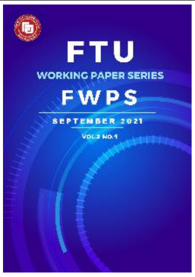

Working Paper 2021.2.3.09
- Vol 2, No 3

PHÁT HIỆN SỰ BẤT THƯỜNG TRONG BÁO CÁO TÀI CHÍNH: NGHIÊN CỨU TRƯỜNG HỢP VIỆT NAM

Vương Ngọc Quỳnh[1]
Sinh viên K56 CLC Kế toán Kiểm toán định hướng nghề nghiệp ACCA – Khoa Kế toán Kiểm toán
Trường Đại học Ngoại thương, Hà Nội, Việt Nam

Đặng Thị Huyền Hương
Giảng viên Khoa Kế toán Kiểm toán
Trường Đại học Ngoại thương, Hà Nội, Việt Nam

Tóm tắt

Sự bất thường trong báo cáo tài chính là vấn đề thường thấy ở Việt Nam. Tính minh bạch thấp trong báo cáo tài chính không chỉ đe dọa lợi ích công mà còn có thể làm suy yếu triển vọng của một quốc gia vì nó tạo ra một môi trường hoàn hảo để che giấu các vụ gian lận. Sự thiếu minh bạch ở Việt Nam, cùng với môi trường kinh doanh đang thay đổi nhanh chóng trong thời đại dữ liệu lớn, đòi hỏi các phương pháp hiệu quả hơn để ngăn ngừa và kiểm soát các sai sót trong báo cáo tài chính. Nghiên cứu này tìm hiểu khả năng áp dụng các phương pháp phát hiện bất thường tiên tiến vào báo cáo tài chính của các doanh nghiệp niêm yết tại Việt Nam. Phương pháp khai thác dữ liệu phân loại, cụ thể là bằng hồi quy logistic và máy vector hỗ trợ, được sử dụng để dự đoán sự bất thường trong báo cáo tài chính của 790 công ty niêm yết trên HOSE, HNX và UPCoM vào năm 2020. Trong tổng số 790 quan sát, có 206 quan sát bất thường với chênh lệch lợi nhuận sau thuế trên 5% trước và sau kiểm toán. Hai máy phân loại đạt được độ chính xác trung bình 70% với dữ liệu mất cân đối này, cho thấy rằng các phương pháp khai thác dữ liệu là hữu ích trong việc phát hiện sớm các báo cáo tài chính bất thường ở Việt Nam. Việc ứng dụng công nghệ chắc chắn là rất quan trọng cho cuộc chiến chống lại sự thiếu minh bạch trên thị trường tài chính. Tuy nhiên, công nghệ chỉ có thể phát huy hiệu quả nếu việc đào tạo và giáo dục được chú trọng. Trên tất cả, hành động của chính phủ có thể là phương tiện hiệu quả nhất để cản trở hoặc thúc đẩy sự minh bạch tài chính của một quốc gia.

Từ khóa: bất thường tài chính, gian lận tài chính, phát hiện bất thường, khai thác dữ liệu phân loại.

DETECTION OF IRREGULARITIES IN FINANCIAL STATEMENTS: THE CASE OF VIETNAM

Abstract

[1] Tác giả liên hệ, Email: k56.1718820067@ftu.edu.vn

Financial statement irregularities are recurrent issues in Vietnam. Low transparency in financial reporting not only threatens domestic public interest but can also undermine the prospects of a country as it creates a perfect environment to conceal frauds. The lack of transparency in Vietnam, coupled with a rapidly changing business environment in the age of big data, calls for more effective methods of preventing and controling irregularities in financial statement. This study aimed to explore the applicability of advanced financial statement irregularity detection methods for publicly listed enterprises in Vietnam. The data mining classification method, by logistic regression and support vector machine in particular, was used to predict irregularity in the financial statements of 790 companies listed on HOSE, HNX, and UPCoM in 2020. Of the total 790 observations, 206 were irregulars whose audited profits after tax deviate over 5% from the unaudited numbers. The two classifiers achieved an average prediction accuracy of 70% on this imbalanced data, suggesting that data mining methods are useful for the early detection of financial statement irregularity in Vietnam. The use of technology is undoubtedly crucial for the fight against opacity in the financial market. However, technology can only be effective if there are adequate training and education. Above all, the actions of the government can be the most effective ways to either hinder or facilitate the financial transparency of a nation.

Keywords: financial irregularities, financial fraud, anomaly detection, data mining classification.

1. Introduction

Financial statement irregularity in general and financial statement fraud is an ongoing global issue. While financial statement fraud is the least common fraud type, they are the most costly form of fraud. According to the Association of Certified Fraud Examiners (ACFE), in 2019, financial statement frauds made up only 10% of fraud cases but caused a median loss of USD 954,000 while other forms of fraud caused a median loss of no more than USD 250,000 (ACFE, 2020). Financial statement frauds can have significant negative effects on the economy, to such an extent that one scandal had led to a regulation reform in the United States (the Enron scandal). In addition to the economy and society, financial statement frauds negatively affect all parties involved: the perpetrators, the stakeholders, the companies, and the local communities. Consequences of frauds range from criminal charges, career and business losses to rising unemployment, declining tax revenue and low market confidence (Zahra, 2005).

The prevention and detection of financial statement frauds are arduous as this type of fraud is often committed by the top managers, who have better means to conceal their actions, and therefore can go uncovered for prolonged periods of time, possibly for years (ACFE, 2020). Despite the many anti-fraud controls available, from audits, fraud training to data monitoring, most cases of fraud are uncovered because they were tipped off and active measures like audit are much less effective (ACFE, 2020). The situation is even more alarming in Vietnam, where most fraud cases were initially detected thanks to tipping or by accidents (PwC, 2018). A survey on information disclosure on the securities market in Vietnam showed that only 45.13% of the surveyed firms disclosed their information in a timely manner (VAFE, 2020). As of 10 April 2021, 643 publicly traded firms in Vietnam were found to have discrepancies between their audited and unaudited financial statements (Vietstock, 2021). The lack of transparency raises doubts about the competency of the accountants, auditors, and regulators alike in Vietnam.

Research on the detection of frauds or irregularities with advanced data analysis is necessary in the new age of technology, of complex and abundant data, as common fraud assessment methods such as ratio analysis would no longer be adequate. Advanced detection methods offer a greater cost benefit trade-off as they reduce processing time and can analyse large amounts of information beyond the capability of humans. This study aims to determine the applicability of advanced anomaly detection methods, data mining classification specifically, in analysing financial information in Vietnam. The research findings could facilitate more informed economic decisions by financial statement users in Vietnam and aid the accountants or auditors in incorporating intelligent decision support systems.

2. Theoretical background

2.1. *Irregularity in financial statement*

In the context of accounting and financial statement, the Institute of Chartered Accountants in England and Wales (ICAEW) refers to "irregularities" as "instances of non-compliance with laws and regulations" (ICAEW, 2021). Following this definition, “irregularity in financial statement” can be defined as “an instance of misstatement in financial statement”, where "misstatement" is defined as the difference between what is reported on the financial statement and what should be reported in order to conform with laws and regulations by the International Federation of Accountants (IFAC) in their International Standards on Auditing (ISA) (IFAC ISA 450, 2009).

Misstatements can arise from either error or fraud, the distinguishing point being whether the underlying action which created the misstatement is unintentional or intentional. It is crucial that research differentiate between the two types of misstatement, as misstatements caused by fraud are more likely to spur negative market reactions and legal actions (Hennes et al., 2007). However, prior fraudulent financial statement detection research in Vietnam has mostly examined misstatements while referring to them as frauds, which is highly misleading. Due to a lack of data on fraud cases in Vietnam, this study examines instances of material misstatements in general, which significantly compromise the credibility of the financial statement, hereafter referred to as irregularity in financial statement.

2.2. Irregularity detection in financial statement

2.2.1. Financial statement irregularity detection methods

Detecting financial statement irregularities is a task and even a responsibility of many professionals such as the financial officers, auditors, and tax authority, etc. Though their positions differ, these professionals may employ the same methods to detect financial statement irregularities. According to Kim et al. (2009), the main financial statement irregularity detection methods are: (1) database queries, (2) ratio analysis, (3) audit sampling, (4) digital analysis, (5) regression or analysis of variance, and (6) data mining classification. Methods (4), (5), and (6) are advanced methods that heavily involve mathematics. Though the advanced methods on average have a higher effectiveness, they are not commonly used due to a lack of necessary company resources and their perceived difficulty (Bierstaker et al., 2006; Kim et al., 2009). This study focuses on advanced detection methods. However, the digital analysis method often requires a large amount of internal transaction data limited to company insiders, e.g. Benford’s law (Nigrini, 2012), and therefore is left out from this study.

2.2.2. Financial statement irregularity detection models

(a) Financial statement irregularity detection models by regression

Financial statement irregularity detection models by regression commonly used logistic regression (LR). The study focuses on LR models that are common in current financial statement irregularity research in Vietnam, namely the fraud triangle based models and F-score model.

3. Fraud triangle based models

These models are referred to as "fraud triangle based" as the fraud triangle itself is not a model but a theory coined by Donald Cressey in the 1950s. The theory is widely applied in the development of fraud prediction models as well as various theories on the motivations behind fraudulent behaviours (Dorminey et al., 2012). Following are some fraud triangle based models from influential research in Vietnam.

Model from Tran et al. (2015):

$$\mathbf{FRAUD = -2.387 - 0.065\ SALAR - 3.446\ INVTA + 3.517\ LEV + 1.183\ BIG4}$$
$$\mathbf{+\ 2.259\ AUDREPORT + 1.052\ RST + \varepsilon\ .}$$

Model from Nguyen et al. (2018):

$$\mathbf{FRAUD = -2.215 - 0.661\ REVTA - 19.908\ ROA - 0.119\ EDU + 0.634\ AUDITOR}$$
$$\mathbf{+\ 3.121\ REPORT + \varepsilon\ .}$$

The predictor variables are proxies of the three factors in the fraud triangle theory, which are the pressure/incentive to commit fraud, the opportunity to commit fraud, and the rationalisation/attitude justifying the action.

4. F-score model

Dechow et al. (2011) developed the F-score model for the detection of material misstatements in financial statement in the United States. Below is the F-score models by Dechow et al. (2011), including the baseline model and the full model with non-financial and market-related variables.

Baseline model:

$$\mathbf{MISSTATEMENT = -7.893 + 0.790\ rsst\_acc + 2.518\ ch\_rec + 1.191\ ch\_inv}$$
$$\mathbf{+\ 1.979\ soft\_assets + 0.171\ ch\_cs - 0.932\ ch\_roa + 1.092\ issue + \varepsilon\ .}$$

Full model:

$$\mathbf{MISSTATEMENT = -7.966 + 0.909\ rsst\_acc + 1.731\ ch\_rec + 1.447\ ch\_inv}$$
$$\mathbf{+\ 2.265\ soft\_assets + 0.160\ ch\_cs - 1.455\ ch\_roa + 0.651\ issue - 0.121\ ch\_emp}$$
$$\mathbf{+\ 0.345\ leasedum + 0.082\ rett + 0.098\ rett{-}1 + \varepsilon\ .}$$

These models were developed and evaluate using the same sample. By performing an out-of-sample test for the full model, Dechow et al. (2011) found that variables ch_emp and rett no longer loaded and variable btm, the book to market value of the company, loaded instead. Further tests showed that while the performance of the model is stable, the utility of different predictors change from year to year.

Dang et al. (2017) applied the baseline F-score model to detect material misstatements in financial statements of firm listed on HOSE in Vietnam and found that only the soft_assets variable was meaningful. By expanding the model with three variables—the returns on assets (ROA), the size of companies by revenue Size, and the financial leverage LV—they were able to develop the model below. However, the model was not tested using a held-out sample and therefore its generalisability is not certain.

$$\mathbf{MISSTATEMENT = -1.599 + 1.567\ Rsstacc + 2.219\ Chrec + 1.257\ Softassets - 11.148\ Roa + \varepsilon\ .}$$

(b) Financial statement irregularity detection models by data mining

One of the fundamental differences between the regression method mentioned above and the data mining method is that the regression method tests specific hypotheses that were formed before the data is known, while data mining is concerned with looking for unsuspected features or interesting patterns in the data (Hand et al., 2001). Data mining models commonly involve various classification algorithms such as Bayesian belief network, genetic algorithm, text mining, response surface methodology, artificial neural network (ANN), logistic regression (LR), group method of data handling, support vector machine (SVM), decision trees, and hybrid methods (West et al., 2014).

Among the research on financial statement irregularity detection by data mining, the research of Perols (2011) especially paid attention to the imbalanced nature of the event—financial statement irregularity is not common and failing to detect irregularity can lead to worse consequences than false accusation of irregularities. Perols (2011) tested six algorithms—ANN, SVM, decision trees, LR, bagging, and stacking—under different assumptions of class probabilities and misclassification costs and found that under more realistic assumptions, LR and SVM are the best performers.

5. Research methodology

This study experimented with the data mining classification method of financial statement irregularity detection (this method also handles LR models). With this, there were no pre-formed hypotheses. The study referenced common and reasonably evaluated extant financial statement irregularity detection models (discussed above) for potential predictors. Based on the findings of Perols (2011), the study chose LR and SVM as classifiers for the experiment.

5.1. Research framework

The study followed the framework of Dechow et al. (2011) and did not incorporate theories like the fraud triangle for it is not a research on behaviours or factors leading to irregularities but a research on the detection of irregularities. The research framework (**Figure 1**) comprises signs from the financial statement indicative of irregularities, including accruals quality, financial performance, non-financial performance, and market-related performance. In accounting, accruals are areas prone to manipulation. Unusually high accruals in accounts such as receivables or inventories can improve performance metrics such as profits or gross margin. Low or declining financial performance can incentivise such manipulation and therefore can be indicative of irregularities. Non-financial performance that does not follow common sense or industry benchmarks can also be an indication of irregularities. In addition, for publicly listed companies, the stock price is an important incentive for manipulation. Therefore, metrics that suggest a need for high stock price may also indicate irregularity in financial statements.

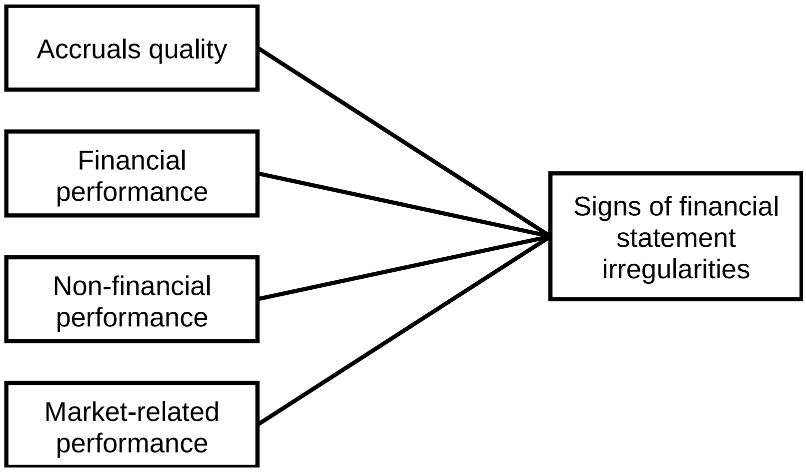

Figure 1. Research framework

Source: By the author

Table 1 presents the potential predictor variables (called "features" in data mining), referenced from prior irregularity detection models (see section 2.2.2). Features that were mathematically complex or unavailable in Vietnam were removed. Only ratio variables were included to minimise the effect of size differences between companies. A lot of non-financial variables as well as market related variables were removed due to a lack of reliable data sources in Vietnam. Although they could still be collected, manual collection was not cost effective and the variables were removed.

With infrequent event such as financial statement irregularities, classification models are prone to errors and misinterpretations as they can easily achieve a high accuracy by classifying all instances as the majority class. Therefore, it is crucial to take into account the occurring frequency of the events and other types of imbalance when developing predictive models. For financial statement irregularities or frauds, in addition to the low probability of occurrence, the cost of failing to detect or prevent is also higher than that of false accusation—especially for frauds. From the perspective of investors, undetected frauds can seriously damage their trust and cause abrupt changes in the market when they are uncovered. From the perspective of auditors, failing to detect frauds can mean litigation costs and a loss of reputation. Therefore, it is highly important to emphasise the detection of irregularities and the minimisation of false negative classifications. In order to do so, we can make use of the expected/prior probabilities of the event and the adjustment of error costs (McCue, 2006). However, as LR and SVM are not algorithms where the relative misclassification costs can be manipulated directly, the imbalances would be controlled through data sampling (see section 0). Following this, the estimated relative error costs of misclassification (ERC) was chosen as the main performance measurement as it accounts for class imbalances (Perols, 2011; West and Bhattacharya, 2016).

Table 1. Summary of research variables

Variable	Calculation[2]	Code
	Accruals quality	
Change in inventories	Δ inventories / [(TA + TA_{t-1})/2]	ch_inv
Change in receivables	Δ accounts receivable / [(TA + TA_{t-1})/2]	ch_rec
Inventories to total assets	inventories / TA	inv_ta
Richardson, Sloan, Soliman and Tuna (RSST) accruals	(Δ WC + Δ NCO + Δ FIN) / [(TA + TA_{t-1})/2], where WC = (current assets – cash and short-term investments) – (current liabilities – debt in current liabilities); NCO = (TA – current assets – investments and advances) – (total liabilities – current liabilities – long-term debt); FIN = (short-term investments + long-term investments) – (long-term debt + debt in current liabilities + preferred stock)	rsst_acc
Soft assets to total assets	(TA – net tangible fixed assets – cash and cash equivalents) / TA	soft_ta
	Financial performance	
Change in cash sales	Δ CS/CS_{t-1}, where CS = net sales – Δ accounts receivable	ch_cs
Change in ROA	PAT / [(TA + TA_{t-1})/2] – PAT_{t-1} / [(TA_{t-1} + TA_{t-2})/2]	ch_roa
ROA	PAT / TA	roa
Sales to total assets	net sales / TA	sale_ta
	Non-financial performance	
Audit opinion	Indicator variable = 1 if audit opinion in year t–1 is a qualified opinion; = 0 otherwise	aud_op
Auditor turnover	Indicator variable = 1 if the auditor changed in the last two years; = 0 otherwise	aud_to
Big 4 auditor	Indicator variable = 1 if auditor in year t – 1 is not a Big 4 firm; = 0 otherwise	big4
Record of past irregularities	The number of instances of financial statement irregularity in the last three years	restate
	Market-related performance	
Book-to-market	equity / market value of equity	btm

[2] If time is not specified, the variable is from year t; year t is 2020. TA = total assets; PAT = profits after tax.

Variable	Calculation[2]	Code
Issuance of securities	Indicator variable = 1 if firm issued securities (stocks, debts, etc.) in the period; = 0 otherwise	issue
Leverage	(short-term borrowings + long-term borrowings) / TA	lev

Source: By the author

5.2. Research process

This section describes in detail the research process from data handling to evaluation results. The collected data was first split into training and test data (80% and 20%). This ensured the objectivity of the models and prevented overfitting. The following tasks were then performed on the train data: data imputation, standardisation, feature selection, and sampling method selection. Through imputation and standardisation, the train data was pre-processed (the test data was pre-processed similarly but separately, right before the final evaluation). The preferred features and sampling method were then selected and used to train each of the two classifiers LR and SVM[3]. Finally, the trained classifiers were evaluated using the prediction results on the pre-processed test data (no sampling involved). To make the most use of the information available in the train data, repeated 10-fold cross validation was employed throughout the process as needed.

Data imputation and standardisation

The features were checked for any missing values. In handling missing values, incomplete records could be dropped completely or the values could be substituted through a process called imputation. As simply dropping incomplete records might cause a loss of valuable information, in this study, the imputation of missing continuous values was performed. Basic imputation methods like using the means or most frequent values are easier to apply but they do not factor in the correlations between features and may be biased. The k-nearest neighbours method uses the similarities between data points to predict values and therefore handles these problems better and was employed.

This method automatically standardised the data, transforming the numerical features to have a mean of 0 and a standard deviation of 1. This could enhance the prediction results. For instance, SVMs are strongly influenced by the scale of features as they work by measuring the distances between data points to determine similarity. Features with higher volume are likely to have more weight and cause bias in SVMs. Scaling the data beforehand ensured that all features contribute equally to the model.

Feature selection and sampling method selection

Reducing the number of features prevents overfitting and improves the generalisability of predictive models. Feature selection can be done manually by fitting models, eliminating less significant variables, and refitting—a process known as backward feature elimination. In this study, recursive feature elimination (RFE), essentially a backward elimination algorithm, was employed. The sampling method selection were performed independently of the feature selection but by a similar approach—sample the train data, train and test the model, resample, and repeat.

[3] Penalised LR and radial basis function kernel SVM from the R library *caret*.

The sampling method with the best average result would be chosen for training the classifiers. The tested sampling methods included under-sampling of majority class, over-sampling of minority class, and the hybrid method random over-sampling examples (ROSE).

5.3. Research data

Due a lack of data on frauds, for prior research in Vietnam, the determination of irregularity was based on whether there were material income restatements (Dang et al., 2017; Tran et al., 2015). Income restatement is calculated as the change in profit after tax (PAT) before and after an audit:

$$\textbf{Change in PAT} = \left|\frac{\textbf{PAT after audit} - \textbf{PAT before audit}}{\textbf{PAT before audit}}\right|.$$

On 16 November 2020, the Ministry of Finance of Vietnam (MOF) issued Circular 96/2020/TT-BTC, requiring publicly traded enterprises to justify change in PAT of more 5% before and after audit and profit-loss reversal (MOF, 2020). Based on this, 5% was selected as the threshold between non-material and material restatements. Financial statements where PAT changes from positive to negative (profit to loss) or alternatively, loss to profit, were also considered irregular.

The necessary data were collected through Vietstock financial data service. The data was collected for the period 2017–2020 but the mainly examined financial year was 2020 as a number of predictor variables required data from three years prior. The data was collected for publicly traded companies that were not financial service entities (banks, insurance or securities firms are subject to different reporting requirements). Vietstock offers data available on these stock exchange platforms: HOSE, HNX, UPCoM, over-the-counter (OTC), and others. However, the data on OTC or other minor stock exchanges was severely missing and therefore removed entirely. Firms with missing data on either audited or unaudited PAT were also removed. As some predictor variables—e.g. records of past irregularities—required data from three years prior, only firms that have been traded and have data on the platforms since 2017 or earlier could be used. The final sample comprises 790 publicly traded firms, of which 206 were found to have material income restatements in 2020. The probability of financial statement irregularities was 26.08%. See

Table 2 for a summary of the sample selection.

Table 2. Sample selection

Total sample	
Publicly traded firms with 2020 financial data on Vietstock	3,161
Less: Banks, insurance or securities firms	(330)
Less: Firms whose stocks were not listed on HNX, HOSE or UPCoM	(1,239)
Less: Firms missing data on profits after tax in 2020	(608)
Less: Firms that were active since 2018 or later	(194)
Usable observations/Total sample size	790
Classes in the sample	
Firms with irregularities	206
Firms without irregularities	584
Probability of irregular observations	26.08%

Source: By the author

In prior research, it was common to sample irregular observations first and then collect matching non-irregular samples to control for unmeasured variables and enhance internal validity. Matching samples by size or industry was employed because these studies were aiming to explain the factors that may lead to irregularities, errors or frauds. As this study has a different aim, it did not sample by matching and instead collected all available data. The data provided by Vietstock has one major flaw that it does not differentiate between zero values and missing values. Therefore judgment was used in determining whether a value is missing.

6. Research results and discussion

6.1. Research results

Feature selection and sampling method selection results

In total, ten features were selected for LR and five features were selected for SVM. **Table 3** summarises the most meaningful selected for each classifier.

Table 3. Feature selection result

Classifier	Preferred features
Logistic regression	restate, roa, soft_ta, ch_roa, aud_to, btm, sale_ta, ch_rec, big4, issue
Support vector machine	roa, restate, btm, soft_ta, sale_ta

Source: By the author

In selecting sampling method for LR and SVM models, ROSE consistently returned the worst results. Down-sampling performed slightly better than up-sampling, in either average ERC or result consistency, and therefore was chosen as the sampling method for both LR and SVM classifiers.

Model training result

$$\mathbf{irregularity = -0.707 + 0.767\ restate - 0.297\ roa + 0.336\ soft\_ta + 0.183\ ch\_roa + 0.468\ aud\_to + 0.205\ btm - 0.296\ sale\_ta - 0.170\ ch\_rec + 0.296\ big4 - 0.494\ issue + \varepsilon}$$

Table 4 presents the model training results after 50 resamples with 10-fold cross validation. On average, SVM models performed better than LR models. However, the results of SVM models fluctuated on a wider range. The final LR model included predictors from all elements of the research framework and had a null deviance of 457.48 on 329 degrees of freedom with a residual deviance of 371.57 on 319 degrees of freedom:

$$\mathbf{irregularity = -0.707 + 0.767\ restate - 0.297\ roa + 0.336\ soft\_ta + 0.183\ ch\_roa + 0.468\ aud\_to + 0.205\ btm - 0.296\ sale\_ta - 0.170\ ch\_rec + 0.296\ big4 - 0.494\ issue + \varepsilon}$$

Table 4. Model training result

	ERC[4]			
Classifier	**Mean**	**Min.**	**Median**	**Max.**
Logistic regression	0.47	0.33	0.45	0.62
Support vector machine	0.45	0.29	0.44	0.77

Source: By the author

Model evaluation result

The held-out test data had 157 observations, of which 41 were financial statements with irregularity (the irregularity probability was 26%, equivalent to the probability in the training data and the whole raw data). **Table 5** presents the final evaluation results. With this held-out sample, the LR model had out-performed SVM considerably. This could be expected because the SVM training results had fluctuated on quite a wide range. Nevertheless, the evaluation results did not deviate far from the training results.

Table 5. Model evaluation result

Classifier	Accuracy	Sensitivity	Specificity	ERC
Logistic regression	0.74	0.71	0.75	0.41

[4] Lower ERCs are more desirable.

Support vector machine	0.66	0.68	0.66	0.50

Source: By the author

6.2. Discussion of research results

The research results showed that the data mining classification method, in particular by LR and SVM, could be applied to detect financial statement irregularities in Vietnam. Though SVM had the potential to out-perform LR, the results of LR models were more consistent and therefore had better generalisability. In addition, with class imbalances, down-sampling by randomly removing observations from the majority class was found to be the most optimal inner sampling method. LR was used as a classifier in this research. However, inferences could still be made about the relations between the predicted and predictor variables.

Table 6 presents the expected versus actual directions of relations between the predicted and predictor variables.

The accruals quality variables include change in receivables and percentage of soft assets in total assets. The change in receivables was expected to be directly proportional to the probability of irregularity, as an unusually large change in receivables could mean that the receivable accounts are being used to artificially inflate revenues. However, the result was that change in receivables was in inverse proportion to the predicted variable. This suggests that there were other factors in play, revenues are not being inflated through receivables, and a company may be more likely to have irregularity in financial statements if receivables do not move in line with their financial performance. The percentage of soft assets—assets which are neither cash or property, plant, and equipment—in total assets was in direct proportion with the probability of irregularity as expected. When there are more soft assets, there may be more discretion to manipulate short-term earnings (Dechow et al., 2011).

The financial performance variables include change in ROA, ROA, and the sales to total assets ratio. With the exception of change in ROA, the remaining variables are in inverse proportion to the predicted variable as expected. The direct relation between change in ROA and the probability of irregularity is actually in accordance with the initial hypothesis of Dechow et al. (2011) that management would be inclined to show positive growth in earnings. The inverse relations of ROA and sales to total assets with the irregularity probability, on the other hand, suggest that declining financial performance had pressured management to manipulate earnings.

Table 6. Expected and observed directions of relations between predicted and predictor variables

Predictor/Feature	Expected relation	Observed relation[5]
Accruals quality		
Change in inventories	+	0
Change in receivables	+	–
Inventories to total assets	+	0
RSST accruals	+	0
Soft assets to total assets	+	+
Financial performance		
Change in cash sales	+	0
Change in ROA	–	+
ROA	–	–
Sales to total assets	–	–
Non-financial performance		
Audit opinion	+	0
Auditor turnover	+	+
Big 4 auditor	+	+
Record of past irregularities	+	+
Market-related performance		
Book-to-market	–	+
Issuance of securities	+	–
Leverage	+	0

Source: By the author

The non-financial performance variables include auditor turnover, auditor quality, and record of past irregularities. All variables are in direct proportion to the probability of irregularity as expected. A change in audit firm may raise doubts about the integrity of the company management. On the other hand, it can improve the objectivity of the auditors toward the company. However, according to the research results, companies that had changed audit firm were more likely to have financial statement irregularity. In addition, a lower auditing quality (non-Big 4 audit firms) and a larger record of past irregularities also correlated with higher likelihood of financial statement irregularity.

The market-related performance variables include book-to-market value and actual issuance of securities. Their relation to the predicted variable were both not of the expected direction. The

[5] Predictors/features not included in the final LR models are denoted with "0" observed relation.

inverse proportion of book-to-market to predicted variable was based on the presumption that the executives of a company whose market value is higher than book value would be more preoccupied with earning management. However, the results suggest that a low market value compared to book value had pressured management to manipulate earnings, presumably also due to incentives tied to stock prices. The direct proportion of issuance of securities to irregularity probability was based on the presumption that companies with financing needs would be more inclined to appear better to obtain the financing. Results showed an inverse relation between the issuance of securities and the probability of irregularity, suggesting that companies with financial statement irregularity were not able to obtain financing and were manipulating earnings to better their chances. Nevertheless, these are simple inferences and deviations from the expected should be further studied.

7. Conclusion

This study explored the applicability of classification models in detecting financial irregularities in Vietnam. Two classifiers—logistic regression and support vector machine—were employed to detect financial statements containing irregularities from publicly traded enterprises on HOSE, HNX, and UPCoM in 2020. Based on the results of the main experiments, it can be concluded that both classifiers are applicable for the detection of irregularities in Vietnam while taking into account the imbalanced nature of the event. The detection models performed best when there were five to ten features/predictors in a model. Some of the features found to have provided the most utility are the record of past irregularities, return on assets, book to market value, and the percentage of soft assets in total assets of the firm.

While the classification models returned results which are not as positive as that of prior research in Vietnam, this study differs from these literature in that, instead of explaining the indicators of fraud, it focused on real-world applicability by accounting for class and misclassification cost imbalances in both training and testing. However, generalisability still remains an issue as the data used in this study is limited to one financial year and to publicly traded enterprises only. The models will not be able to perform as well for other firm years or for smaller enterprises. Nevertheless, as public and private enterprises have many differing characteristics, it may be more preferable to have different models based on firm types.

In addition, the performance of classification models is limited by the quality of public financial data in Vietnam as well as the quantity of data. The auditors, those who have access to private transaction data, would be able to build more informative models and also employ additional detection methods. The utility of the models is also limited due to how broad and unfocused the object of detection is. A change in PAT of 5% before and after audit may have been due to either fraud or error, events whose nature are vastly different. The determination of irregularity was also entirely based on the opinions of the auditors, which could be highly subjective. In general, information on fraud cases in Vietnam is either not available to the public or poorly reported and in general not reliable. Improving the quality of financial information in Vietnam—e.g. more transparent fraud investigation and reporting, application of eXtensible Business Reporting Language, etc.—would facilitate the development of more reliable and specialised detection models that make use of both structured and unstructured data.

Despite its limitations, the study has addressed the gaps in irregularity detection research in Vietnam by experimenting on data of firms on multiple stock exchanges (HOSE, HNX, and

UPCoM) and accounting for class and cost imbalances. The study is expected to contribute to and encourage studies on the detection of financial statement irregularities, frauds, and irregular events in general in Vietnam, especially studies that involve computational intelligence or artificial intelligence. The study has various aspects that can be expanded on, such as classifiers (neural networks, ensemble methods, etc.), sampling methods, etc. for strictly technical subjects. For theory wise aspects, open areas include finding and explaining meaningful indicators of fraud, determining the causes of frauds, using time series data or event study to identify frauds, etc.

References

ACFE (2020), "Report to the Nations on Occupational Fraud and Abuse", *Global Fraud Study*, Available at: https://www.acfe.com/report-to-the-nations/2020 (Accessed: 02 May 2021).

Bierstaker, J.L., Brody, R.G. and Pacini, C. (2006), "Accountants' perceptions regarding fraud detection and prevention methods", *Managerial Auditing Journal*, Vol. 21 No.5, pp. 520 – 535.

Dang, N.H., Hoang, T.V.H & Dang, T.B. (2017), "Application of F-score in predicting fraud, errors: Experimental research in Vietnam", *International Journal of Accounting and Financial Reporting*, Vol. 7 No. 2, pp. 303 – 322.

Dechow, P.M., Ge, W., Larson, C.R. & Sloan, R.G. (2011), "Predicting Material Accounting Misstatements", *Contemporary Accounting Research*, Vol. 28, pp. 17 - 82,

Dorminey, J., Fleming, A.S., Kranacher, M.J. & Riley Jr, R.A. (2012), "The evolution of fraud theory", *Issues in Accounting Education*, Vol. 27 No. 2, pp. 555 - 579,

Hand, D.J., Mannila, H. & Smyth, P. (2001), *Principles of Data Mining*, MIT Press Books.

Hennes, K.M., Leone, A.J. & Miller, B.P. (2008), "The importance of distinguishing errors from irregularities in restatement research: The case of restatements and CEO/CFO turnover", *The Accounting Review*, Vol. 83 No.6, pp. 1487 - 1519.

ICAEW (2021), "How to report on irregularities, including fraud, in the auditor's report – a guide for auditors", Available at: https://www.icaew.com/technical/audit-and-assurance/audit/reporting-and-completion/how-to-report-on-irregularities (Accessed: 12 May 2021).

IFAC (2009), *International Standard On Auditing (ISA 450)*, Available at: https://www.ifac.org/sites/default/files/downloads/a021-2010-iaasb-handbook-isa-450.pdf (Accessed: 12 May 2021).

Kim, H.J., Mannino, M. & Nieschwietz, R.J. (2009), "Information technology acceptance in the internal audit profession: Impact of technology features and complexity", *International Journal of Accounting Information Systems*, Vol. 10 No. 4, pp. 214 – 228.

McCue, C. (2006), *Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis*, Elsevier.

MOF (2020), *Thông tư Hướng dẫn công bố thông tin trên thị trường chứng khoán*, Available at: http://vbpl.vn/tw/Pages/vbpq-van-ban-goc.aspx?ItemID=146048 (Accessed: 29 June 2021).

Nguyen, T.H., Huynh, V.S. & Nguyen, T.D. (2018), "Fraud of Financial Statements at Listed Enterprises on Ho Chi Minh City Securities Department", *VNU Journal of Science: Economics and Business*, Vol. 34 No. 4.

Nigrini, M.J. (2012), *Benford's Law: Applications for forensic accounting, auditing, and fraud detection,* John Wiley & Sons.

Perols, J. (2011), "Financial statement fraud detection: An analysis of statistical and machine learning algorithms", *Auditing: A Journal of Practice & Theory,* Vol. 30 No. 2, pp. 19 – 50.

PwC (2018), "Pulling fraud out of the shadows – Global Economic Crime and Fraud Survey 2018: Vietnam Perspectives", Available at: https://www.pwc.com/vn/en/publications/vietnam-publications/economic-crime-fraud-survey-2018.html (Accessed: 29 June 2021).

Tran, T.G.T., Nguyen, T.T., Dinh, N.T., Hoang, T.H. & Nguyen, D.H.U. (2015), "Đánh giá rủi ro gian lận báo cáo tài chính của các công ty niêm yết tại Việt Nam", *Tạp chí Phát triển kinh tế*, Vol. 26 No. 1, pp. 74-94, Available at: http://jabes.ueh.edu.vn/Home/SearchArticle?article_Id=6a169fe7-595c-4398-9573-cc4930dc8dbf (Accessed: 09 May 2021).

VAFE (2020), "Báo cáo khảo sát về công bố thông tin trên thị trường chứng khoán năm 2020", Available at: http://vafe.org.vn/Bao-cao-khao-sat-ve-cong-bo-thong-tin-tren-thi-truong-chung-khoan-nam-2020-830-765799.htm (Accessed: 08 May 2021).

Vietstock (2021), "Bức tranh kiểm toán 2020: 'Muôn hình vạn trạng'", Available at: https://vietstock.vn/2021/04/buc-tranh-kiem-toan-2020-8216muon-hinh-van-trang8217-737-845830.htm (Accessed: 29 June 2021).

West, J. & Bhattacharya, M. (2016), "Intelligent financial fraud detection: a comprehensive review", *Computers & Security,* Vol. 57, pp. 47 - 66, https://doi.org/10.1016/j.cose.2015.09.005 (Accessed: 29 June 2021).

West, J., Bhattacharya, M. & Islam, R. (2014), "Intelligent financial fraud detection practices: an investigation", *International Conference on Security and Privacy in Communication Networks*, pp. 186 - 203, https://doi.org/10.1007/978-3-319-23802-9_16 (Accessed: 29 June 2021).

Zahra, S.A., Priem, R.L. and Rasheed, A.A. (2005), "The antecedents and consequences of top management fraud", *Journal of Management*, Vol. 31 No. 6, pp. 803 - 828, https://doi.org/10.1177%2F0149206305279598 (Accessed: 29 June 2021).