DỰ ĐOÁN GIAN LẬN BÁO CÁO TÀI CHÍNH TẠI VIỆT NAM BẰNG PHƯƠNG PHÁP HỌC MÁY

NCS. Vũ Văn Đức, Lê Minh Hòa, Nguyễn Thanh Mai, Đàm Khánh Linh, Lê Diễm Quỳnh
Trường Đại học Ngoại thương
Tác giả liên hệ: minhhoah2003@gmail.com

Ngày nhận: 14/4/2025
Ngày nhận bản sửa: 18/4/2025
Ngày duyệt đăng: 24/4/2025

Tóm tắt

Gian lận báo cáo tài chính gây thiệt hại tài chính và làm giảm niềm tin. Tại Việt Nam, nhiều vụ đã được phát hiện với tổn thất lớn; các phương pháp kiểm toán truyền thống khó phát hiện gian lận tinh vi. ML (Machine Learning) và AI (Artificial Intelligence) mở ra cơ hội mới. Nghiên cứu so sánh Logistic Regression, SVM, Random Forest và Neural Networks trên dữ liệu niêm yết Việt Nam 2019 - 2023, dựa trên Lý thuyết Tam giác gian lận (Cressey, 1953). Lựa chọn đặc trưng và xác thực chéo đảm bảo tính khách quan và khả năng khái quát. Kết quả sơ bộ cho thấy ML vượt trội phương pháp truyền thống trong phát hiện gian lận, cung cấp hiểu biết quan trọng cho cơ quan quản lý, nhà đầu tư và doanh nghiệp tăng cường giám sát tài chính và minh bạch thị trường.

Từ khóa: *Gian lận báo cáo tài chính, hiệu suất, học máy, so sánh, Việt Nam.*

Predicting Financial Statement Fraud in Vietnam Using Machine Learning Techniques

Vu Van Duc, Le Minh Hoa, Nguyen Thanh Mai, Dam Khanh Linh, Le Diem Quynh
Foreign Trade University
Corresponding Author: minhhoah2003@gmail.com

Abstract

Financial statement fraud results in substantial financial losses and erodes stakeholder trust. In Vietnam, several high-profile cases have been revealed, often with significant repercussions, while traditional auditing methods frequently fail to detect complex fraudulent schemes. Emerging technologies, such as Machine Learning (ML) and Artificial Intelligence (AI) present new avenues for fraud detection. This study evaluates the effectiveness of Logistic Regression, Support Vector Machine (SVM), Random Forest, and Neural Networks by analyzing data from Vietnamese listed companies from 2019 to 2023, grounded in the Fraud Triangle Theory (Cressey, 1953). We employ feature selection and k-fold cross-validation to enhance the objectivity and generalizability of our findings.. Preliminary results indicate that ML techniques outperform traditional methods in detecting financial fraud, offering valuable insights for regulators, investors, and firms seeking to improve financial oversight and market transparency.

Keywords: *Comparative analysis, Financial statement fraud, Machine learning, Model performance, Vietnam.*

1. Đặt vấn đề

Gian lận báo cáo tài chính (BCTC) là hành vi cố tình sai lệch hoặc che giấu sai phạm trong báo cáo tài chính, gây tổn thất lớn về tài chính và niềm tin của nhà đầu tư. Trên toàn cầu, gian lận báo cáo tài chính là hình thức ít phổ biến nhất (chiếm 5% số vụ), nhưng lại gây thiệt hại lớn nhất, với mức tổn thất trung vị lên tới 766.000 USD mỗi vụ (Association of Certified Fraud Examiners, 2024). Tại Việt Nam, những vụ bê bối như Trịnh Xuân Thanh (PVC), MobiFone - AVG, Vinalines và Thuduc House đã làm xói mòn niềm tin của nhà đầu tư và yêu cầu các công cụ phát hiện gian lận hiệu quả hơn. Phương pháp kiểm toán truyền thống với tần suất kiểm tra định kỳ và trọng tâm tuân thủ chuẩn mực kế toán thường tỏ ra kém hiệu quả trước các thủ đoạn gian lận tinh vi. Các vụ bê bối lớn như Enron và WorldCom cho thấy gian lận có thể kéo dài nhiều năm mà không bị phát hiện, gây hậu quả nghiêm trọng (DeFond & Francis, 2005; Perols, 2011). Với sự bùng nổ của khoa học dữ liệu, đặc biệt là học máy (ML) và trí tuệ nhân tạo (AI), các mô hình này hứa hẹn khả năng phân tích khối lượng dữ liệu lớn và nhận

diện các mẫu gian lận phức tạp mà phương pháp truyền thống không thể chạm tới. Nhiều nghiên cứu cho thấy ML có thể phát hiện các bất thường tiềm ẩn trong dữ liệu tài chính và phi tài chính, từ đó, nâng cao độ chính xác và độ nhạy (Adhikari & Bansal, 2018; Wang & cộng sự, 2016; Zhang & cộng sự, 2019). Tuy nhiên, nhược điểm của các mô hình ML là tính khó giải thích (hộp đen) và thiếu so sánh giữa các phương pháp trong bối cảnh Việt Nam (Rajula & cộng sự, 2020; Churpek & cộng sự, 2016).

Nghiên cứu này sẽ so sánh hiệu suất của các thuật toán học máy như Support Vector Machine (SVM), Random Forest (RF) và Neural Networks (NN) với phương pháp lượng truyền thống (Logistic Regression - LR) để tìm ra mô hình phù hợp nhất cho các doanh nghiệp niêm yết trên HOSE và HNX trong giai đoạn 2019 - 2023. Các biến tài chính và phi tài chính sẽ được tiền xử lý và lựa chọn đặc trưng qua phân tích tương quan, sau đó, được huấn luyện và kiểm định qua xác thực chéo K-fold. Các chỉ số như Accuracy, Precision, Recall, F1-Score và AUC-ROC sẽ được sử dụng để đánh giá hiệu suất của các mô hình. Đồng thời, kỹ thuật xử lý mất cân bằng dữ liệu như SMOTE sẽ được áp dụng để cải thiện khả năng phát hiện gian lận hiếm gặp.

2. Cơ sở lý thuyết

2.1. Gian lận Báo cáo tài chính

Theo Hiệp hội các nhà Kiểm tra Gian lận (Association of Certified Fraud Examiners - ACFE, 2016), gian lận là những hành vi vi phạm pháp luật được thực hiện một cách cố ý nhằm những mục đích nhất định như thao túng hoặc cung cấp thông tin sai lệch cho các bên liên quan. Dạng gian lận này thường được thực hiện thông qua việc thao túng báo cáo tài chính nhằm che giấu tình hình thực tế của doanh nghiệp, từ đó, trục lợi từ các bên liên quan.

Tại Việt Nam, theo Chuẩn mực Kiểm toán Việt Nam số 240, gian lận BCTC được định nghĩa là hành vi cố ý của một hoặc nhiều cá nhân trong Ban Quản trị, Ban Giám đốc, nhân viên hoặc bên thứ ba nhằm trình bày sai lệch báo cáo tài chính một cách trọng yếu thông qua các hành vi gian dối với mục đích thu lợi bất chính hoặc bất hợp pháp.

2.2. Lý thuyết Tam giác gian lận

Năm 1950, Donald Cressey bắt đầu nghiên cứu về hành vi gian lận và nhận thấy rằng các hành vi này thường có ba đặc điểm chung. *Thứ nhất,* người thực hiện hành vi gian lận có cơ hội để tiến hành gian lận. *Thứ hai,* cá nhân này đối mặt với vấn đề về tài chính không thể chia sẻ (áp lực). *Thứ ba,* người gian lận hợp lý hóa hành động của mình, tin rằng nó phù hợp với chuẩn mực đạo đức cá nhân. Năm 1953, Cressey tổng hợp ba yếu tố: áp lực, cơ hội và sự hợp lý hóa và biểu diễn dưới dạng mô hình Tam giác gian lận.

3. Phương pháp nghiên cứu

3.1. Đề xuất mô hình nghiên cứu

Nghiên cứu này đề xuất mô hình dự đoán xác suất gian lận theo công thức:

P(Gian lận) = f(Áp lực, Cơ hội, Sự hợp lý hóa)

- Biến phụ thuộc FIF_1: Khoản chênh lệch lợi nhuận:

Biến FIF_1 có giá trị là 1, tức doanh nghiệp có khả năng gian lận nếu mức chênh lệch này từ 5% trở lên, ngược lại, FIF_1 có giá trị là 0.

Chênh lệch lợi nhuận = |(Lợi nhuận ròng sau kiểm toán - Lợi nhuận ròng trước kiểm toán)/Lợi nhuận ròng trước kiểm toán|

- Biến phụ thuộc FIF_2: F-Score

Theo Situngkir & Triyanto (2020), các công ty có giá trị F-Score lớn hơn 1 đồng nghĩa với việc có khả năng gian lận trong báo cáo tài chính. Ngược lại, nếu giá trị F-Score nhỏ hơn 1, công ty không có nguy cơ gian lận trong báo cáo tài chính.

F-Score = accrual quality (chất lượng dồn tích) + financial performance (hiệu suất tài chính)

- Biến phụ thuộc FIF_3: M-score

Beneish (1999) đề xuất mức ngưỡng tới hạn (cutoff) phù hợp là 3,75%, hay nói cách khác, công ty có M-Score cao hơn -1,78 được coi là có khả năng gian lận tài chính.

M-Score được tính theo công thức:

$$M\text{-}Score = -4{,}84 + 0{,}92 \times DSRI + 0{,}528 \times GMI + 0{,}404 \times AQI + 0{,}892 \times SGI + 0{,}115 \times DEPI - 0{,}172 \times SGAI + 4{,}679 \times TATA - 0{,}327 \times LVGI$$

- Biến độc lập

Dựa trên lý thuyết Tam giác gian lận, mô hình bao gồm 3 nhóm chính: Áp lực, Cơ hội và Sự hợp lý hóa. Các nhân tố của 3 nhóm chính được tính toán dựa theo nghiên cứu của Đặng Ngọc Hùng & cộng sự (2022) và nghiên cứu của Skousen & cộng sự (2009).

3.2. Phương pháp nghiên cứu

Về phương pháp xây dựng mô hình, nghiên cứu sử dụng các thuật toán học máy có giám sát để dự báo gian lận: Hồi quy Logistic, SVM (Support Vector Machine), Random Forest và Mạng nơ-ron (Neural Network).

Về đánh giá hiệu suất mô hình, nghiên cứu sử dụng các tiêu chí sau: Accuracy, Precision, Recall, F1-Score và AUC-ROC. Ngoài sử dụng các chỉ số trên, nghiên cứu còn sử dụng phương pháp xác thực chéo K

lần (K-fold cross validation) để kiểm tra hiệu suất các mô hình được ứng dụng trong bài nghiên cứu. Đối với phương pháp này, tập dữ liệu được chia thành 10 phần (10-fold) để tiến hành huấn luyện (Training) và kiểm tra mô hình (Testing) theo phương pháp xác thực chéo (10-fold cross validation). Trong mỗi lần lặp, 1 phần dữ liệu được sử dụng để kiểm tra, trong khi 9 phần còn lại dùng để huấn luyện mô hình.

4. Kết quả và thảo luận

4.1. Kết quả nghiên cứu

Kết quả cho thấy các mô hình học máy (SVM, Random Forest, Neural Network) đều vượt trội hơn phương pháp truyền thống. Trong đó, Random Forest cho hiệu suất cao nhất ở tất cả tiêu chí, nổi bật là Accuracy 92,62% và AUC-ROC 97,9% trong Mô hình 2, cho thấy khả năng phát hiện gian lận chính xác và hạn chế dương tính giả (Hossain & cộng sự, 2024). Neural Network cũng thể hiện khả năng dự báo mạnh, theo sau là SVM và Logistic Regression. Logistic Regression có hiệu suất thấp nhất, phản ánh hạn chế khi xử lý dữ liệu phức tạp.

Bảng 1. Kết quả đánh giá hiệu suất mô hình

		Accuracy	Precision	Recall	F1-Score	AUC-ROC
Mô hình 1	LR	73,07%	72,2%	75,24%	73,64%	80,08%
	SVM	77,72%	77,75%	77,85%	77,75%	85,76%
	RF	86,74%	86,52%	87%	86,75%	94,13%
	NN	79,42%	78,48%	81,26%	79,81%	86,29%
Mô hình 2	LR	77,54%	76,55%	79,46%	77,94%	85,22%
	SVM	82,73%	81,52%	84,66%	83,04%	90,51%
	RF	92,62%	92,4%	92,9%	92,63%	97,9%
	NN	86,64%	85,07%	88,92%	86,94%	92,34%
Mô hình 3	LR	63,15%	61,71%	69,61%	65,37%	66,89%
	SVM	69,5%	67,99%	73,81%	70,72%	76,31%
	RF	84,53%	87,99%	79,96%	83,76%	90,59%
		73,89%	72,2%	77,65%	74,81%	79,31%

Nguồn: Các tác giả tổng hợp từ phần mềm Python

4.2. Thảo luận

Thứ nhất, Logistic Regression giúp xác định ảnh hưởng từng biến nhưng hạn chế khi xử lý quan hệ phi tuyến. Ngược lại, mô hình học máy, đặc biệt là học sâu, dù khó diễn giải nhưng hiệu quả hơn trong dự báo gian lận nhờ khai thác quan hệ ẩn (Hossain & cộng sự, 2024).

Thứ hai, Random Forest vượt trội hơn các mô hình khác như Logistic Regression, SVM hay Neural Networks nhờ khả năng tổng hợp và khái quát hóa tốt (Hamal & Senvar, 2021; Hajek & Henriques, 2017). Recall cao giúp hạn chế bỏ sót gian lận (Bao & cộng sự, 2020).

Thứ ba, mỗi phương pháp đo lường gian lận đều có ưu nhược điểm riêng. Chênh lệch lợi nhuận trước - sau kiểm toán dễ gây hiểu nhầm. F-Score có độ chính xác cao nhưng dễ bị ảnh hưởng bởi quản lý lợi nhuận (Dechow & cộng sự, 2009), còn M-Score đơn giản nhưng kém hiệu quả trong bối cảnh Việt Nam (Nam & cộng sự, 2023).

5. Kết luận và khuyến nghị giải pháp

Trong bối cảnh gian lận BCTC ngày càng tinh vi và các phương pháp kiểm toán truyền thống dần bộc lộ hạn chế, dưới đây là một số hàm ý chính sách cho các bên liên quan nhằm nâng cao hiệu quả phát hiện và phòng ngừa gian lận tại các doanh nghiệp niêm yết.

Đối với các cơ quan quản lý nhà nước như Bộ Tài chính, Kiểm toán Nhà nước hay Ủy ban Chứng khoán Nhà nước, cần tăng cường áp dụng công nghệ học máy vào hệ thống cảnh báo sớm và phát hiện bất thường trong BCTC của doanh nghiệp. F-Score là chỉ báo hiệu quả đã được chứng minh trong

mô hình, có thể tích hợp vào bộ chỉ báo giám sát định lượng bắt buộc hoặc khuyến nghị cho doanh nghiệp niêm yết.

Đối với các công ty kiểm toán và tổ chức tài chính (bao gồm ngân hàng, công ty chứng khoán, quỹ tín dụng, quỹ đầu tư,...), để giảm phân loại nhầm một doanh nghiệp không gian lận thành doanh nghiệp gian lận (sai sót loại I) các đơn vị ưu tiên áp dụng Mô hình 2: F-Score và Mô hình 3: M-Score với thuật toán Random Forest, vốn cho kết quả Precision cao (92,4% và 87,99%). Để hạn chế bỏ sót các doanh nghiệp gian lận (sai sót loại II), các đơn vị nên chú trọng đến các mô hình có chỉ số Recall cao như Mô hình 1 với thuật toán Random Forest (87%) và Mô hình 2 khi kết hợp Random Forest hoặc Neural Networks (92,9% và 88,92%). Tuy nhiên, hai thuật toán trên đều gặp khó khăn về khả năng diễn giải, các cơ quan có thể khắc phục bằng cách sử dụng điểm số về mức độ quan trọng của đặc trưng (Feature Importance Scores) hoặc áp dụng mô hình lai (hybrid models) kết hợp với các thuật toán diễn giải tốt như Logistic Regression để tối ưu hóa hiệu quả dự báo và tính minh bạch trong diễn giải kết quả. Khi các đơn vị cần sự cân bằng giữa precision và recall, chỉ số F1-Score được xem là tiêu chí phù hợp để lựa chọn mô hình. Trong nghiên cứu, các mô hình dùng Random Forest đều đạt F1-Score cao và vượt trội so với các mô hình còn lại. Ngoài ra, để giảm thiểu rủi ro tín dụng, các tổ chức tài chính nên ưu tiên mô hình có khả năng phân loại tổng thể tốt, thể hiện qua các chỉ số như Accuracy và AUC-ROC. Kết quả nghiên cứu cho thấy Mô hình 2 (Random Forest) đạt Accuracy 92,62% và AUC-ROC 97,90%, trong khi Mô hình 1 cũng có kết quả tích cực với Accuracy 86,74% và AUC-ROC 94,13%.

Tài liệu tham khảo

Adhikari, S., & Bansal, A. (2018). An empirical study of financial fraud detection using machine learning techniques. *Journal of Financial Crime*, 25(4), 1123-1140. DOI: 10.1108/JFC-06-2017-0052

Association of Certified Fraud Examiners. (2016). *Report to the Nations on Occupational Fraud and Abuse*. Truy cập ngày 23/04/2025, từ https://www.acfe.com/fraud-resources/report-to-the-nations-archive.

Association of Certified Fraud Examiners. (2024). *Occupational fraud 2024: A report to the nations (13th ed.)*. Truy cập ngày 23/04/2025, từ https://legacy.acfe.com/report-to-the-nations/2024/.

Bao, Y., Ke, B., Li, B., Yu, Y. J., & Zhang, J. (2020). Detecting accounting fraud in publicly traded US firms using a machine learning approach. *Journal of Accounting Research*, 58(1), 199-235. DOI: 10.1111/1475-679X.12292

Beneish, M. D. (1999). The Detection of Earnings Manipulation. *Financial Analysts Journal,* 55(5), 24-36.

Churpek, M. M., Yuen, T. C., Winslow, C., Meltzer, D. O., Kattan, M. W., & Edelson, D. P. (2016). Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. *Critical Care Medicine*, 44(2), 368-374.

Cressey, D. R. (1953). *Other people's money.* Patterson Smith.

Dechow, P. M., Ge, W., & Schrand, C. (2009). Understanding earnings quality: A review of the proxies, their determinants and their consequences. Working paper, University of California, Berkeley, University of Washington, and University of Pennsylvania.

DeFond, M. L., & Francis, J. R. (2005). Audit research after Sarbanes-Oxley. Auditing: *A Journal of Practice & Theory,* 24(s-1), 5–30.

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. DOI: 10.1016/j.knosys.2017.05.001

Hamal, S., & Senvar, O. (2021). Comparing performances and effectiveness of machine learning classifiers in detecting financial accounting fraud for Turkish SMEs. *International Journal of Computational Intelligence Systems*, 14(1), 769-782.

Hung, D. N., Diep, P. T. H., & Nhien, C. T. (2022). Nghiên cứu gian lận báo cáo tài chính tiếp cận theo thuật toán rừng ngẫu nhiên. *Tạp chí Khoa học và Công nghệ - Đại học Công nghiệp Hà Nội*, 58(2), 155-161.

Hossain, M. Z., Raja, M. R., & Hasan, L. (2024). Developing predictive models for detecting financial statement fraud: A machine learning approach. *Applied Sciences*, 2(6), 271-290.

Nam, N. H. P., Thoa, T. T. K., Dung, T. P., Thao, N. P., & Hung, T. K. (2023). Nghiên cứu mô hình phát hiện gian lận báo cáo tài chính. *Tạp chí Công Thương,* (8), 430-434.

Perols, J. (2011). Financial statement fraud detection: An analysis of statistical and machine learning algorithms. Auditing: *A Journal of Practice & Theory,* 30(2), 19-50.

Rajula, H. S. R., Verlato, G., Manchia, M., Antonucci, N., & Fanos, V. (2020). Comparison of conventional statistical methods with machine learning in medicine: Diagnosis, drug development, and treatment. *Medicina*, 56(9), 455.

Situngkir, N. C., & Triyanto, D. N. (2020). Detecting fraudulent financial reporting using fraud score model and fraud pentagon theory: Empirical study of companies listed in the L.Q. 45 Index. *The Indonesian Journal of Accounting Research*, Vol. 23, No 3.

Skousen, C. J., Smith, K. R., & Wright, C. J. (2009). Detecting and predicting financial statement fraud: The effectiveness of the fraud triangle and SAS No. 99. *International Journal for Researcher Development,* 13. DOI: 10.1108/MRR-09-2015-0216

Wang, Y., Wang, J., & Wang, H. (2016). A novel approach for financial fraud detection based on machine learning techniques. *Expert Systems with Applications*, 66, 131-141. DOI: 10.1016/j.eswa.2016.09.037

Zhang, Y., Hu, J., & Liu, Y. (2019). Financial fraud detection based on a hybrid model of machine learning. *IEEE Access*, 7, 86686-86694. DOI: 10.1109/ACCESS.2019.2925298