

DỰ BÁO RỦI RO PHÁ SẢN CÁC CÔNG TY NIÊM YẾT VIỆT NAM VỚI DỮ LIỆU KẾ TOÁN VÀ THỊ TRƯỜNG

● LÊ THỊ HẢI BÌNH¹ - LÊ THỊ MỸ NGUYỆT^{1*}

- LÊ THỊ HUYỀN TRANG² - NGUYỄN THỊ NHƯ NGỌC² - NGUYỄN THỊ ÁNH DƯƠNG²

- MAI NGỌC THẢO LY² - VÕ NGUYỄN ANH THƯ²

¹Giảng viên, Viện Tài chính Kế toán,

Trường Đại học Công nghiệp Thành Phố Hồ Chí Minh

²Sinh viên, Viện Tài chính Kế toán,

Trường Đại học Công nghiệp Thành Phố Hồ Chí Minh

*Email: lethimyinguyet@iuh.edu.vn

TÓM TẮT:

Nghiên cứu kiểm chứng mô hình dự báo rủi ro phá sản cho các công ty niêm yết tại Việt Nam dựa trên dữ liệu kế toán và dữ liệu thị trường trong giai đoạn 2016 - 2024. Bài viết sử dụng phương pháp học máy dựa trên thuật toán hồi quy Logistic trên mẫu 103 cặp doanh nghiệp đối chứng. Kết quả cho thấy, mô hình dữ liệu kế toán có hiệu quả dự báo cao nhất với AUC đạt 92,5%, 97,75% và 89%. Mô hình dữ liệu thị trường chưa thể hiện tốt khả năng dự báo rủi ro phá sản. Khi kết hợp 2 nguồn dữ liệu này, AUC không tăng thêm, chứng tỏ dữ liệu thị trường không đem lại giá trị thông tin tăng thêm để dự báo rủi ro phá sản trong mô hình dữ liệu kế toán. Kết quả nhấn mạnh vai trò vượt trội của thông tin kế toán trong dự báo rủi ro phá sản tại Việt Nam.

Từ khóa: rủi ro phá sản, kiệt quệ tài chính, dữ liệu kế toán, dữ liệu thị trường, mô hình hồi quy Logistic.

1. Đặt vấn đề

Theo dữ liệu từ Cục Thống kê - Bộ Tài Chính, năm 2025, số doanh nghiệp tạm ngừng kinh doanh có thời hạn là 114.400 doanh nghiệp, tăng 14,3% so với năm trước; 76.900 doanh nghiệp ngừng hoạt động chờ làm thủ tục giải thể, tăng 0,9%; 35.900 doanh nghiệp hoàn tất thủ tục giải thể, tăng 66,1%. Bình

quân một tháng có 18.900 doanh nghiệp rút lui khỏi thị trường (Cục Thống kê, 2026). Trước tình hình này, rủi ro phá sản (RRPS) hay kiệt quệ tài chính là đề tài thu hút sự quan tâm từ nhiều nhà đầu tư, nhà quản trị doanh nghiệp và cơ quan quản lý nhà nước. Các nhà nghiên cứu quốc tế nói chung và Việt Nam nói riêng đều chú trọng nhiều đến dự báo RRPS của các

doanh nghiệp, nhằm đưa ra những mô hình dự báo sớm cho các nhà quản lý, các nhà đầu tư và người sử dụng thông tin khác.

Trên thế giới, nghiên cứu về dự báo RRPS đã được phát triển từ khá sớm với nhiều mô hình và phương pháp tiếp cận khác nhau. Các nghiên cứu phát triển từ mô hình truyền thống với dữ liệu kế toán (DLKT) như Z-Score của Altman (1968), O-Score của Ohlson (1980), hoặc với dữ liệu thị trường (DLTT) như Merton (1974) đến các nghiên cứu xây dựng mô hình kết hợp giữa DLKT và DLTT như Shumway (2001), M.-Y. L. Li & Miu (2010), cũng như ứng dụng các mô hình học máy (machine learning) như Barboza et al. (2017).

Trong khi đó, tại Việt Nam, nhiều nghiên cứu đã xây dựng mô hình dự báo RRPS dựa trên sự kết hợp giữa DLKT và DLTT. Cụ thể, Phạm Võ Ninh et al. (2018) sử dụng dữ liệu giai đoạn 2003-2016 của các công ty niêm yết Việt Nam, kết hợp nhiều yếu tố trong mô hình dự báo gồm yếu tố kế toán, thị trường và chỉ số kinh tế vĩ mô. Võ et al. (2019) xem xét tình trạng khó khăn tài chính của các công ty niêm yết ở cấp độ ngành tại Việt Nam trong thời kỳ khủng hoảng tài chính toàn cầu 2007-2009 và sau đó 2010-2017 với mô hình dựa trên biến kế toán, biến thị trường và biến vĩ mô.

Mặc dù đã có các nghiên cứu xây dựng mô hình dự báo RRPS ở Việt Nam, nhưng các nghiên cứu trước xem xét đến bối cảnh kinh tế giai đoạn trước đây, khác với tình hình hiện tại. Đồng thời, phạm vi nghiên cứu trong nước giới hạn với một ngành và giai đoạn cụ thể. Vì vậy, mục tiêu nghiên cứu này tập trung đánh giá các kết quả dự báo RRPS của các mô hình dựa trên DLKT, DLTT và kết hợp cả 2 nguồn với dữ liệu các công ty niêm yết trong bối cảnh giai đoạn 2016-2024.

Nghiên cứu này thực hiện xây dựng mô hình dự báo RRPS trên 103 công ty có RRPS và 103 công ty có tình hình tài chính lành mạnh đối ứng trong giai đoạn từ năm 2016 đến 2024. Giai đoạn 2016-2024 từ trạng thái tăng trưởng ổn định (2016-2019), rơi vào cú sốc đại dịch COVID-19 (2020-2021) và chuyển sang khủng hoảng niềm tin, thất chặt thanh

khoản (2022-2024). Việc nghiên cứu RRPS trong khoảng thời gian này giúp kiểm chứng tính hiệu quả của các mô hình dự báo dựa trên DLKT và DLTT, đồng thời mang lại giá trị thực tiễn cao cho nhà đầu tư và các tổ chức tín dụng.

2. Tổng quan nghiên cứu

Các nghiên cứu về khả năng dự báo RRPS dựa vào DLKT và DLTT có sự đa dạng về bối cảnh ứng dụng cũng như phương pháp tiếp cận, là chủ đề thu hút rất nhiều nhà nghiên cứu khai thác. Nghiên cứu Hillegeist et al. (2004) thực hiện nhằm đánh giá khả năng dự báo RRPS giữa chỉ số Z-Score của Altman (1968), O-Score của Ohlson (1980) và BSM-Prob. Nghiên cứu này đã xây dựng mô hình rủi ro rời rạc để ước tính xác suất phá sản và chỉ ra thước đo BSM-Prob dựa trên DLTT có khả năng dự báo vượt trội hơn so với 2 chỉ số dựa trên DLKT.

Ở quốc tế, Nouri & Soltani (2016), Waqas et al. (2018) và Saad et al. (2025) là 3 nghiên cứu cùng sử dụng phương pháp phân tích hồi quy Logit trong quá trình thực hiện mô hình dự báo RRPS. Nghiên cứu Nouri & Soltani (2016) và Saad et al. (2025) tập trung thiết kế phát triển mô hình dự báo phá sản. Trong khi đó, Waqas et al. (2018) xác định các yếu tố dự đoán khó khăn tài chính cho các công ty niêm yết tại Pakistan. Kết quả nghiên cứu của Nouri & Soltani (2016) cho thấy, các biến thông tin kế toán dự báo RRPS tốt hơn so với thông tin thị trường, nhưng mô hình kết hợp cả 2 yếu tố đã mang lại khả năng dự báo RRPS vượt trội hơn hẳn. Kết quả nghiên cứu của Saad et al. (2025) cho thấy, mô hình kết hợp biến kế toán và biến thị trường mang lại khả năng dự báo tốt hơn. Các nghiên cứu của Ruxanda et al. (2018), Lotfi et al. (2024) và Correia (2025) chủ yếu tập trung vào chủ đề khám phá vai trò của thông tin kế toán, tìm hiểu sức mạnh của các mô hình dự báo RRPS và so sánh các kỹ thuật giám sát, xác định các chỉ số kinh tế đóng vai trò là chỉ số cảnh báo sớm cho RRPS của doanh nghiệp.

Tại Việt Nam, nghiên cứu của Hoàng Thị Hồng Vân (2020) đánh giá khả năng áp dụng mô hình Z-Score vào thực tiễn Việt Nam. Tác giả tập trung vào

việc kiểm định độ chính xác của mô hình trong việc phân loại doanh nghiệp phá sản và doanh nghiệp đang hoạt động. Mô hình Z-score của Altman có khả năng dự báo chính xác RRPS của các doanh nghiệp Việt Nam ở mức 76,67% đối với dữ liệu 1 năm trước khi phá sản và 70% đối với dữ liệu 2 năm trước khi phá sản. Theo xu hướng ứng dụng công nghệ hiện đại, Nguyễn & Ngô (2024) đã thực hiện mục tiêu xây dựng và đánh giá các mô hình học máy gồm hồi quy Logistic (LR), cây quyết định, XGBoost và Mạng nơ-ron nhân tạo (ANN) trong dự báo rủi ro vỡ nợ của doanh nghiệp vừa và nhỏ tại Việt Nam. Kết quả cho thấy, ANN và XGBoost cho kết quả dự báo rủi ro vỡ nợ chính xác vượt trội hơn các mô hình còn lại, với độ chính xác lần lượt 93% và AUC đạt 95%, khẳng định hiệu quả cao trong quản lý rủi ro tín dụng.

3. Phương pháp nghiên cứu

3.1. Dữ liệu nghiên cứu

Nghiên cứu sử dụng phương pháp định lượng với dữ liệu thứ cấp từ 206 doanh nghiệp niêm yết trên HOSE và HNX giai đoạn 2016 - 2024. Thông qua kỹ thuật ghép cặp đối ứng 1:1 giữa công ty có RRPS và công ty có tình hình tài chính lành mạnh, dựa trên ngành, quy mô tài sản và tuổi niêm yết, mẫu nghiên cứu cuối cùng bao gồm 103 cặp doanh nghiệp. Trong đó, nhóm RRPS được xác định dựa trên các tiêu chí: bị hủy niêm yết bắt buộc, lỗ ròng 3 năm liên tiếp hoặc âm vốn chủ sở hữu. Dữ liệu được thu thập trong giai đoạn 3 năm trước khi sự kiện RRPS xảy ra.

3.2. Giả thuyết nghiên cứu

Nghiên cứu này xây dựng các giả thuyết dựa trên 2 nhóm thông tin chính là thông tin kế toán và thông tin thị trường.

Các thông tin kế toán phản ánh toàn diện tình hình tài chính và hiệu quả hoạt động của doanh nghiệp. Pham Vo Ninh et al. (2018) phân tích dữ liệu lớn với 800 công ty niêm yết tại Việt Nam giai đoạn 2003-2016. Kết quả khẳng định 4 biến đại diện kế toán (khả năng thanh khoản, khả năng sinh lợi, hiệu quả sử dụng tài sản và khả năng thanh toán) có mối quan hệ nghịch biến với xác suất vỡ nợ. Lâm

Thị Thanh Huyền & Nguyễn Thị Quỳnh Châm (2024) áp dụng mô hình hồi quy Logistic trên 79 công ty bất động sản niêm yết (2015-2022). Kết quả chứng minh các thông tin kế toán có khả năng dự báo tốt với độ phân loại chính xác cho toàn mẫu lên tới 88,92%. Correia (2025) dựa trên mẫu dữ liệu từ 1980 đến 2023, nghiên cứu khẳng định sức mạnh dự báo của các mô hình kế toán (Z-Score, O-Score) không hề suy giảm theo thời gian. Đặc biệt, trong những giai đoạn khủng hoảng khi thông tin thị trường bị suy giảm sức mạnh, thông tin kế toán vẫn duy trì sự ổn định và bù đắp hiệu quả cho các mô hình dự báo. Vì vậy, nhóm tác giả đề xuất giả thuyết nghiên cứu H1 như sau:

Giả thuyết H1: Các thông tin kế toán có khả năng dự báo RRPS của các công ty niêm yết tại Việt Nam.

Các thông tin thị trường phản ánh cách nhà đầu tư đánh giá tình hình tài chính và triển vọng của doanh nghiệp. Theo Saad et al. (2025), các biến số thị trường như sự thay đổi về lợi nhuận cổ phiếu, lợi nhuận vượt trội trong quá khứ và tỷ lệ giá trị thị trường trên giá trị sổ sách ảnh hưởng đến RRPS của các công ty. Cao et al. (2024) giới thiệu mô hình Ensemble cải tiến tích hợp các mô hình học máy với cấu trúc 2 tầng của cơ chế xếp chồng để tiến hành dự đoán khả năng phá sản của doanh nghiệp. Theo đó, yếu tố thông tin thị trường đo lường lợi nhuận cổ phiếu, biến động, cổ tức, biến động giá xuống và thanh khoản cho thấy khả năng dự đoán RRPS mạnh nhất. Pham Vo Ninh et al. (2018) phát hiện có 3 biến thị trường có ý nghĩa thống kê 1% trong việc dự báo RRPS là giá cổ phiếu, giá trị thị trường của vốn chủ sở hữu và tỷ lệ đòn bẩy. Trong giai đoạn hậu khủng hoảng, nghiên cứu của Vo et al. (2019) cho thấy, 3 biến thị trường này có tác động đáng kể đến tình trạng kiệt quệ tài chính của các công ty niêm yết.

Giả thuyết H2: Các thông tin thị trường có khả năng dự báo RRPS của các công ty niêm yết tại Việt Nam.

Thông tin kế toán và thông tin thị trường phản ánh RRPS ở những góc độ khác nhau và có tính bổ

trợ lẫn nhau. Correia (2025) so sánh nhiều mô hình và kết luận mô hình kết hợp có khả năng dự báo vượt trội hơn tất cả các mô hình chỉ dùng DLKT hoặc DLTT. Saad et al. (2025) phát hiện sự kết hợp giữa biến kế toán và thị trường đóng góp tích cực và đáng kể vào độ chính xác của dự báo phá sản.

Trên cơ sở đó, nghiên cứu kỳ vọng việc kết hợp thông tin kế toán và thông tin thị trường sẽ giúp nâng cao khả năng dự báo RRPS của các công ty niêm yết tại Việt Nam.

Giả thuyết H3: Mô hình kết hợp DLKT và DLTT cho kết quả dự báo RRPS cải thiện hơn so với chỉ dùng một loại dữ liệu.

3.3. Mô hình nghiên cứu

Nghiên cứu sử dụng 3 mô hình Logistic Regression xây dựng mô hình học máy trên R để dự báo RRPS dựa trên nguồn dữ liệu khác nhau: DLKT, DLTT và kết hợp cả 2.

Mô hình 1: Các biến dự báo thu thập từ DLKT

$$\text{Logit}(p_i) = \beta_0 + \beta_1 * CR_i + \beta_2 * QR_i + \beta_3 * (D/A_i) + \beta_4 * ICR_i + \beta_5 * ROA_i + \beta_6 * NPM_i + \beta_7 * (OCF/TA_i) + \beta_8 * (WC/TA_i) + \beta_9 * TAT_i + \epsilon_i \quad (1)$$

Mô hình 2: Các biến dự báo thu thập từ DLTT

$$\text{Logit}(p_i) = \beta_0 + \beta_1 * (P/B_i) + \beta_2 * VOL_i + \beta_3 * MCAP_i + \beta_5 * AR_i + \epsilon_i \quad (2)$$

Mô hình 3 là mô hình kết hợp giữa DLKT và DLTT:

$$\text{Logit}(p_i) = \beta_0 + \beta_1 * CR_i + \beta_2 * QR_i + \beta_3 * (D/A_i) + \beta_4 * ICR_i + \beta_5 * ROA_i + \beta_6 * NPM_i + \beta_7 * (OCF/TA_i) + \beta_8 * (WC/TA_i) + \beta_9 * TAT_i + \beta_{10} * (P/B_i) + \beta_{11} * VOL_i + \beta_{12} * MCAP_i + \beta_{13} * AR_i + \epsilon_i \quad (3)$$

p_i là xác suất xảy ra RRPS ($BR_i = 1$) của công ty i. BR đại diện cho RRPS, mã hóa bằng 1 nếu công ty gặp RRPS và bằng 0 nếu công ty có tình trạng tài chính lành mạnh. Các biến độc lập trong 3 mô hình được thể hiện trong Bảng 1.

Thay vì sử dụng phương pháp hợp lý cực đại như hồi quy Logistic truyền thống, nghiên cứu này áp dụng thuật toán Penalized Logistic Regression. Phương pháp này giúp tự động thu nhỏ các hệ số hồi quy của các biến kém quan trọng về 0, qua đó kiểm soát hiện tượng đa cộng tuyến và ngăn chặn việc mô hình bị quá khớp.

Bảng 1. Mô tả các biến trong mô hình

Ký hiệu	Diễn giải	Cách tính	Nghiên cứu đã sử dụng
Các biến thu thập từ dữ liệu kế toán			
CR	Khả năng thanh toán hiện hành	Tài sản ngắn hạn/ Nợ ngắn hạn	Ohlson (1980), Muttia & Sutrisno (2022)
QR	Khả năng thanh toán nhanh	(Tài sản ngắn hạn -hàng tồn kho) /Nợ ngắn hạn	Phan Trần Trung Dũng et al. (2022)
D/A	Tỷ lệ nợ trên tổng tài sản	Tổng nợ phải trả/ Tổng tài sản	Phan Trần Trung Dũng et al. (2022), Muttia & Sutrisno (2022)
ICR	Lợi nhuận trước thuế và lãi vay/ Chi phí lãi vay	(Lợi nhuận từ hoạt động kinh doanh + chi phí lãi vay)/ Chi phí lãi vay	Altman (1968)
ROA	Tỷ suất sinh lợi trên tổng tài sản	Lợi nhuận sau thuế/ Tổng tài sản bình quân	Hoàng Thị Hồng Vân (2020), Phan Trần Trung Dũng et al. (2022)

Ký hiệu	Diễn giải	Cách tính	Nghiên cứu đã sử dụng
Các biến thu thập từ dữ liệu kế toán			
NPM	Biên lợi nhuận ròng	Lợi nhuận sau thuế/ Doanh thu	Muttia & Sutrisno (2022)
OCF/TA	Tỷ suất dòng tiền hoạt động trên tổng tài sản	Tổng dòng tiền thuần hoạt động kinh doanh/ Tổng tài sản bình quân	(Lâm Thị Thanh Huyền & Nguyễn Thị Quỳnh Châm, 2024)
WC/TA	Vốn lưu động trên tổng tài sản	(Tài sản ngắn hạn- nợ ngắn hạn) / Tổng tài sản	Altman (1968)
TAT	Vòng quay tài sản	Doanh thu thuần/ Tổng tài sản bình quân	Altman (1968),
Các biến thu thập từ dữ liệu thị trường			
P/B	Tỷ số giá thị trường trên giá trị sổ sách	Tỷ số giá thị trường trên mỗi cổ phiếu/ giá trị sổ sách trên mỗi cổ phiếu	Phan Trần Trung Dũng et al. (2022)
VOL	Tỷ suất quay vòng cổ phiếu	Số cổ phiếu giao dịch trong kỳ / Số cổ phiếu lưu hành bình quân	L. Li & Faff (2019)
MCAP	Giá trị vốn hóa thị trường	Ln (Giá cổ phiếu hiện tại x Số cổ phiếu đang lưu hành)	Shumway (2001)
AR	Tỷ suất sinh lợi bất thường	Lợi nhuận thực tế - Lợi nhuận kỳ vọng	Shumway (2001)

Nguồn: Tổng hợp của nhóm tác giả

Nghiên cứu tiếp cận theo phương pháp học máy. Đầu tiên, dữ liệu được tiền xử lý, áp dụng phép biến đổi Winsor (thay thế các giá trị ngoại lai bằng các giá trị tại 2 ngưỡng phân vị 1% và 99%). Tiếp theo, dữ liệu được chia thành 2 tập dữ liệu gồm tập huấn luyện (80% dữ liệu) được sử dụng để xây dựng mô hình và tập kiểm tra (20% dữ liệu) cho kiểm định. Ngoài ra, để tránh hiện tượng quá khớp và tìm ra bộ tham số ổn định nhất, phương pháp kiểm chứng chéo lặp K-Fold được áp dụng bằng cách chia tập huấn luyện thành 5 phần để luân phiên huấn luyện và đánh giá qua tổng cộng 15 vòng lặp. Sau đó, nghiên cứu thực hiện chuẩn hóa dữ liệu và áp dụng thuật toán mô hình Penalized Logistic Regression, đồng thời thực hiện tối ưu hóa siêu tham số để chọn ra mô hình có khả năng dự báo tốt nhất.

4. Kết quả và thảo luận

Nghiên cứu đã tiến hành kiểm định hiện tượng

đa cộng tuyến dựa vào hệ số phóng đại phương sai VIF. Kết quả kiểm định cho thấy, biến QR và CR có giá trị VIF >10. Điều này cho thấy, có hiện tượng đa cộng tuyến xảy ra trong mô hình. Do đó, để khắc phục hiện tượng này, biến QR được loại bỏ ra khỏi mô hình.

Nghiên cứu vận dụng học máy trên R với thuật toán Penalized Logistic Regression để dự báo RRPS. Bảng 2 trình bày chi tiết kết quả từng mô hình, tương ứng với các nguồn dữ liệu và độ trễ thời gian 1 năm (t-1), 2 năm (t-2) và 3 năm (t-3) trước khi sự kiện công ty gặp RRPS xảy ra.

Kết quả nghiên cứu từ Bảng 2 cho thấy, DLKT có khả năng dự báo tốt RRPS với cả 3 năm quan sát. Độ chính xác dự báo trên tập kiểm tra DLKT của năm t-1, t-2 và t-3 lần lượt là 87,5%; 90% và 77,5%. Trong khi đó, chỉ số AUC đo lường khả năng phân biệt của mô hình giữa 2 nhóm đối với DLKT của

Bảng 2. Tóm tắt kết quả dự báo RRPS năm t của các mô hình

Dữ liệu năm	Nguồn dữ liệu	Độ chính xác (Accuracy) (%)	Chỉ số AUC (Area Under the Curve) (%)
t-1	Kế toán (mô hình 1)	87,5	95,25
	Thị trường (mô hình 2)	55,0	61,75
	Kết hợp (mô hình 3)	87,5	93,75
t-2	Kế toán (mô hình 1)	90,0	97,75
	Thị trường (mô hình 2)	55,0	59,25
	Kết hợp (mô hình 3)	87,5	97,25
t-3	Kế toán (mô hình 1)	77,5	89,00
	Thị trường (mô hình 2)	65,0	69,00
	Kết hợp (mô hình 3)	77,5	89,00

năm t-1, t-2 và t-3 lần lượt là 95,25%; 97,75%; 89%. Kết quả này tương đồng với kết quả nghiên cứu của Lâm Thị Thanh Huyền & Nguyễn Thị Quỳnh Châm (2024), Correia (2025). Giả thuyết H1 được chấp thuận.

Bên cạnh đó, kết quả nghiên cứu cho thấy, DLTT thể hiện khả năng dự báo RRPS kém hơn rất nhiều. Chỉ số AUC chỉ hơn mức ngẫu nhiên một chút (61,75%; 59,25%; 69%) chứng tỏ DLTT chưa phân biệt tốt công ty RRPS. Kết quả này trái ngược phát hiện của Saad et al. (2025), Correia (2025), Cao et al. (2024) DLTT có khả năng dự báo tốt RRPS. Giả thuyết H2 bị bác bỏ.

Hơn nữa, đối với mô hình kết hợp DLKT và DLTT, kết quả nghiên cứu cho thấy hiệu suất dự báo không được cải thiện, chỉ số AUC còn giảm nhẹ so với việc chỉ dùng DLKT trong ngắn hạn (năm t-1 và t-2). Điều này cho thấy, trong bối cảnh giai đoạn

nghiên cứu, các biến thị trường không mang lại thêm giá trị thông tin so với thông tin báo cáo tài chính đã phản ánh. Kết quả nghiên cứu này trái ngược so với Saad et al. (2025), Correia (2025). Giả thuyết H3 bị bác bỏ.

5. Kết luận

Kết quả nghiên cứu chỉ ra, trong mẫu dữ liệu giai đoạn 2016 - 2024, mô hình dự báo RRPS sử dụng DLKT có khả năng dự báo cao nhất. Mô hình DLTT có kết quả kém hơn. Đối với mô hình kết hợp, DLTT không ảnh hưởng đáng kể tăng thêm năng lực dự báo RRPS của mô hình. Như vậy, nghiên cứu đã xây dựng mô

hình dự báo RRPS của các công ty niêm yết trong bối cảnh gần nhất của Việt Nam, góp phần tạo thêm bằng chứng đối với nguồn dữ liệu tại thị trường Việt Nam.

Về đóng góp thực tiễn, kết quả nghiên cứu có thể hỗ trợ nhà đầu tư và các tổ chức tín dụng sử dụng mô hình dự báo RRPS đánh giá rủi ro của các công ty. Họ nên sử dụng mô hình như một công cụ hỗ trợ ra quyết định kết hợp với các phân tích tài chính chuyên môn, giúp giảm thiểu rủi ro và nâng cao hiệu quả trong hoạt động đầu tư và cấp tín dụng.

Tuy nhiên, nghiên cứu vẫn có những hạn chế nhất định. Nghiên cứu tập trung vào việc xây dựng và kiểm định mô hình theo cách chọn mẫu theo cặp. Cách chọn mẫu này chưa phản ánh đúng tổng thể do trong thực tế tỷ lệ các công ty có RRPS nhỏ hơn các công ty có tình hình tài chính lành mạnh. Nhóm tác giả sẽ khắc phục trong các nghiên cứu tương lai ■

TÀI LIỆU THAM KHẢO:

Cục Thống kê (2026). Báo cáo tình hình kinh tế - xã hội quý IV và năm 2025. Truy cập tại <https://www.nso.gov.vn/du-lieu-va-so-lieu-thong-ke/2026/01/bao-cao-tinh-hinh-kinh-te-xa-hoi-quy-iv-va-nam-2025/>.
Hoàng Thị Hồng Vân (2020). Vận dụng mô hình Z-score trong dự báo khả năng phá sản doanh nghiệp tại Việt Nam. Tạp chí Khoa học và Đào tạo Học viện Ngân hàng, (217), 43-51.

Lâm Thị Thanh Huyền, & Nguyễn Thị Quỳnh Châm. (2024). Dự báo rủi ro phá sản của các doanh nghiệp bất động sản niêm yết trên thị trường chứng khoán Việt Nam thông qua mô hình hồi quy Logistic. *Tạp chí Nghiên cứu Tài chính Kế toán*, 260(2), 80-84.

Phan Trần Trung Dũng, Nguyễn Thị Hà Thanh, Võ Minh Thu. (2022). Nghiên cứu nhân tố tác động đến rủi ro phá sản của các doanh nghiệp bất động sản niêm yết tại Việt Nam. *Tạp chí Quản lý và Kinh tế Quốc tế*, (147).

Nguyễn Minh Nhật, Ngô Hoàng Khánh Duy (2024). Dự báo khả năng vỡ nợ của doanh nghiệp nhỏ và vừa tại Việt Nam: Nghiên cứu trên các mô hình học máy. *Tạp chí Kinh tế - Luật và Ngân hàng*, 266(7), 51-64. Truy cập tại <https://doi.org/10.59276/JELB.2024.07CD.2762>.

Altman E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589-609. <https://doi.org/10.1111/j.1540-6261.1968.tb00843.x>.

Barboza F., Kimura H., & Altman E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83, 405-417. <https://doi.org/10.1016/j.eswa.2017.04.006>.

Cao Y., Luo Y., Wei P., Zhai J., & Shi S. (2024). Bankruptcy forecasting market information with ensemble model. *The British Accounting Review*, 101530. <https://doi.org/10.1016/j.bar.2024.101530>.

Correia M. (2025). Accounting and corporate failure: the evolving role of accounting information in bankruptcy prediction. *Accounting and Business Research*, 55(5), 510-537. <https://doi.org/10.1080/00014788.2025.2516273>.

Hillegeist S. A., Keating E. K., Cram D. P., & Lundstedt K. G. (2004). Assessing the Probability of Bankruptcy. *Review of Accounting Studies*, 9(1), 5-34. <https://doi.org/10.1023/B:RAST.0000013627.90884.b7>.

Li L., & Faff R. (2019). Predicting corporate bankruptcy: What matters? *International Review of Economics & Finance*, 62, 1-19. <https://doi.org/10.1016/j.iref.2019.02.016>.

Li M. Y. L., & Miu P. (2010). A hybrid bankruptcy prediction model with dynamic loadings on accounting-ratio-based and market-based information: A binary quantile regression approach. *Journal of Empirical Finance*, 17(4), 818-833. <https://doi.org/None>.

Lotfi B., Bahri Sales J., Jabbarzadeh Kangarlouei S., & Heydari M. (2024). Comparison of the combination model with the structural and accounting model in predicting the financial distress. *International Journal of Nonlinear Analysis and Applications*, 14(1), 277-290. <https://doi.org/http://dx.doi.org/10.22075/ijnaa.2023.30444.4405>.

Merton R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance*, 29(2), 449. <https://doi.org/10.2307/2978814>.

Muttia E. E., & Sutrisno (2022). Financial Performance and Prediction of Financial Distress in Food and Beverage Companies Listed on the Indonesia Stock Exchange. *International Journal of Economics, Business and Management Research*, 6(5), 154-167. <https://doi.org/10.51505/IJEBMR.2022.6513>.

Nouri B. A., & Soltani M. (2016). Designing a Bankruptcy Prediction Model Based on Account, Market and Macroeconomic Variables (Case Study: Cyprus Stock Exchange). *Iranian Journal of Management Studies*, 9(1), 125-147. <https://doi.org/10.22059/ijms.2016.55038>.

Ohlson J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), 109. <https://doi.org/10.2307/2490395>

Pham Vo Ninh B., Do Thanh T., & Vo Hong D. (2018). Financial distress and bankruptcy prediction: An appropriate model for listed firms in Vietnam. *Economic Systems*, 42(4), 616-624. <https://doi.org/10.1016/j.ecosys.2018.05.002>.

Ruxanda G., Zamfir C., & Muraru A. (2018). predicting financial distress for Romanian companies. *Technological and Economic Development of Economy*, 24(6), 2318-2337. <https://doi.org/10.3846/tede.2018.6736>.

Saad M., Hussain R. Y., & Irshad H. (2025). Step towards a general bankruptcy prediction model for listed non-financial firms: The use of accounting, market, and macroeconomic variables. *Studia Historiae Oeconomicae*, 43(1), 209-240. <https://doi.org/10.14746/sho.2025.43.1.009>.

Shumway T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business*, 74(1), 101-124. <https://doi.org/10.1086/209665>

Vo D. H., Pham B. N. V., Ho C. M., & McAleer M. (2019). Corporate Financial Distress of Industry Level Listings in Vietnam. *Journal of Risk and Financial Management*, 12(4), 155. <https://doi.org/10.3390/jrfm12040155>.

Waqas H., Md-Rus R., & Elgammal M. M. (2018). Predicting financial distress: Importance of accounting and firm-specific market variables for Pakistan's listed firms. *Cogent Economics & Finance*, 6(1). <https://doi.org/10.1080/23322039.2018.1545739>.

Ngày nhận bài: 11/03/2026

Ngày phản biện đánh giá và sửa chữa: 20/03/2026

Ngày chấp nhận đăng bài: 05/04/2026

FORECASTING BANKRUPTCY RISK FOR VIETNAMESE LISTED COMPANIES USING ACCOUNTING AND MARKET DATA

● LE THI HAI BINH¹

● LE THI MY NGUYET¹

● LE THI HUYEN TRANG²

● NGUYEN THI NHU NGOC²

● NGUYEN THI ANH DUONG²

● MAI NGOC THAO LY²

● VO NGUYEN ANH THU²

¹Lecturer, School of Finance And Accounting,
Industrial University of Ho Chi Minh City

²Student, School of Finance And Accounting,
Industrial University of Ho Chi Minh City

ABSTRACT:

This study develops a bankruptcy risk prediction model for listed companies in Vietnam using accounting and market data over the period 2016–2024. Based on a sample of 103 firms, the analysis employs a machine learning approach with logistic regression to evaluate predictive performance across different data specifications. The results indicate that the model based on accounting data achieves the highest predictive accuracy, with AUC values of 92.5%, 97.75%, and 89%, while the market-based model demonstrates comparatively weaker performance. Moreover, combining accounting and market data does not enhance predictive power, suggesting that market variables provide limited incremental information beyond accounting indicators. Overall, the findings underscore the dominant role of accounting information in forecasting bankruptcy risk in the Vietnamese context.

Keywords: bankruptcy risk, financial distress, accounting data, market data, Logistic regression model.