

NGHIÊN CỨU GIAN LẬN BÁO CÁO TÀI CHÍNH TIẾP CẬN THEO THUẬT TOÁN RỪNG NGẪU NHIÊN

RESEARCH ON FRAUD FINANCIAL STATEMENTS APPROACH BY RANDOM FOREST ALGORITHM

Đặng Ngọc Hùng^{1,*},
Phạm Thị Hồng Diệp¹, Cao Thị Nhiên²

TÓM TẮT

Nghiên cứu sử dụng các thuật toán rừng ngẫu nhiên dựa trên lý thuyết tam giác gian lận, để xem xét ảnh hưởng của các yếu tố đến gian lận báo cáo tài chính của các doanh nghiệp ở Việt Nam, sử dụng bộ dữ liệu bảng với 2235 quan sát trong giai đoạn 2014 - 2020. Nghiên cứu xem xét các mức chênh lệch lợi nhuận trước và sau kiểm toán trên 5%. Kết quả dự báo theo thuật toán rừng ngẫu nhiên đạt ở mức 91%. Đồng thời nghiên cứu đã xác định yếu tố ảnh hưởng lớn nhất là sự ổn định tài chính của doanh nghiệp. Nghiên cứu này có giá trị khi ứng dụng trong việc đánh giá tình trạng gian lận báo cáo tài chính của doanh nghiệp khả năng dự báo với độ chính xác cao.

Từ khóa: Gian lận; báo cáo tài chính; học máy; tam giác gian lận; rừng ngẫu nhiên.

ABSTRACT

Research using random forest algorithms based on fraudulent triangle theory, to examine the influence of factors on fraudulent financial statements of enterprises in Vietnam, using a dataset Table data with 2235 observations for the period 2014 - 2020. The study considers the difference in profit before and after the audit over 5%. Forecast results according to random forest algorithm reach 91%. At the same time, the study has determined that the most influential factor is the financial stability of enterprises. This study is valuable when it comes to assessing the financial reporting fraud situation of enterprises with the ability to predict with high accuracy.

Keywords: Fraud; financial report; machine learning; fraud triangle; random forest.

¹Trường Đại học Công nghiệp Hà Nội

²Trường Cao đẳng Công nghệ Bách khoa Hà Nội

*Email: dangngochung@haui.edu.vn; hungdangngockt@yahoo.com.vn

Ngày nhận bài: 11/02/2022

Ngày nhận bài sửa sau phản biện: 11/3/2022

Ngày chấp nhận đăng: 25/4/2022

1. GIỚI THIỆU

Gian lận báo cáo tài chính (BCTC) trên thế giới ngày càng gia tăng và trở thành vấn đề nghiêm trọng đối với các doanh nghiệp, chính phủ, và các nhà đầu tư. Đặc biệt là tại các quốc gia có thị trường vốn gian lận BCTC đã đe dọa đến niềm tin của công chúng vào thông tin trên thị trường. Tại Hoa Kỳ, hằng năm thiệt hại do gian lận BCTC gây ra ước

tính vào khoảng 572 tỉ USD (ACFE, 2008). Các doanh nghiệp trung bình hằng năm bị thất thoát khoảng 5 - 6% doanh thu do gian lận BCTC. Trường hợp điển hình, một trong những vụ gian lận báo cáo lớn nhất lịch sử - gian lận BCTC ở công ty Enron năm 2000 đã gây thiệt hại 80 tỉ USD trên thị trường vốn hóa cho những cá nhân, tổ chức đầu tư vào công ty này. Tại Việt Nam, những gian lận BCTC gần đây của công ty gỗ Trường Thành, công ty thủy sản Hùng Vương, và một loạt các công ty niêm yết trên sàn chứng khoán đã bị phát hiện đã gây ra tâm lý nghi ngờ cho nhà đầu tư, ảnh hưởng tới hoạt động của thị trường vốn. Trong bối cảnh hiện nay, quá trình sử dụng thông tin đăng tải trên BCTC ngày càng thể hiện vai trò quan trọng trong quản lý và đầu tư. Với tâm lý hạn chế rủi ro, các nhà đầu tư muốn biết rõ về dòng tiền mà mình đầu tư sẽ được sử dụng có hiệu quả không. Ở Việt Nam, trong thời gian gần đây, hiện tượng chênh lệch kết quả giữa các BCTC trước và sau kiểm toán đã tạo ra những tâm lý nghi ngại. Nhân tố này có tác động tiêu cực đến việc ra quyết định của các nhà đầu tư. Điều này có thể dẫn đến việc hạn chế dòng tiền lưu thông không hiệu quả, gây ra nhiều hệ lụy cho cả nền kinh tế nếu không có những biện pháp cải thiện tình trạng này. Hiện nay, Bộ Tài chính cũng đã ban hành chuẩn mực kiểm toán số 240 (VSA 240) quy định về trách nhiệm của kiểm toán viên liên quan đến gian lận BCTC, yêu cầu kiểm toán viên phải đánh giá rủi ro sai sót trọng yếu trên BCTC dựa vào các yếu tố Động cơ/Áp lực, cơ hội và thái độ hoặc khả năng hợp lý hóa. Mặc dù vậy, các chuẩn mực này trong quá trình áp dụng vào thực tế còn gặp nhiều khó khăn. Các nhân tố này phụ thuộc lớn vào quyết định suy xét của kiểm toán viên. Đứng về góc độ của nhà đầu tư, chuẩn mực này chưa thực sự hữu ích trong quá trình ra quyết định. Thực tiễn này đòi hỏi một nghiên cứu thực tiễn đưa ra một mô hình có thể dễ dàng tiếp cận và áp dụng hơn đối với các phán đoán về mức độ gian lận BCTC. Câu hỏi đặt ra là những nhân tố nào có tầm quan trọng nhất trong dự báo khả năng gian lận BCTC? và việc ứng dụng các thuật toán học trong dự báo có khả thi không? Do đó mục tiêu nghiên cứu là vận dụng thuật toán rừng ngẫu nhiên vào việc xem xét các biến số tài chính và phi tài chính tiếp cận theo lý thuyết tam giác gian lận ảnh hưởng đến khả năng gian lận BCTC. Trong nhiều mô hình dự báo, nghiên cứu này chọn rừng ngẫu nhiên vì đây là một công nghệ tiên tiến, đang

được ứng dụng rộng rãi trên thế giới song còn mới mẻ tại Việt Nam.

2. TỔNG QUAN NGHIÊN CỨU

2.1. Nghiên cứu về sử dụng các thước đo tài chính

Các thước đo tài chính là những chỉ tiêu đánh giá có thể được trình bày bằng đơn vị tiền tệ hoặc là những hệ số tài chính được tính toán từ những chỉ tiêu tiền tệ. Những chỉ tiêu tài chính mà người sử dụng thông tin khá quan tâm có thể kể đến như tổng tài sản, doanh thu, lợi nhuận, khoản phải thu, dòng tiền.

M.D Beneish [10] đã tiến hành phân tích 363 mẫu được thu thập và thu được 49 doanh nghiệp vi phạm. Ngoài ra, ông còn thu thập thêm 15 doanh nghiệp có hệ thống kế toán bị nghi ngờ bởi thông tin truyền thông trong khoảng thời gian từ 1987 đến 1993, cùng với việc sử dụng mô hình điều chỉnh các khoản dồn tích của Jones [24] đã giúp cho M.D Beneish [10] đưa ra kết luận tốc độ tăng trưởng doanh số bán hàng, đòn cân nợ và tổng ước tính kế toán trên tổng tài sản là khả hữu ích trong việc phát hiện các doanh nghiệp vi phạm và cả cho những doanh nghiệp đang tích cực lợi dụng các khoản trích trước cho việc điều chỉnh thu nhập. M. D Beneish [11] đã cho thấy rằng những chỉ số tài chính có giá trị cao càng có khả năng lợi nhuận đang bị thổi phồng. Cụ thể là những tỷ số tài chính có liên quan đến tỷ lệ khoản phải thu trên doanh thu, lợi nhuận biên, chất lượng tài sản, tăng trưởng doanh thu và các khoản trích trước.

Persons [37] đã thực hiện phân tích hồi quy logit giữa hai nhóm mẫu nghiên cứu được thu thập gồm 103 doanh nghiệp có gian lận kết hợp với 103 doanh nghiệp không có gian lận. Persons đã chọn lựa 10 biến trong đó có 8 biến là các chỉ số tài chính được đưa vào để phát triển hai mô hình logistics gồm mô hình cho năm gian lận và năm trước năm gian lận nhằm tăng độ tin cậy cho kết quả dự báo. Tương tự như nhận định của Persons [37], nghiên cứu của [17] cũng cho rằng những doanh nghiệp nào có đòn bẩy tài chính càng cao, cùng với những khoản bảo đảm cho nợ vay thì động cơ thực hiện hành vi gian lận càng cao. Ngoài ra, khả năng tự tài trợ vốn của doanh nghiệp chi phối nhiều đến nhu cầu huy động vốn từ bên ngoài, một doanh nghiệp càng có khả năng tự tài trợ vốn một cách tự chủ thì khả năng thực hiện hành vi gian lận càng thấp.

Lou & Wang [32] đã dựa vào những nghiên cứu trước đây để tìm ra những biến có liên quan đến hành vi gian lận đại diện cho áp lực, cơ hội và thái độ và cùng với việc thực hiện phương pháp hồi quy logit giúp ông kết luận đòn bẩy tài chính có mối quan hệ có ý nghĩa với khả năng BCTC có gian lận. Một số nghiên cứu khác cũng có kết quả tương đồng các công ty niêm yết có đòn bẩy tài chính càng lớn thì khả năng sai sót BCTC của công ty càng cao [11, 16, 22, 44, 45, 47].

2.2. Nghiên cứu về sử dụng các thước đo phi tài chính

Thước đo phi tài chính là những thông tin không được trình bày bằng đơn vị tiền tệ hay được tính toán bằng đơn vị tiền tệ. Ưu điểm nổi bật của thước đo phi tài chính là khả

năng lý giải các vấn đề không được trình bày trên BCTC như thông tin về lợi thế cạnh tranh, thị phần... Các thước đo phi tài chính còn có khả năng phản ánh một cách chính xác, khách quan mọi mặt của doanh nghiệp. Ngoài ra, chúng rất đa dạng về phương pháp tính toán, đo lường và đặc biệt có thể dự báo kết quả tài chính trong tương lai tốt hơn so với những thước đo tài chính. Chính vì vậy đã có rất nhiều những nghiên cứu về chủ đề gian lận trước đây chỉ ra rằng không chỉ các thông tin tài chính, mà cả những thông tin phi tài chính cũng góp phần dự báo BCTC có gian lận một cách hiệu quả. Có thể kể đến như nghiên cứu của M. S. Beasley [7], Dechow et al. [17], Peasnell, Pope, Young [34].

Những nghiên cứu này đã xây dựng các mô hình để phát hiện hành vi gian lận BCTC dựa trên việc xem xét các nhân tố trong nội bộ quản trị doanh nghiệp như cấu trúc chủ sở hữu, đặc điểm Ban giám đốc. Nghiên cứu của M. S. Beasley [7] từ mô hình hồi quy logit phát hiện rằng tỷ lệ gian lận BCTC thường thấp đối với các doanh nghiệp có thành viên bên ngoài chiếm tỷ lệ phần trăm cao hơn đáng kể trong hội đồng quản trị. Nhận định về mức độ lập của thành viên Hội đồng quản trị có ảnh hưởng đến khả năng BCTC có gian lận cũng đã được chứng minh ở nghiên cứu của Dechow et al [7] và nhận thấy rằng một trong những động cơ quan trọng cho hành vi gian lận BCTC là muốn thu hút nguồn vốn đầu tư bên ngoài với chi phí thấp và giảm áp lực về tài chính. Nghiên cứu tập trung ở ba khía cạnh của hành vi gian lận là lựa chọn phương pháp kế toán, tổng các khoản kế toán dồn tích và ước lượng khoản kế toán dồn tích không hợp lý. Khi quy mô của HĐQT quá nhỏ hoặc quá lớn đều có thể làm cho HĐQT hoạt động kém hiệu quả. Lipton & Lorsch [30], Jensen [24] cho rằng quy mô HĐQT có ảnh hưởng đến tính hữu hiệu của việc giám sát. HĐQT càng lớn, sự giám sát càng kém hiệu quả, một hội đồng lí tưởng chỉ nên gồm từ 8 đến 9 thành viên. Các nghiên cứu của Yermack [46], Abbott et al. [1] đã chỉ ra quy mô HĐQT nhỏ sẽ hiệu quả hơn bởi vì giữa các thành viên có thể giao tiếp tốt hơn, dễ dàng tham gia vào giám sát DN tốt hơn.

Mark S Beasley [7] cho rằng nhà đầu tư thường xem việc hiện diện của các thành viên không điều hành trong HĐQT như là một tín hiệu quản trị tốt. Theo [17, 27] HĐQT càng "độc lập" thì DN càng ít khả năng thực hiện hành vi chi phối lợi nhuận. Gul & Leung [20] kết luận thành viên HĐQT không điều hành có ảnh hưởng tích cực đến chất lượng BCTC vì 3 lí do: Họ kiểm soát người quản lí và giảm chi phí đại diện; Họ xây dựng mạng lưới các mối quan hệ với các đối tác bên ngoài của DN; sự hiện diện của họ truyền đạt một tín hiệu của sự minh bạch tài chính. Perry & Peyer [36] cho rằng thành viên HĐQT không điều hành có thể nâng cao giá trị DN vì để nâng cao danh tiếng, họ sẽ tăng cường giám sát và phát hiện các hành vi gian lận nhiều hơn.

Từ tổng quan nghiên cứu các nhân tố ảnh hưởng đến gian lận BCTC theo thước đo tài chính và thước đo phi tài chính, chúng tôi sẽ xây dựng mô hình xem xét ảnh hưởng của các nhân tố này đến gian lận BCTC trên cơ sở áp dụng thuật toán rừng ngẫu nhiên, với kỳ vọng cho kết quả dự đoán với độ chính xác cao.

Bảng 1. Thống kê mẫu nghiên cứu

Năm	Industry									Tổng
	Bất động sản và Xây dựng	Công nghệ	Công nghiệp	Dịch vụ	Hàng tiêu dùng	Năng lượng	Nông nghiệp	Nguyên vật liệu	Y tế	
2014	105	10	40	32	27	29	36	28	11	318
2015	109	10	42	33	28	29	37	28	11	327
2016	111	10	42	33	28	30	36	29	11	330
2017	112	10	42	31	27	30	36	29	11	328
2018	107	10	42	30	25	30	36	28	11	319
2019	111	10	42	30	27	30	37	29	11	327
2020	96	9	39	29	24	26	26	27	10	286
Tổng	751	69	289	218	186	204	244	198	76	2,235
Tỷ lệ	33,6%	3,1%	12,9%	9,8%	8,3%	9,1%	10,9%	8,9%	3,4%	100%

3. PHƯƠNG PHÁP NGHIÊN CỨU

Đo lường biến phụ thuộc

Gian lận BCTC được biểu hiện thông qua các chỉ tiêu tài chính cơ bản như lợi nhuận, doanh thu, chi phí, tài sản, nợ. Ở các nước phát triển, cơ quan có thẩm quyền/tổ chức liên quan công bố các công ty có gian lận BCTC. Tại Việt Nam, chưa có tổ chức nào công bố các doanh nghiệp một cách đầy đủ và toàn diện về các công ty vi phạm gian lận BCTC, do đó trong nghiên cứu này gian lận được xác định bằng tỷ lệ chênh lệch lợi nhuận sau thuế số trước và sau kiểm toán, sau đó tính theo giá trị tuyệt đối. Vận dụng nguyên tắc trọng yếu trong kiểm toán, mức trọng yếu là từ 5% đối với chỉ tiêu lợi nhuận, theo đó các chênh lệch lợi nhuận từ 5% trở lên được xem là có gian lận. Biến phụ thuộc là chênh lệch lợi nhuận sau thuế được tính toán như sau:

$$\text{Tỷ lệ chênh lệch lợi nhuận} = \frac{\text{Lợi nhuận sau kiểm toán} - \text{Lợi nhuận trước kiểm toán}}{\text{Lợi nhuận sau kiểm toán}} \times 100$$

Đo lường các biến độc lập

Trên cơ sở lý thuyết tam giác gian lận, lý thuyết quản trị công ty và tổng quan các nghiên cứu như [2, 7, 11, 26, 37, 42], chúng tôi xây dựng và thiết lập 60 thuộc tính của 6 nhóm gồm: Sự ổn định tài chính; Dòng tiền trong doanh nghiệp; Áp lực từ bên thứ ba; Khả năng tạo doanh thu; Đặc điểm của ngành và các hoạt động của doanh nghiệp; Giám sát của HĐQT và ban kiểm soát và thái độ và các yếu tố khác. Các thuộc tính được mô tả tóm tắt ở Phụ lục 1.

Thuật toán rừng ngẫu nhiên

Random Forest (rừng ngẫu nhiên) là phương phân lớp thuộc tính được phát triển bởi Leo Breiman tại đại học California, Berkeley. Breiman cũng là đồng tác giả của phương pháp CART (Classification and Regression Trees) được đánh giá là một trong mười phương pháp khai phá dữ liệu. Trong random forest, việc cải thiện một cách đáng kể

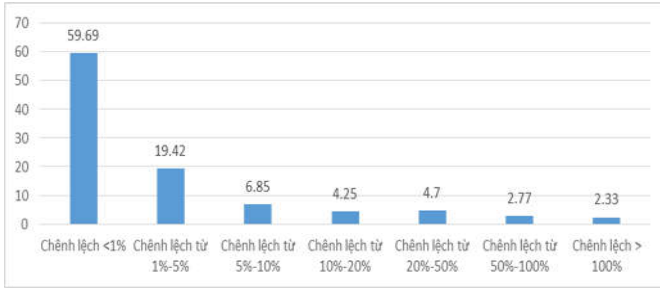
trong độ chính xác phân lớp là kết quả có được từ sự phát triển của một tập hợp các cây, mỗi cây trong tập hợp sẽ “bỏ phiếu” cho lớp phổ biến nhất. Để phát triển các tập hợp cây này thông thường các véc tơ ngẫu nhiên được tạo ra, các véc tơ này sẽ chi phối sự phát triển của mỗi cây trong các tập nói trên. Đối với cây thứ k trong tập các cây, một véc tơ ngẫu nhiên V_k được tạo ra, véc tơ này độc lập với các véc tơ được tạo ra trước đó V_1, V_2, \dots, V_{k-1} nhưng sự phân bố của các véc tơ này là tương tự nhau. Một cây được phát triển dựa vào tập tập huấn và véc tơ V_k kết quả là được một phân lớp $h(x, V_k)$ trong đó x là véc tơ đầu vào. Sau khi một số lượng lớn các cây được tạo ra các cây này “bỏ phiếu” cho lớp phổ biến nhất.

Dữ liệu nghiên cứu

Nghiên cứu này sử dụng dữ liệu được thu thập thị trường chứng khoán Việt Nam trong giai đoạn 2014 - 2020. Dữ liệu được thu thập từ các BCTC đã được kiểm toán của doanh nghiệp niêm yết sau khi loại bỏ các doanh nghiệp thuộc lĩnh vực ngân hàng, chứng khoán, bảo hiểm. Sau khi xác định các chỉ tiêu, dữ liệu được sử dụng để thực hiện phân tích và dự báo là 2235 quan sát, được trình bày tại bảng 1 theo năm và theo ngành. Trên cơ sở dữ liệu nghiên cứu, tiếp theo để huấn luyện, lựa chọn và kiểm tra kết quả của mô hình chúng ta sẽ phân chia một cách ngẫu nhiên, không trùng lặp bộ dữ liệu thành các tập train /test.

4. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

Theo kết quả hình 1, trong tổng số 2235 công ty được khảo sát, không có chênh lệch và chênh lệch lợi nhuận sau thuế dưới 1% chiếm tỷ lệ 59,69%, số doanh nghiệp có chênh lệch lợi nhuận sau thuế không trọng yếu từ 1% - 5% chiếm tỷ lệ 19,42%. Số doanh nghiệp có sự chênh lệch lợi nhuận sau thuế từ 5% đến 10% chiếm tỷ lệ 6,85%, số doanh nghiệp có chênh lệch lợi nhuận sau thuế từ 10% đến 20% là 4,7% và đặc biệt số doanh nghiệp có chênh lệch lợi nhuận sau thuế từ 50% - 100% và trên 100% lần lượt là 2,77% và 2,33%.



Hình 1. Tổng hợp mức độ chênh lệch lợi nhuận sau thuế trước và sau kiểm toán

Trong số 2235 quan sát, có 1428 quan sát quan sát có chênh lệch lợi nhuận sau thuế trước và sau kiểm toán với tỷ lệ là 63,9%, trong đó chênh lệch giảm chiếm tỷ lệ 33,1% trong khi đó số công ty chênh lệch tăng chiếm tỷ lệ 30,8%. Như vậy các công ty có xu hướng chênh lệch lợi nhuận sau thuế trước và sau kiểm toán là giảm, nghĩa là lợi nhuận sau thuế sau khi thực hiện kiểm toán thấp hơn so với lợi nhuận sau thuế chưa được kiểm toán. Như vậy các công ty có xu hướng thổi phồng lợi nhuận nhiều hơn các công ty giấu bớt lợi nhuận.

Bảng 2. Chênh lệch lợi nhuận sau thuế trước và sau kiểm toán theo xu hướng

Năm	Chênh lệch			Tổng
	Giảm	Không thay đổi	Tăng	
2014	96	133	89	318
2015	97	137	93	327
2016	115	108	107	330
2017	112	106	110	328
2018	114	103	102	319
2019	115	118	94	327
2020	90	102	94	286
Tổng	739	807	689	2,235
Tỷ lệ	33,1%	36,1%	30,8%	100,0%

Theo kết quả thống kê ở bảng 3, tỷ lệ số quan sát thuộc diện gian lận BCTC tương ứng với các ngưỡng trên 5%, là 20,9% như vậy có sự mất cân bằng về dữ liệu giữa nhóm quan sát không gian lận và gian lận BCTC. Để xử lý dữ liệu mất cân bằng, chúng tôi sử dụng các kỹ thuật SMOTE (Synthetic Minority Over-sampling) trong quá trình phân tích và dự báo.

Bảng 3. Tổng hợp các ngưỡng gian lận

Ngưỡng gian lận	Không gian lận		Gian lận	
	Số lượng	Tỷ lệ (%)	Số lượng	Tỷ lệ (%)
Chênh lệch trên 5%	1768	79,1%	467	20,9%

Kết quả bảng 4, dựa trên thuật toán Random Forest là thuật toán thuộc lớp mô hình kết hợp (ensemble model). Kết quả của thuật toán cho thấy độ chính xác của theo các thước đo Precision, Recall, F1-score đều trên 90%. Cụ thể tại ngưỡng gian lận lợi nhuận trên 5%, mức độ dự báo Precision là 90%, Recall là 92% và F1-score là 91%, mức độ

dự báo tổng quát là 91%. Trong 60 chỉ số được đưa vào mô hình dự báo, kết quả nghiên cứu có sự thống nhất cao về tầm quan trọng ảnh hưởng của các thuộc tính đến khả năng dự báo gian lận BCTC. Tại phụ lục 2, có 3 chỉ số có ảnh hưởng lớn nhất đến gian lận BCTC lần lượt là Lợi nhuận sau thuế trên tổng tài sản (X10), Lợi nhuận trước thuế trên tổng tài sản (X2), Lợi nhuận sau thuế trên vốn chủ sở hữu (X7), (Phụ lục 2). Kết quả nghiên cứu này tương đồng và độ dự báo chính xác hơn so với nghiên cứu của [11, 13, 19, 23]. Kết quả nghiên cứu này phù hợp, đồng thuận với các nghiên cứu của [13, 23, 28, 42], khi cho rằng 3 yếu tố động cơ, cơ hội và thái độ có mối quan hệ với hành vi gian lận BCTC.

Bảng 4. Kết quả dự báo chính xác theo ngưỡng gian lận dựa trên thuật toán Random Forest

Chênh lệch lợi nhuận		Precision	Recall	F1-score
Trên 5%	Không gian lận (No Fraud)	0.92	0.9	0.91
	Gian lận (Fraud)	0.9	0.92	0.91
	Trung bình (weighted avg)	0.91	0.91	0.91

5. KẾT LUẬN VÀ KHUYẾN NGHỊ

Với mẫu dữ liệu bao gồm 2235 quan sát trong giai đoạn 2014 - 2020, kết quả nghiên cứu cho thấy mô hình được xây dựng dựa trên hai yếu tố Động cơ (sự ổn định về tài chính): Lợi nhuận sau thuế trên tổng tài sản (X10), Lợi nhuận trước thuế trên tổng tài sản (X2), Lợi nhuận sau thuế trên vốn chủ sở hữu (X7) là 3 thuộc tính có ảnh hưởng lớn nhất. Mô hình này có khả năng dự báo chính xác trên 91% khi sử dụng thuật toán rừng ngẫu nhiên. Kết quả nghiên cứu đã giúp trả lời cho câu hỏi nghiên cứu những nhân tố quan trọng có ảnh hưởng đến gian lận BCTC đã được xác định với khả năng dự đoán chính xác cao. Trên cơ sở kết quả nghiên cứu, chúng tôi đề xuất một số khuyến nghị:

- Sự bất ổn tài chính của doanh nghiệp, dẫn đến áp lực từ bên ngoài, hay nói cách khác là sự kỳ vọng của bên thứ ba dẫn đến gian lận BCTC. Do đó, để hạn chế cũng như ngăn chặn hành vi này, doanh nghiệp cần gia tăng hiệu quả giám sát, xây dựng cơ chế lương thưởng, thù lao cho ban lãnh đạo một cách hợp lý. Thêm vào đó, cần chú ý đến các chỉ tiêu này qua các năm và từ đó so với số bình quân của ngành.

- Sự tác động qua HĐQT và ban kiểm soát với hành vi gian lận BCTC, điều này là phù hợp với thực tế, vì những người có trình độ học vấn càng cao sẽ phát hiện ra được những bất ổn trên BCTC và từ đó họ sẽ đưa ra những biện pháp khắc phục phù hợp trong việc ngăn chặn gian lận BCTC.

- Tăng cường tính giám sát bằng cách ban hành các văn bản nhằm nâng cao vai trò, phân định rõ trách nhiệm của ủy ban chứng khoán Việt Nam, sở giao dịch chứng khoán trong quá trình giám sát hoạt động công bố thông tin của doanh nghiệp niêm yết. Có như vậy các doanh nghiệp niêm yết mới có trách nhiệm và nghĩa vụ thực hiện một cách nghiêm túc.

- Nâng cao chất lượng nhân lực kế toán, các chênh lệch lợi nhuận sau thuế trên một phần là gian lận mà người lập BCTC cố ý tạo nên, cũng có một phần không nhỏ là từ những nguyên nhân chủ quan như trình độ nghiệp vụ của người làm kế toán còn yếu kém, thiếu tính trung thực khách quan, cũng như khả năng quan sát, phân tích tổng hợp.

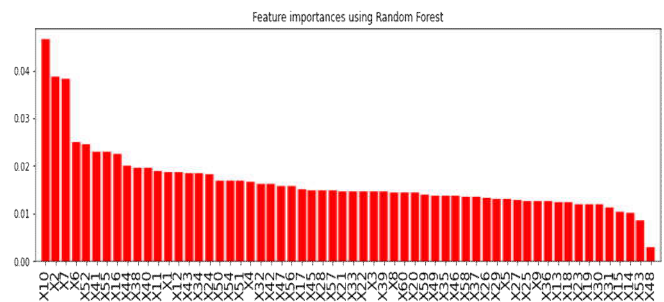
Phụ lục 1. Các thuộc tính trong mô hình nghiên cứu

Ký hiệu	Tên biến	Đo lường
	Biến phụ thuộc	
	Gian lận	Gian lận = 1: Chênh lệch lợi nhuận lớn hơn 5%; 10%; 20%; 50% Ngược lại = 0
	Biến độc lập	
	Động cơ/Áp lực	
I	Sự ổn định tài chính	
X1	Lợi nhuận từ hoạt động sản xuất kinh doanh trên doanh thu thuần	Lợi nhuận từ hoạt động sản xuất kinh doanh t/Doanh thu thuần t
X2	Lợi nhuận trước thuế trên tổng tài sản	Lợi nhuận trước thuế t/Tổng tài sản t
X3	Lợi nhuận gộp trên tài sản cố định	Lợi nhuận gộp t/Tài sản cố định t
X4	Lợi nhuận sau thuế trên giá vốn hàng bán	Lợi nhuận sau thuế t/Giá vốn hàng bán t
X5	Lợi nhuận gộp trên doanh thu thuần	Lợi nhuận gộp t/Doanh thu thuần t
X6	Lợi nhuận gộp trên tổng tài sản	Lợi nhuận gộp t/Tổng tài sản t
X7	Lợi nhuận sau thuế trên vốn chủ sở hữu	Lợi nhuận sau thuế t/Vốn chủ sở hữu t
X8	Giá vốn hàng bán trên lợi nhuận gộp	Giá vốn hàng bán t/Lợi nhuận gộp t
X9	Lợi nhuận gộp trên giá vốn hàng bán	Lợi nhuận gộp t/Giá vốn hàng bán t
X10	Lợi nhuận sau thuế trên tổng tài sản	Lợi nhuận sau thuế t/Tổng tài sản t
X11	Lợi nhuận từ hoạt động sản xuất kinh doanh trên lợi nhuận gộp	Lợi nhuận từ hoạt động sản xuất kinh doanh t/Lợi nhuận gộp t
X12	Lợi nhuận sau thuế trên doanh thu thuần	Lợi nhuận sau thuế t/Doanh thu thuần t
II	Dòng tiền trong doanh nghiệp	
X13	Tổng biến đổi tích trên tổng tài sản	(Lợi nhuận sau thuế t- Dòng tiền thuần từ hd sxkd t)/Tổng tài sản t
X14	Dòng tiền thuần từ hoạt động sản xuất kinh doanh trên giá vốn hàng bán	Dòng tiền thuần từ hoạt động sản xuất kinh doanh t/Giá vốn hàng bán t
X15	Dòng tiền thuần từ hoạt động sản xuất kinh doanh trên doanh thu thuần	Dòng tiền thuần từ hoạt động sản xuất kinh doanh t/doanh thu thuần t
X16	Lợi nhuận sau thuế trên dòng tiền thuần từ hoạt động sản xuất kinh doanh	Lợi nhuận sau thuế t/Dòng tiền thuần từ hoạt động sản xuất kinh doanh t

X17	Dòng tiền thuần từ hoạt động sản xuất kinh doanh trên tổng tài sản	Dòng tiền thuần từ hoạt động sản xuất kinh doanh t/Tổng tài sản t
X18	Dòng tiền thuần từ hoạt động sản xuất kinh doanh trên lợi nhuận sau thuế	Dòng tiền thuần từ hoạt động sản xuất kinh doanh t/Lợi nhuận sau thuế t
	Áp lực từ bên thứ ba	
X19	Vốn lưu động trên tổng tài sản	(Tài sản ngắn hạn t – Nợ ngắn hạn t)/Tổng tài sản t
X20	Tiền và tương đương tiền trên nợ ngắn hạn	Tiền và tương đương tiền t/Nợ ngắn hạn t
X21	Tiền và tương đương tiền trên tài sản ngắn hạn	Tiền và tương đương tiền t/Tài sản ngắn hạn t
X22	Tiền và tương đương tiền trên tổng tài sản	Tiền và tương đương tiền t / tổng tài sản t
X23	Tài sản ngắn hạn trên nợ ngắn hạn	Tài sản ngắn hạn t/Nợ ngắn hạn t
X24	Nợ ngắn hạn trên tổng nợ	Nợ ngắn hạn t/Tổng nợ t
X25	Vốn chủ sở hữu trên tổng nợ	Vốn chủ sở hữu t/Tổng nợ t
X26	Tài sản ngắn hạn trên tổng nợ	Tài sản ngắn hạn t/Tổng nợ t
X27	Đòn cân nợ	Tổng nợ t/ Tổng tài sản t
X28	Dòng tiền thuần từ hoạt động sản xuất kinh doanh trên tổng nợ	Dòng tiền thuần từ hoạt động sản xuất kinh doanh t/Tổng nợ t
X29	Vốn chủ sở hữu trên tổng tài sản	Vốn chủ sở hữu t/Tổng tài sản t
X30	Vốn chủ sở hữu trên tài sản dài hạn	Vốn chủ sở hữu t/(Tổng tài sản t – Tài sản ngắn hạn t)
X31	Hệ số thanh toán Nhanh	(Tài sản ngắn hạn t – Hàng tồn kho t)/Nợ ngắn hạn t
III	Khả năng tạo doanh thu	
X32	Doanh thu thuần trên tài sản cố định	Doanh thu thuần t/Tài sản cố định t
X33	Doanh thu thuần trên tổng tài sản	Doanh thu thuần t/Tổng tài sản t
X34	Tài sản cố định trên tổng tài sản	Tài sản cố định t/Tổng tài sản t
X35	Tỷ lệ doanh thu thuần trên giá vốn hàng bán	(Doanh thu thuần t/Giá vốn hàng bán t)/ (Doanh thu thuần t-1/Giá vốn hàng bán t-1)
X36	Tỷ lệ doanh thu thuần trên tổng chi phí cơ bản	(Doanh thu thuần t/(Giá vốn hàng bán t + Chi bán hàng t + Chi phí quản lý doanh nghiệp t))/ (Doanh thu thuần t-1/(Giá vốn hàng bán t-1 + Chi bán hàng t-1 + Chi phí quản lý doanh nghiệp t-1))
X37	Tỷ lệ doanh thu thuần trên chi phí bán hàng và quản lý doanh nghiệp	(Doanh thu thuần t/(Chi bán hàng t + Chi phí quản lý doanh nghiệp t))/ (Doanh thu thuần t-1/(Chi bán hàng t-1 + Chi phí quản lý doanh nghiệp t-1))
IV	Đặc điểm của ngành hay các hoạt động của doanh nghiệp	
X38	Hàng tồn kho trên tổng tài sản	Hàng tồn kho t/Tổng tài sản t
X39	Hàng tồn kho trên tài sản ngắn hạn	Hàng tồn kho t/Tài sản ngắn hạn t

X40	Hàng tồn kho trên giá vốn hàng bán	Hàng tồn kho t/Giá vốn hàng bán t
X41	Tỉ lệ nợ phải thu trên doanh thu	(Phải thu khách hàng ngắn hạn t – Dự phòng nợ khó đòi t + Phải thu khách hàng dài hạn t – Dự phòng phải thu dài hạn t)/Doanh thu thuần t
V Giám sát của HĐQT và ban kiểm soát		
X42	Sự kiêm nhiệm giữa chức danh chủ tịch hội đồng quản trị và giám đốc điều hành	Biến định danh. CEO có giá trị là 1 nếu chủ tịch hội đồng quản trị kiêm nhiệm giám đốc điều hành, ngược lại, biến có giá trị là 0
X43	Số lượng thành viên HĐQT là cổ đông lớn	Số lượng thành viên HĐQT nắm giữ cổ phần trên 5%
X44	Quy mô của HĐQT	Số lượng thành viên trong HĐQT
X45	Tỷ lệ thành viên HĐQT có tham gia điều hành	Số lượng thành viên HĐQT thanh gia điều hành/Tổng số thành viên HĐQT
X46	Tỷ lệ thành viên HĐQT là nữ	Số thành viên HĐQT là nữ/Tổng số thành viên HĐQT
X47	Số lần họp của HĐQT trong năm	Số buổi họp HĐQT trong năm
X48	Quy mô ban kiểm soát	Số thành viên trong ban soát
X49	Tỷ lệ nữ trong ban kiểm soát	Số lượng thành viên nữ trong ban kiểm soát/Tổng số thành viên ban kiểm soát
X50	Tỷ lệ thành viên trong ban kiểm soát có chuyên môn tài chính.	Số lượng thành viên có chuyên môn tài chính/Tổng số thành viên ban kiểm soát
X51	Số lần họp ban kiểm soát trong năm	Số buổi họp ban kiểm soát trong năm
VI Thái độ và các yếu tố khác		
X52	Kiểm toán viên thuộc nhóm Big 4	Biến định danh. BIG4 có giá trị là 1 nếu được kiểm toán bởi công ty thuộc nhóm không phải là Big 4.
X53	Báo cáo tài chính hợp nhất	Biến định danh. Báo cáo tài chính hợp nhất là 1, còn lại là 0.
X54	Quy mô doanh nghiệp	Logarit (Tổng tài sản)
X55	Thời gian niêm yết	Số năm niêm yết
X56	Tăng trưởng doanh thu	(Doanh thu t – Doanh thu t-1)/Doanh thu thuần t-1
X57	Chỉ số Z	Mô hình Z-score [4]
X58	Chỉ số M-Score	Mô hình M-score của M. D Beneish [11]
X59	Sở hữu Nhà nước	Tỷ lệ sở hữu nhà nước
X60	Sở hữu nước ngoài	Tỷ lệ sở hữu nước ngoài

Phụ lục 2. Tầm quan trọng của các thuộc tính



TÀI LIỆU THAM KHẢO

[1]. Abbott L. J., Parker S., Peters G. F., 2004. *Audit committee characteristics and restatements*. Auditing: A Journal of Practice & Theory, 23(1), 69-87.

[2]. Albizri A., Appelbaum D., Rizzotto N., 2019. *Evaluation of financial statements fraud detection research: a multi-disciplinary analysis*. International Journal of Disclosure and Governance, 16(4), 206-241.

[3]. Albrecht W. S., Romney M. B., Cherrington D. J., Payne I. R., Roe A. J., Romney M. B., 1986. *Red-flagging management fraud: A validation*. Advances in Accounting, 3, 323-333.

[4]. Altman E. I., 1968. *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy*. The journal of finance, 23(4), 589-609.

[5]. Baber W. R., Kang S., Liang L., 2005. *Strong boards, management entrenchment, and accounting restatement*. George Washington University.

[6]. Bayley L., Taylor S. L., 2000. *Identifying earnings overstatements: A practical test*. Available at SSRN 995957.

[7]. Beasley M. S., 1996. *An empirical analysis of the relation between the board of director composition and financial statement fraud*. Accounting review, 71(4), 443-465. doi: <https://www.jstor.org/stable/248566>

[8]. Beasley M. S., 1996. *An empirical analysis of the relation between the board of director composition and financial statement fraud*. Accounting review, 443-465.

[9]. Beasley M. S., Hermanson D. R., Carcello J. V., Nea, T. L., 2010. *Fraudulent financial reporting: 1998-2007: An analysis of US public companies*.

[10]. Beneish M. D., 1997. *Detecting GAAP violation: Implications for assessing earnings management among firms with extreme financial performance*. Journal of Accounting and Public Policy, 16(3), 271-309.

[11]. Beneish M. D., 1999. *Incentives and penalties related to earnings overstatements that violate GAAP*. The Accounting Review, 74(4), 425-457.

[12]. Chen G., Firth M., Gao D. N., Rui O. M., 2006. *Ownership structure, corporate governance, and fraud: Evidence from China*. Journal of Corporate Finance, 12(3), 424-448.

[13]. Chen S., 2016. *Detection of fraudulent financial statements using the hybrid data mining approach*. SpringerPlus, 5(1), 1-16.

[14]. Cressey D. R., 1953. *Other people's money; a study of the social psychology of embezzlement*.

[15]. Cybenko G., 1992. *Approximation by superpositions of a sigmoidal function*. Mathematics of Control, Signals and Systems, 5(4), 455-455.

[16]. Dechow P. M., Ge W., Larson C. R., Sloan R. G., 2011. *Predicting material accounting misstatements*. Contemporary accounting research, 27(1), 17-82.

- [17]. Dechow P. M., Sloan R. G., Sweeney A. P., 1996. *Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC*. Contemporary Accounting Research, 13(1), 1-36.
- [18]. Feroz E. H., Kwon T. M., Pastena V. S., 2000. *The efficacy of red flags in predicting the SEC's targets: an artificial neural networks approach*. Intelligent Systems in Accounting, Finance & Management, 9(3), 145-157.
- [19]. Green B. P., Choi J. H., 1997. *Assessing the risk of management fraud through neural network technology*. Auditing, 16, 14-28.
- [20]. Gul F. A., Leung S., 2004. *Board leadership, outside directors' expertise and voluntary corporate disclosures*. Journal of Accounting and Public Policy, 23(5), 351-379.
- [21]. Hornik K., Stinchcombe M., White H., 1989. *Multilayer feedforward networks are universal approximators*. Neural networks, 2(5), 359-366.
- [22]. Hung D. N., Ha H. T. V., Binh D. T., 2017. *Application of F-Score in Predicting Fraud, Errors: Experimental Research in Vietnam*. International Journal of Accounting and Financial Reporting, 7(2), 303-322.
- [23]. Jan C., 2018. *An effective financial statements fraud detection model for the sustainable development of financial markets: Evidence from Taiwan*. Sustainability, 10(2), 513.
- [24]. Jensen M. C., 1993. *The modern industrial revolution, exit, and the failure of internal control systems*. The journal of finance, 48(3), 831-880.
- [24]. Jones J. J., 1991. *Earnings management during import relief investigations*. Journal of accounting research, 29(2), 193-228. doi: DOI: 10.2307/2491047
- [26]. Kanapickienė R., Grundienė Ž., 2015. *The model of fraud detection in financial statements by means of financial ratios*. Procedia-Social and Behavioral Sciences, 213, 321-327.
- [27]. Klein A., 2002. *Audit committee, board of director characteristics, and earnings management*. Journal of accounting and economics, 33(3), 375-400.
- [28]. Kotsiantis S., Koumanakos E., Tzelepis D., Tampaka, V. (2006). *Predicting fraudulent financial statements with machine learning techniques*. Paper presented at the Hellenic Conference on Artificial Intelligence.
- [29]. Lenard M. J., Ala P., 2009. *An historical perspective on fraud detection: From bankruptcy models to most effective indicators of fraud in recent incidents*. Journal of Forensic & Investigative Accounting, 1(1), 1-27.
- [30]. Lipton M., Lorsch J. W., 1992. *A modest proposal for improved corporate governance*. The business lawyer, 48(1), 59-77.
- [31]. Loebbecke J. K., Eining M. M., Willingham J. J., 1989. *Auditors experience with material irregularities-frequency, nature, and detectability*. Auditing: A Journal of Practice & Theory, 9(1), 1-28.
- [32]. Lou Y. I., Wang M. L., 2011. *Fraud risk factor of the fraud triangle assessing the likelihood of fraudulent financial reporting*. Journal of Business & Economics Research, 7(2).
- [33]. McNelis P. D., 2005. *Neural networks in finance: gaining predictive edge in the market*. Academic Press.
- [34]. Peasnell K. V., Pope P. F., Young S., 2005. *Board monitoring and earnings management: Do outside directors influence abnormal accruals?* Journal of Business Finance & Accounting, 32(7-8), 1311-1346.
- [35]. 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.
- [36] Perry T., Peyer U., 2005. *Board seat accumulation by executives: A shareholder's perspective*. The journal of finance, 60(4), 2083-2123.
- [37]. Persons O. S., 1995. *Using financial statement data to identify factors associated with fraudulent financial reporting*. Journal of Applied Business Research, 11(38-46).
- [38]. Rezaee Z., 2002. *The three Cs of fraudulent financial reporting*. Internal Auditor, 59(5), 56-62.
- [39]. Romney M. B., Albrecht W. S., Cherrington D. J. (1980). *Auditors and the Detection of Fraud*. Journal of Accountancy, 149(5), 63-69.
- [40]. Sharma A., Panigrahi P. K., 2013. *A review of financial accounting fraud detection based on data mining techniques*. arXiv preprint arXiv:1309.3944.
- [41]. Simunic D. A., Stein M. T., 1987. *Product differentiation in auditing: Auditor choice in the market for unseasoned new issues: Canadian Certified General*.
- [42]. Skousen C. J., Smith K. R., Wright C. J., 2008. *Detecting and predicting financial statement fraud: The effectiveness of the fraud triangle and SAS No. 99 In Corporate Governance and Firm Performance*. Advances in Financial Economics, 53-81.
- [43]. Summers S. L., Sweeney J. T., 1998. *Fraudulently misstated financial statements and insider trading: An empirical analysis*. Accounting review, 131-146.
- [44]. Tran M. D., Dang N. H., Hoang T. V. H., 2017. *Research on Misstatements in Financial Statements: The Case of Listed Firms on Ho Chi Minh City Stock Exchange*. International Journal of Applied Business and Economic Research, 15(23(Part 2)), 499-518.
- [45]. Wuerges A. F. E., Borba J. A., 2010. *Accounting fraud detection: is it possible to quantify undiscovered cases?* Available at SSRN 1718652.
- [46]. Yermack D., 1996. *Higher market valuation of companies with a small board of directors*. Journal of financial economics, 40(2), 185-211. doi: https://doi.org/10.1016/0304-405X(95)00844-5
- [47]. Zainudin E. F., Hashim H. A., 2016. *Detecting fraudulent financial reporting using financial ratio*. Journal of Financial Reporting and Accounting.
- [48]. Zhang G., Patuwo B. E., Hu M. Y., 1998. *Forecasting with artificial neural networks: The state of the art*. International journal of forecasting, 14(1), 35-62.

AUTHORS INFORMATION

Dang Ngoc Hung¹, Pham Thi Hong Diep¹, Cao Thi Nhien²

¹Hanoi University of Industry

²Hanoi Polytechnic College