriation companies and non-asset misappropriation companies (Chen et al. (2017); Chen et al. (2019)), 42 companies that experienced asset misappropriation in the period from 1996 through 2019 are selected. Their financial and corporate governance indicators are further determined for the current and the previous three years. During the same period, 84 companies that did not commit asset misappropriation are selected from the same industry and similar total assets as that of the companies that experienced asset misappropriation. The frequency of asset misappropriation and the number of implicit cases of asset misappropriation among the collected samples are larger in the period from 1997 through 2002 than it is in other periods. Therefore, the training dataset is sampled for the period from 1997 through 2000. Further, 54 financial and corporate governance indicators of 19 companies that experienced asset misappropriation and 114 financial and corporate governance indicators of 44 companies that did not experience asset misappropriation are determined from the modules of IFRS finance and corporate governance in the TEJ database. The testing dataset is sampled for the period 2001-2019, of which 45 financial and corporate governance indicators of 23 companies that experienced asset misappropriation and 86 financial and corporate governance indicators of 39 companies that did not experience asset misappropriation are determined from the modules of IFRS finance and corporate governance in the TEJ database.

2. Detecting Corporate Asset Misappropriation

The datasets established in Step 1. are input into the proposed QGA-SVM model (Fig. 1) for the training and testing of the detection of fraud. Table 4 illustrates the parameter settings. This study uses the sign test (Glancy & Yadav, 2011) to evaluate the clustering results. The null hypothesis is that the QGA-SVM model cannot discriminate between corporate and non-corporate asset misappropriation; the probability distribution in each cluster is 0.5. Table 5 indicates that the classification results can help to identify potential asset misappropriation at a >0.01 significance level.

Table 6 adopts the type I and type II error to represent the classification results of corporate and non-corporate asset misappropriation (Ravisankar et al. 2011; Zhou & Kapoor 2011; Chen et al. 2017; Chen et al. 2019). The p-value for 19 out of the 23 companies that experienced misappropriation classification correctly is 0.0243; the result for 31 out of the 36 companies that did not experience asset misappropriation classification correctly is p=2.9032x10^{-7}. Therefore, the testing of the proposed detection model of this study demonstrates that it discriminates at a significant level between corporate and non-corporate asset misappropriation.

Table 4: Parameter Settings for the QGA-SVM Detection Model

Parameter Name	Value Set
QGA Population	50
QGA Evolution	200
QGA Threshold	0.9
SVM Cross Validation	5
c and g of SVM	Based on the results of QGA

Table 5: QGA-SVM Training Results and Detection at a Significance Level of 0.01

Training Sample	Total	Correctly Identified	Incorrectly Identified	P-value	Accuracy (%)
Companies	59	52	7	1.2274x10^{-23}	88.1356

Table 6: QGA-SVM Testing Results and Detection at a Significance Level of 0.01

Testing Sample	Total	Correctly Identified	Incorrectly Identified	P-value	Upper	Lower	Accuracy (%)
Asset Misappropriation Companies	23	19	4	0.0243	18	4	82.6087
Non-Asset Misappropriation Companies	39	34	5	2.9033x10^{-7}	34	15	87.1795

4.2 Comparison of Detection Accuracy

In this section, the training samples are first divided into "included corporate governance indicators" and "non-included corporate governance indicators," which are input into the QGA-SVM detection model adopted in this study. The purpose is to determine the effects of corporate governance indicators on the accuracy of the detection of corporate asset misappropriation. As presented in Table 7, corporate governance indicators can effectively enhance the accuracy of the detection of corporate asset misappropriation. Furthermore, in order to prove that QGA-SVM has better classification accuracy, the samples described in Section 4.1 are used to compare the accuracy of the most common types of classification algorithms - K-Nearest Neighbors (KNN), Naïve Bayes, Logistic Regression, Decision Tree, Random Forest, CART, PSO-SVM, GA-SVM, Grid-SVM and QGA-SVM in detecting corporate asset misappropriation. Table 8 presents the accuracy results of these models. The comparison results indicate that the accuracy of the QGA-SVM detection model used in this study is higher than that of the other classification models.

Table 7: The Effect of Corporate Governance Indicator on Detection Accuracy

Detection Model	Dataset	*c*	*g*	Accuracy (%)
QGA-SVM	Included Corporate Governance	4.1941	6.9895	82.6149
	Non-Included Corporate Governance	6.1082	5.7145	68.1022

Table 8: Accuracy Comparison for the Classifiers

Classifier	*c*	*g*	Accuracy (%)
KNN	---	---	77.6180
Naïve Bayes	---	---	79.4371
Logistic Regression	---	---	65.6802
Decision Tree	---	---	78.4465
Random Forest	---	---	80.4589
CART	---	---	71.2468
PSO-SVM	8.2637	1.0893	74.2345
GA-SVM	5.2644	8.4151	78.9465
Grid-SVM	5.2780	0.0068	72.7493
QGA-SVM	**4.1943**	**6.9893**	**83.0564**

5. CONCLUSION AND FUTURE RESEARCH

This study considers the financial feature indicators of financial structure, solvency, operating capacity, profitability, cash flow, and growth ability. It also examines the non-financial feature indicators of shareholding structure, board composition, related party transactions, and management style in corporate governance. Further, the study establishes feature indicators for the detection of corporate asset misappropriation using principal component analysis and stepwise regression. Subsequently, SVM and QGA algorithms are integrated into a detection model for corporate asset misappropriation. The main results and contributions of this study are summarized as follows.

- The feature indicators for corporate asset misappropriation detection: For the existing financial and non-financial indicators used in corporate fraud research, the feature indicators for detecting corporate asset misappropriation were established using the techniques of principal component analysis and stepwise regression. The established feature indicators can be used to detect corporate asset misappropriation in other countries.
- The model for corporate asset misappropriation detection: To effectively detect the fraud of corporate asset misappropriation, the algorithm of QGA-SVM was developed by training the fraudulent and non-fraudulent listed Taiwanese companies. According to the experimental results, the accuracy of the detection of corporate asset misappropriation is higher than 80%. The developed model can also be used to detect the various types of fraud, such as tax evasion and insurance fraud.
- The methodology for developing a corporate asset misappropriation detection method: This study provides a valuable reference model that can be applied for the development of other fraud detection methods (such as financial statement fraud) or financial (crisis) forecasting.

In the practical experiment, the detection accuracy of the proposed method in the study was better than that of other methods, considering the fact that corporate governance indicators were shown to increase accuracy in the detection of corporate asset misappropriation by 14%.

Although it is the most common type of occupational fraud, especially research on the use of information technology to detect asset misappropriation, asset misappropriation was given less attention in previous literature as well as in the audit practice. The results of this study realize the use of information technology to detect corporate asset misappropriation and enhance fraud detection accuracy to provide investors and creditors with a reference for making decisions regarding investment and thereby increase investment profitability.

However, some limitations on developing and using the proposed approach are identified. First, several studies show that the fraud risk of asset misappropriation is related to the corporate culture of different countries. Therefore, the relevant factors of the corporate culture in different countries were not considered in developing the fraud detection method. Furthermore, the experimental samples in the study were all from the listed companies in Taiwan; the results of the experiment cannot explain the effectiveness of the detection of corporate asset misappropriation applicable to various countries.

Based on the developed approach in this study, future research should address three directions.

- Other external indicators that may also affect the detection of corporate asset misappropriation, such as industrial environment and macroeconomics, should be explored to detect corporate asset misappropriation more objectively.
- Financial news and public opinion should also be examined to reflect enterprise operating conditions objectively. Therefore, financial news and data on public opinion should be retrieved from online social media and used to detect the possibility of corporate asset misappropriation.
- Textual analysis can be used to analyze the negatives and uncertainties of textual financial and non-financial information (such as annual reports and litigation documents) to detect whether the company is fraudulent or not.

REFERENCES

Achilles, W. W. (2006). An experimental analysis of the impact of goal orientation, ethical orientation, and personality traits on managers' and accountants' abilities to recognize misappropriation of assets, *Dissertation, Business, Virginia Commonwealth University*.

Akman, E., Karaman, A. S. & Kuzey, C. (2020). Visa trial of international trade: evidence from support vector machines and neural networks, *Journal of Management Analytics*, 7(2), 231-252.

Bartlett, M. S. (1951). A further note on tests of significance in factor analysis, *British Journal of Statistical Psychology*, 4(1), 1-2.

Brazel, J. F., Jones, K. L. & Zimbelman, M. F. (2009). Using nonfinancial measures to assess fraud risk, *Journal of Accounting Research*, 47(5), 1135-1166.

Button, M. (2008). *Doing security: critical reflections and an agenda for change*, Palgrave MacMillan.

Button, M. & Gee, J. (2013). Countering fraud for competitive advantage, *John Wiley & Sons Ltd*, *Chichester*.

Cecchini, M., Aytug, H., Koehler, G. J. & Pathak, P. (2010). Making words work: using financial text as a predictor of financial events, *Decision Support Systems,* 50(1), 164-175.

Chen, M. Y. (2011a). Predicting corporate financial distress based on integration of decision tree classification and logistic regression, *Expert Systems with Applications*, 38(9), 11261-11272.

Chen, M. Y. (2011b). Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches, *Computers & Mathematics with Applications*, 62(12), 4514-4524.

Chen, W. S. & Du, Y. K. (2009). Using neural networks and data mining techniques for the financial distress prediction model, *Expert Systems with Applications*, 36(2), Part 2, 4075-4086.

Chen, H. J., Huang, S. Y. & Kuo, C. L. (2009). Using the artificial neural network to predict fraud litigation: some empirical evidence from emerging markets, *Expert Systems with Applications*, 36(2), Part 1, 1478-1484.

Chen, Y. J., Liou, W. C., Chen, Y. M. & Wu, J. H. (2019). Fraud detection for financial statements of business groups, *International Journal of Accounting Information Systems*, 32, 1-23.

Chen, Y. J., Wu, C. H., Chen, Y. M., Li, H. Y. & Chen, H. K. (2017). Enhancement of fraud detection for narratives in annual reports, *International Journal of Accounting Information Systems*, 26, 32-45.

Claessens, S., Djankov, S., Fan, P. H. & Lang, H. P. (2002). Disentangling the incentive and entrenchment effects of large shareholdings, *The Journal of Finance*, 57(6), 2741-2771.

Connelly, J. T., Limpaphayom, P. & Nagarajan, N. J. (2012). Form versus substance: the effect of ownership structure and corporate governance on firm value in Thailand, *Journal of Banking & Finance*, 36(6), 1722-1743.

Cortes, C. & Vapnik, V. (1995). Support-vector networks, *Machine Learning*, 20(3), 273-297.

Depczynski, U., Frost, V. J. & Molt, K. (2000). Genetic algorithms applied to the selection of factors in principal component regression, *Analytica Chimica Acta*, 420(2), 217-227.

Farhan, A. S., Tareq, W. Z. & Awad, F. H. (2015). Solving *N* queen problem using genetic algorithm, *International Journal of Computer Applications*, 122(12), 11-14.

Gates, S. & Sullivan, C. (2013). Asset misappropriation and the death of a mayor, *Journal of Business Cases & Applications*, 8.

Glancy, F. H. & Yadav, S. B. (2011). A computational model for financial reporting fraud detection, *Decision Support Systems*, 50(3), 595-601.

Goel, R. (2013). Integrating a systems approach to detect risk of insider trading, *International Journal of Humanities and Social Science*, 3(19), 21-26, 2013.

Hajek, P. & Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud - a comparative study of machine learning methods, *Knowledge-based Systems*, 128, 139-152.

Harris, T. (2015). Credit scoring using the clustered support vector machine. *Expert Systems with Applications*, 42(2), 741-750.

Hsiao, H. F., Lin, S. H. & Hsu, A. C. (2010). Earnings management, corporate governance, and auditor's opinions: a financial distress prediction model, *Investment Management and Financial Innovations*, 7(3), 29-40.

Joachims, T. (1998). Text categorization with support vector machines: learning with many relevant features, *European Conference on Machine Learning, Springer*, 137-142.

Jolliffe, I. T. (2002). Principal component analysis, *Springer Series in Statistics*, Edition 2.

Kaiser, H. F. (1974). An index of factorial simplicity, *Psychometrika*, 39(1), 31-36.

Karim, Z. A., Said, J. & Bakri, H. H. M. (2015). An exploratory study on the possibility of assets misappropriation among Royal Malaysian police officials, *Procedia Economics and Finance*, 31, 625-631.

Kazemian, S., Said, J., Hady Nia, E. & Vakilifard, H. (2019), Examining fraud risk factors on asset misappropriation: evidence from the Iranian banking industry, *Journal of Financial Crime*, 26(2), 447-463.

Kassem, R. (2014). Detecting asset misappropriation: a framework for external auditors, *International Journal of Accounting Auditing and Performance Evaluation*, 10(1), 1-42.

Kim, S. Y. & Upneja, A. (2014). Predicting restaurant financial distress using decision tree and AdaBoosted decision tree models, *Economic Modeling*, 36, 354-362.

Koomson, T. A. A., Owusu, G. M. Y., Bekoe, R. A. & Oquaye, M. (2020), Determinants of asset misappropriation at the workplace: the moderating role of perceived strength of internal controls, *Journal of Financial Crime*, 27(4), 1191-1211.

Kumar, P. R. & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques - a review, *European Journal of Operational Research*, 180(1), 1-28.

Lehman, C. M. (2015). Asset misappropriation schemes: short cases for use in the classroom, *Journal of Forensic & Investigative Accounting*, 7(2), 340-362.

Li, H., Li, C. J., Wu, X. J. & Sun, J. (2014). Statistics-based wrapper for feature selection: an implementation on financial distress identification with support vector machine, *Applied Soft Computing*, 19, 57-67.

Li, Y. & Zhang, Q. (2011). The application of principal component analysis on financial analysis in real estate listed company, *Procedia Engineering*, 15, 4499-4503.

Lin, R. H., Wang, Y. T., Wu, C. H. & Chuang, C. L. (2009). Developing a business failure prediction model via RST, GRA and CBR, *Expert Systems with Applications*, 36(2), Part 1, 1593-1600.

Nia, E. H. & Said, J. (2015). Assessing fraud risk factors of assets misappropriation: evidences from Iranian banks, *Procedia Economics and Finance*, 31, 919-924.

Mustafa, S. T. & Ben Youssef, N. (2010). Audit committee financial expertise and misappropriation of assets, *Managerial Auditing Journal*, 25(3), 208-225.

Olmo, J., Pilbeam, K. & Pouliot, W. (2011). Detecting the presence of insider trading via structural break tests, *Journal of Banking & Finance*, 35(11), 2820-2828.

Pai, P. F., Hsu, M. F. & Wang, M. C. (2011). A support vector machine-based model for detecting top management fraud, *Knowledge-Based Systems*, 24(2), 314-321.

Park, Y. S. & Lee, J. (2010). Detecting insider trading: the theory and validation in Korea Exchange, *Journal of Banking & Finance*, 34(9), 2110-2120.

Parng, Y. J. & Fu, C. J. (2011). Detecting firms with going-concern risk based on the industry affiliation, corporate governance characteristics, and financial performance, *African Journal of Business Management*, 5(22), 9249-9261.

Platt, H. and Platt, M. (2012). Corporate board attributes and bankruptcy, *Journal of Business Research*, 65(8), 1139-1143.

Prost, L., Makowski, D. and Jeuffroy, M. H. (2008). Comparison of stepwise selection and Bayesian model averaging for yield gap analysis, *Ecological Modelling*, 219(1-2), 66-76.

Ramakalyan, A., Sivakumar, A., Aravindan, C., Kannan, K., Swaminathan, V. & Sarala, D. (2016). Development of KSVGRNN: A hybrid soft computing technique for estimation of boiler flue gas components, *Journal of Industrial Information Integration*, 4, 42-51.

Ravisankar, P., Ravi, V., Rao, G. R. & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques, *Decision Support Systems*, 50(2), 491-500.

Sameer Taher, M. (2003). Misapprpriation of assets: a test of SAS no. 82 risk factors, *Thesis*, *Cleveland State University.*

Schwartz, M. (2013). Developing and sustaining an ethical corporate culture: the core elements, *Bus. Horiz*, 56, 39-50.

Shen, L., Wang, H., Xu, L., Ma, X., Chaudhry, S. & He, W. (2016). Identity Management based on PCA and SVM, *Information Systems Frontiers*, 18, 711-716.

Ssegane, H., Tollner, E. W., Mohamoud, Y. M., Rasmussen, T. C. & Dowd, J. F. (2012). Advances in variable selection methods I: causal selection methods versus stepwise regression and principal component analysis on data of known and unknown functional relationships, *Journal of Hydrology*, 438-439, 16-25.

Stern, H., Chassidim, Y. & Zofi, M. (2006). Multi-agent visual area coverage using a new genetic algorithm selection scheme, *European Journal of Operational Research*, 175(3), 1890-1907.

Thi Thu, H. L. & Manh Dung, T. (2018). The effect of internal control on asset misappropriation: the case of Vietnam, *Business and Economic Horizons (BEH)*, 14(4), 941-953.

Tinoco, M. H. & Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables, *International Review of Financial Analysis*, 30, 394-419.

Tong, S. & Koller, D. (2001). Support vector machine active learning with applications to text classification, *Journal of Machine Learning Research*, 45-66.

Tsai, C. F. (2009). Feature selection in bankruptcy prediction, *Knowledge-Based Systems*, 22(2), 120-127.

Vafeiadis, T., Dimitriou, N., Ioannidis, D., Wotherspoon, T., Tinker, G. & Tzovaras, D. (2018). A framework for inspection of dies attachment on PCB utilizing machine learning techniques, *Journal of Management Analytics*, 5(2), 81-94. Wang, X., Chen, X. & Bi, Z. (2015). Support vector machine and ROC curves for modeling of aircraft fuel consumption, *Journal of Management Analytics*, 2(1), 22-34.

Yazdani, M., Babagolzadeh, M., Kazemitash, N. & Saberi, M. (2019). Reliability estimation using an integrated support vector regression-variable neighborhood search model, *Journal of Industrial Information Integration*, 15, 103-110.

Yuan, R., Li, Z., Guan, X. & Xu, L. (2010). An SVM-based machine learning method for accurate Internet traffic classification, *Information Systems Frontiers*, 12, 149-156.

Zager, L., Malis, S. S. & Novak, A. (2016). The role and responsibility of auditors in prevention and detection of fraudulent financial reporting, *Procedia Economics and Finance*, 39, 693-700.

Zhang, L., Zhang, L., Teng, W. & Chen, Y. (2013). Based on information fusion technique with data mining in the application of finance early-warning, *Procedia Computer Science*, 17, 695-703.

Zhou, W. & Kapoor, G. (2011). Detecting evolutionary financial statement fraud, *Decision Support Systems*, 50(3), 570-575.