



ISSN 3093 - 3390

TẠP CHÍ

Kinh tế - Tài chính

CƠ QUAN CỦA BỘ TÀI CHÍNH

Kỳ 2 tháng 8/2026 (Số 32)



TĂNG TRƯỞNG KINH TẾ CỦA VIỆT NAM: CƠ HỘI VÀ THÁCH THỨC



<https://tapchikinhthetataichinh.vn>

THÔNG TIN LÝ LUẬN VÀ NGHIỆP VỤ KINH TẾ - TÀI CHÍNH

GIỚI THIỆU, CÔNG BỐ KẾT QUẢ NGHIÊN CỨU KHOA HỌC

BAN BIÊN TẬP

Tổng Biên tập:

TS. PHẠM THU PHONG

Phó Tổng Biên tập:

BÙI ĐỨC HẢI

TS. VŨ THỊ ÁNH HỒNG

ThS., NGUYỄN HẢI HỒNG

ThS., BÙI THỊ THU HƯƠNG

HỘI ĐỒNG BIÊN TẬP

GS. TS. HOÀNG VĂN CƯỜNG

GS. TS. PHẠM HỒNG CHƯƠNG

GS. TS. TRẦN THỌ ĐẠT

GS. TS. NGÔ THẮNG LỢI

GS. TS. NGUYỄN CÔNG NGHIỆP

GS. TS. SỬ ĐÌNH THÀNH

PGS. TS BÙI HUY NHƯỠNG

PGS. TS. TRẦN TRỌNG NGUYỄN

PGS. TS. NGUYỄN HỒNG SƠN

PGS.TS. NGUYỄN ĐÀO TÙNG

MỤC LỤC

TĂNG TRƯỞNG KINH TẾ CỦA VIỆT NAM: CƠ HỘI VÀ THÁCH THỨC

- 7 Việt Nam với mục tiêu tăng trưởng 2 con số: Cơ hội và thách thức Lê Thị Trường An
11 Vai trò của chính sách lãi suất đối với tăng trưởng giai đoạn 2001-2025 và hàm ý
cho giai đoạn 2026-2030 Nguyễn Thạc Hoát
19 Đầu tư công, tăng trưởng kinh tế và chuyển dịch cơ cấu ngành: Bằng chứng so sánh
từ dữ liệu cấp tỉnh tại các địa phương Lê Công Hoàng, Phạm Thị Tường Vân, Trần Văn Trí
26 Già hóa dân số: Cơ hội và thách thức đối với tăng trưởng kinh tế Việt Nam
Lê Thị Hồng Dương
29 Tác động của sai lệch tỷ giá với tăng trưởng kinh tế của Việt Nam
Nguyễn Thị Tinh Nhi, Phạm Thị Khánh Hòa, Hoàng Mai Chi, Lê Khánh Linh, Nguyễn Thị Nguyệt
34 Phát triển kinh tế tư nhân trở thành động lực quan trọng nhất của nền kinh tế:
Những vấn đề đặt ra Hà Thị Hằng

TÀI CHÍNH - KINH DOANH

- 38 Xây dựng cơ chế tài chính hỗn hợp để thúc đẩy hình thành trung tâm tài chính quốc tế
ở Việt Nam Vũ Thanh Nguyên, Lương Văn Hà
43 Giải pháp thúc đẩy chuyển đổi số ngành ngân hàng ở Việt Nam Đỗ Diệu Linh
47 Nghịch lý số hóa trong quản trị tín dụng ngân hàng: Vai trò trung gian chuỗi của áp
lực KPI và kiệt sức nghề nghiệp Tôn Thất Viên
52 Động lực biến động giá cổ phiếu ngân hàng dẫn dắt tại Việt Nam: Tiếp cận bằng
mô hình học máy Dương Thị Hải Phương, Trần Minh Thy
57 Ứng dụng trí tuệ nhân tạo trong lĩnh vực công nghệ tài chính tại Việt Nam:
Thực trạng và khuyến nghị Trần Thế Sao, Phan Hồng Hạnh
61 Tác động của tự do kinh tế và cạnh tranh đến các ngân hàng thương mại cổ phần
Việt Nam Đặng Bửu Kiếm, Nguyễn Thị Thanh Hòa
65 Dự báo khoảng chuỗi thời gian tài chính tại thị trường mới nổi: Tiếp cận bằng
ARIMA mở dựa trên biến động trong ngày Lê Thanh Hoa
69 Nhận diện sớm nguy cơ khủng hoảng tài chính của doanh nghiệp niêm yết ngành
xây dựng: Ứng dụng mô hình Z-score và phương pháp Beaver Nguyễn Thị Thanh Thủy
72 Ứng dụng mô hình ARIMA trong dự báo ROE: Trường hợp Công ty Cổ phần Xây dựng
COTECCONS Nguyễn Thị Ngọc Thu
78 Rủi ro của nhà đầu tư trong thị trường tài sản số: Phân tích đa chiều và khoảng trống
bảo vệ tại Việt Nam Trịnh Quang Long
84 Nhân tố ảnh hưởng đến tăng trưởng tín dụng tại các ngân hàng thương mại Việt Nam
Đàm Thị Thảo
89 Nghịch lý cạnh tranh cho vay khách hàng cá nhân trong thời đại số: Bằng chứng từ
nhận thức của nhân viên ngân hàng và khách hàng vay vốn tại Việt Nam Phan Ngọc Minh
94 Nhân tố ảnh hưởng đến quyết định phát triển chiến lược kinh doanh của doanh nghiệp
Lê Mỹ Hà
99 Các yếu tố tài chính và phi tài chính ảnh hưởng đến rủi ro tín dụng trong cho vay
khách hàng tại các ngân hàng thương mại Việt Nam
Lê Hoàng Vinh, Nguyễn Như Tuyết Trinh
106 Thu nhập ngoài lãi và xu hướng đa dạng hóa nguồn thu tại các ngân hàng thương
mại Việt Nam giai đoạn 2014-2025 Phạm Xuân Dũng, Phạm Thị Hoàng Anh
111 Chuyển đổi số trong quản trị tài chính doanh nghiệp ở Việt Nam: Thực trạng và
giải pháp Phạm Thị Mai Huyền
116 Môi trường đa văn hóa và hiệu quả nhân viên: Đề xuất khung phân tích cho ngành
bảo hiểm nhân thọ tại Việt Nam Đoàn Văn Tuấn, Trần Tuấn Minh
121 Rủi ro tín dụng tại ngân hàng thương mại Việt Nam: Bằng chứng thực nghiệm từ
các nhân tố đặc thù ngân hàng Nguyễn Phan Yến Phương, Phạm Thị Hoàng Anh
125 Năng lực phân tích kinh doanh và hiệu quả quản trị tại các doanh nghiệp sản xuất
ở Việt Nam Vũ Đình Dũng, Phan Tiến Dũng, Trần Mạnh Kiên
128 Đề xuất mô hình nghiên cứu các yếu tố ảnh hưởng tới cấu trúc vốn của doanh nghiệp
Nguyễn Linh Quỳnh Ngân
133 Tác động của hợp tác Fintech tới chuyển đổi số hoạt động ngân hàng bán lẻ ở Việt Nam:
Nghiên cứu trường hợp điển hình tại BIDV Phan Thị Thanh Nhân, Phan Thanh Đức
137 Tác động của sở hữu tổ chức đến hiệu quả hoạt động doanh nghiệp: Bằng chứng
thực nghiệm từ các công ty bất động sản niêm yết tại Việt Nam
Nguyễn Minh Ánh, Lê Thị Cẩm Tú, Trần Thị Thu Hằng, Nguyễn Phương Thảo, Đào Thị Trang, Đỗ Văn Anh
143 Phân tích hiệu quả tài chính các doanh nghiệp điện tái tạo tại Việt Nam: Hướng
tiếp cận CRITIC-TOPSIS Nguyễn Minh Nhật
147 Tác động của mức độ quan tâm công chúng đến biến động giá carbon: Bằng chứng
thực nghiệm từ dữ liệu Google Trends trong thị trường EU ETS Bùi Quang Huy
151 Biến động chính sách và tác động đối với thị trường tiền kỹ thuật số tại Việt Nam:
Bằng chứng từ mô hình EGARCH Lê Thị Ngọc Phương
155 Ảnh hưởng của quản trị doanh nghiệp đến tính thích hợp của thông tin báo cáo tài chính:
Bằng chứng thực nghiệm tại Việt Nam Nguyễn Thị Minh Hằng, Dương Nguyễn Bảo Trân,
Nguyễn Mai Loan, Nguyễn Thị Hải Yến, Võ Trọng Khoa, Lê Hoàn Thiên Bảo



ĐỘNG LỰC BIẾN ĐỘNG GIÁ CỔ PHIẾU NGÂN HÀNG DẪN ĐẦU TẠI VIỆT NAM: TIẾP CẬN BẰNG MÔ HÌNH HỌC MÁY

DƯƠNG THỊ HẢI PHƯƠNG¹, TRẦN MINH THY²

Tóm tắt

Nghiên cứu giải mã động lực xu hướng giá ngắn hạn của 3 cổ phiếu ngân hàng dẫn đầu (VCB, TCB, MBB) giai đoạn 2021 - 2025. Bài toán phân loại nhị phân được giải quyết bằng thuật toán Random Forest, kết hợp chiến lược phân chia dữ liệu tuyến tính và kiểm thử cuốn chiếu nhằm triệt tiêu rò rỉ thông tin. Đóng góp chính là việc ứng dụng học máy giải thích được (XAI) qua giá trị SHAP để bóc tách hành vi vi mô. Kết quả thực chứng ghi nhận hiệu năng dự báo suy giảm trên tập kiểm thử 2025, củng cố thuyết thị trường hiệu quả (EMH) dạng yếu. Phân tích SHAP bộc lộ sự phân hóa cấu trúc: cổ phiếu quốc doanh (VCB) mang tính phòng thủ và vận động theo RSI; khối tư nhân (TCB, MBB) nhạy cảm với thanh khoản đầu cơ nhưng chịu sự chi phối nền tảng bởi xu hướng trung hạn (SMA_50, MACD). Qua đó, nghiên cứu đề xuất các hàm ý quản trị danh mục và minh bạch thông tin vĩ mô.

Từ khóa: Dự báo giá cổ phiếu, ngân hàng thương mại, random forest, SHAP values, học máy giải thích được (XAI)

PRICE VOLATILITY DYNAMICS OF MARKET-LEADING BANKING STOCKS IN VIET NAM: A MACHINE LEARNING APPROACH

Abstract

This study examines the short-term price trend dynamics of three market-leading banking stocks (VCB, TCB, MBB) in the Vietnamese stock market from 2021 to 2025. Formulated as a binary classification task, the problem is addressed using the Random Forest algorithm in conjunction with a temporal split strategy and walk-forward validation to mitigate data leakage. Crucially, Explainable AI (XAI) via SHAP values is applied to dissect the micro-behavior of the models. Empirical results reveal a decline in forecasting performance on the 2025 test set, robustly supporting the Weak-form Efficient Market Hypothesis (EMH). Notably, SHAP analysis uncovers a profound structural divergence: the state-owned stock (VCB) exhibits a defensive nature driven by the RSI; whereas private banks (TCB, MBB) are highly sensitive to speculative liquidity, yet ultimately governed by underlying mid-term trends (SMA_50, MACD). Consequently, the study proposes significant implications for portfolio management and macroeconomic information transparency.

Keywords: Stock price forecasting, commercial banks, random forest, SHAP values, Explainable AI (XAI)

Ngày nhận bài: 18/6/2026; Ngày hoàn thiện biên tập: 28/7/2026; Ngày duyệt đăng: 8/8/2026

GIỚI THIỆU

Trong bối cảnh tài chính định lượng hiện đại, dự báo xu hướng biến động giá tài sản tài chính luôn là một bài toán đầy thách thức. Thị trường chứng khoán không vận hành như một cỗ máy cơ học tuyến tính đơn giản mà mang đặc trưng của một hệ thống thích nghi phức tạp, nơi hành vi của các tác nhân tương tác liên tục với thông tin vĩ mô, chính sách tiền tệ và tâm lý đám đông (Chen và cộng sự, 2023). Tại thị trường Việt Nam, nhóm cổ phiếu ngân hàng đóng vai trò là kênh phân bổ vốn chủ đạo của nền kinh tế, dẫn dắt dòng tiền và phản ánh trực tiếp sức khỏe tài chính quốc gia. Mặc dù thuyết thị trường hiệu quả cho rằng

mọi thông tin đều được phản ánh tức thời vào giá, thực tế tại các thị trường đang phát triển cho thấy các yếu tố tâm lý hành vi, độ trễ truyền dẫn chính sách và nút thắt thanh khoản thường tạo ra các mẫu hình động lượng có tính chu kỳ (Deniz và cộng sự, 2020). Do đó, dữ liệu tài chính tần suất ngày luôn bị chi phối bởi các yếu tố ngẫu nhiên khiến tỷ số tín hiệu trên nhiễu cực kỳ thấp, đặt ra yêu cầu cấp thiết về một hệ thống cảnh báo sớm bằng học máy phi tuyến.

Để giải quyết bài toán nhiễu động dữ liệu, nghiên cứu này tiếp cận việc dự báo dưới dạng phân loại nhị phân thay vì mô hình hồi quy truyền thống. Dưới góc độ thực tiễn của phân tích định lượng, việc nhận diện

¹ TS., Trường Đại học Kinh tế, Đại học Huế; Email: dthphuong@hce.edu.vn

² Email: C25035002@hce.edu.vn



chính xác hướng dịch chuyển chủ đạo của giá (tăng hoặc giảm) giúp triệt tiêu các sai số bình phương trung bình lớn do các điểm cực trị ngẫu nhiên gây ra. Phương pháp phân loại này tập trung nhận diện các mẫu hình động lượng, hoàn toàn phù hợp với đặc thù chu kỳ thanh toán T+2 hiện nay, từ đó mang lại độ khả thi và tính ổn định cao hơn cho các quyết định phòng hộ rủi ro danh mục. Mục tiêu chính của nghiên cứu là xây dựng và đánh giá năng lực dự báo xu hướng giá ngắn hạn của thuật toán Random Forest (Rừng ngẫu nhiên). Đây là một thuật toán học máy thuộc nhóm học có giám sát được sử dụng phổ biến để giải quyết các bài toán phân loại và hồi quy. Thuật toán hoạt động bằng cách xây dựng nhiều cây quyết định và kết hợp kết quả của các cây để đưa ra dự đoán cuối cùng, giúp tăng độ chính xác và giảm hiện tượng quá khớp. Thuật toán này áp dụng đối với nhóm cổ phiếu ngân hàng dẫn dắt (VCB, TCB, MBB) trong giai đoạn biến động 2021-2025.

Điểm khác biệt và đóng góp chính của bài báo là việc ứng dụng học máy giải thích được (Explainable AI, XAI) thông qua giá trị SHAP (Shapley Additive exPlanations) nhằm minh bạch hóa cấu trúc hợp đen của các mô hình dự báo truyền thống. Thông qua lăng kính XAI, nghiên cứu sẽ cung cấp một phân tích so sánh trực diện nhằm bóc tách hành vi vi mô của các chỉ báo kỹ thuật đối với hai mô hình hoạt động khác biệt: khối ngân hàng quốc doanh (đại diện bởi VCB mang tính chất phòng thủ) và khối ngân hàng tư nhân (đại diện bởi TCB, MBB có tính nhạy cảm cao về thanh khoản) (Trần Thị Mỹ Linh, 2025). Các phát hiện này sẽ trực tiếp định lượng ảnh hưởng của từng loại chỉ báo, cung cấp cơ sở khoa học cho các nhà đầu tư và cơ quan quản lý trên thị trường chứng khoán.

CƠ SỞ LÝ THUYẾT VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Cơ sở lý thuyết và khung phân tích

Trong nhiều thập kỷ, dự báo chuỗi thời gian tài chính chủ yếu dựa vào các cấu trúc kinh tế lượng tham số. Các mô hình kinh điển như tự hồi quy tích hợp trung bình trượt (ARIMA) do Box & Jenkins (1970) phát triển, hay phương sai sai số thay đổi có điều kiện tổng quát (GARCH) do Bollerslev (1986) giới thiệu, đã tạo nền tảng vững chắc trong định giá và kiểm soát rủi ro. Tuy nhiên, các mô hình này bộc lộ điểm yếu khi phụ thuộc lớn vào giả định chuỗi lợi suất tuân theo phân phối chuẩn. Trong thực tế, biến động giá cổ phiếu thường xuyên vi phạm giả định này, thể hiện rõ đặc tính phân phối đuôi dày (Fat-tailed) với sự xuất hiện của các sự kiện cực đoan nằm ngoài phân phối chuẩn (Cont, 2001).

Để khắc phục giới hạn tuyến tính, các thuật toán học máy đã được ứng dụng rộng rãi. Điển hình là

thuật toán Random Forest do Breiman (2001) đề xuất, sử dụng cơ chế học kết hợp giúp giảm thiểu phương sai và chống lại hiện tượng quá khớp hiệu quả trước dữ liệu nhiễu cao. Dù đem lại độ chính xác cao, tính chất hộp đen của học máy lại cản trở khả năng diễn giải hành vi vi mô trong tài chính. Do đó, lĩnh vực học máy giải thích được (XAI) đã ra đời, nổi bật là phương pháp SHAP (Shapley Additive exPlanations) do (Lundberg & Lee, 2017) phát triển dựa trên lý thuyết trò chơi hợp tác (Winter, 2002). SHAP cho phép tính toán mức độ đóng góp biên của từng biến số vào kết quả dự báo, biến hộp đen thành một công cụ phân tích trong suốt.

Trên cơ sở kế thừa các lý thuyết trên, nghiên cứu nhận thấy một khoảng trống đáng kể trong việc ứng dụng XAI để đối chiếu cấu trúc hành vi vi mô giữa các mô hình kinh doanh ngân hàng khác nhau tại Việt Nam vẫn còn hạn chế. Từ đó, khung phân tích được thiết lập theo quy trình chuỗi thời gian nghiêm ngặt: Dữ liệu thô (OHLCV) được tiền xử lý và trích xuất đặc trưng thành các chỉ báo kỹ thuật; sau đó áp dụng phân chia tuyến tính để huấn luyện mô hình Random Forest độc lập cho từng mã; và cuối cùng sử dụng SHAP để bóc tách động lực dẫn dắt (RSI, thanh khoản, dải Bollinger) riêng biệt giữa khối ngân hàng quốc doanh và tư nhân.

Phương pháp nghiên cứu

Mô hình định lượng và lựa chọn biến

Bài toán được định nghĩa dưới dạng phân loại nhị phân nhằm dự báo hướng dịch chuyển của giá (tăng/giảm), giúp triệt tiêu sai số cực trị của phương pháp hồi quy. Cụ thể, để đảm bảo tính thực tiễn trong giao dịch, biến mục tiêu phân loại nhị phân Y_t được gán nhãn dựa trên hàm biến đổi của tỷ suất sinh lời liền sau. Lớp mục tiêu được xác định là $Y_t = 1$ (lớp tăng) nếu tỷ suất sinh lời $R_{t+1} > 0$ và $Y_t = 0$ (lớp giảm/đứng) nếu $R_{t+1} \leq 0$. Việc thiết lập ngưỡng phân tách này giúp thuật toán tập trung nhận diện các xung lượng định hướng thực tế thay vì dự báo biến động nhiễu. Hệ thống đặc trưng được xây dựng đại diện cho bốn khía cạnh động lực học của thị trường: xu hướng, động lượng, biến động và thanh khoản. Thay vì sử dụng dữ liệu thô, nghiên cứu trích xuất 7 nhóm chỉ báo kỹ thuật bao gồm: đường trung bình động đơn giản (SMA_20, SMA_50), chỉ số sức mạnh tương đối (RSI Wilder chu kỳ 14), trung bình động hội tụ phân kỳ (MACD và Signal), độ rộng dải Bollinger (BB_Width chu kỳ 20), tỷ lệ khối lượng giao dịch (Volume_MA_Ratio), tỷ suất sinh lời hàng ngày (Daily Return), và quán tính giá (Momentum 10 phiên).

Thuật toán Random Forest được ứng dụng làm nền tảng thuật toán trọng tâm để giải quyết bài toán phân loại nhị phân. Để giải thích mô hình, giá trị SHAP cục bộ của đặc trưng i được đo lường thông



qua trung bình đóng góp biên trên tất cả các tổ hợp đặc trưng S có thể xảy ra theo công thức (1):

$$\Phi_i = \sum_{S \in N \setminus \{i\}} \frac{|S|! (M - |S| - 1)!}{M!} [f_x(S \cup \{i\}) - f_x(S)] \quad (1)$$

Trong đó, M là tổng số lượng đặc trưng và f_x là hàm dự báo của hệ thống.

Phương pháp thu thập dữ liệu và phạm vi nghiên cứu

Nghiên cứu thu thập dữ liệu lịch sử tần suất ngày của ba mã cổ phiếu dẫn dắt (VCB, TCB, MBB) từ thư viện dữ liệu tài chính vnstock3 (HOSE) giai đoạn 2021-2025. Quy trình tiền xử lý được thực hiện nghiêm ngặt bằng phương pháp điền ngược để xử lý giá trị thiếu, loại bỏ các ngày phi giao dịch, và áp dụng kỹ thuật Winsorization tại phân vị 1% và 99% nhằm kiểm soát dữ liệu cực đoan mà không làm mất tính chu kỳ. Bảng 1 mô tả cấu trúc thống kê của tập dữ liệu.

Phân tích dữ liệu cho thấy sự phân hóa lớn về biên độ dao động: VCB (đại diện quốc doanh) có hệ số biến thiên thấp nhất (chỉ 10,3%), minh chứng cho vai trò nền tảng phòng thủ; trong khi đó, TCB (đại diện tư nhân) sở hữu độ biến động cực lớn với hệ số biến thiên lên tới 31,6% phản ánh sự nhạy cảm cao trước các cú sốc thanh khoản trên thị trường.

Đặc biệt, để đảm bảo tính khoa học và ngăn chặn hiện tượng rò rỉ dữ liệu tương lai, dữ liệu được phân chia tuyến tính thành: tập huấn luyện (2021-2023), tập xác thực (2024) và tập kiểm thử (2025). Đồng thời, kỹ thuật kiểm thử cuốn chiếu với 5 nếp gấp được áp dụng để mô phỏng chính xác quá trình học liên tục của thuật toán trong môi trường thực tiễn.

KẾT QUẢ VÀ THẢO LUẬN

Kết quả nghiên cứu

Năng lực dự báo xu hướng ngắn hạn của thuật toán Random Forest trên tập dữ liệu kiểm thử độc lập được đánh giá toàn diện thông qua các tiêu chí đo lường sự cân bằng của phân loại nhị phân: Precision, Recall, F1-Score và đặc biệt là chỉ số ROC-AUC. Bảng 2 tóm tắt hiệu suất phân loại của ba mã cổ phiếu nghiên cứu.

Kết quả thống kê từ Bảng 2 cho thấy hiệu năng dự báo trên tập dữ liệu tương lai (2025) suy giảm đáng kể so với tập huấn luyện, tiến sát về biên độ của một hệ thống phân phối ngẫu nhiên (ROC-AUC dao động từ 0,4138 đến 0,5067). Hiện tượng này cung cấp minh chứng thực nghiệm vững chắc ủng hộ thuyết thị trường hiệu quả (EMH) dạng yếu tại Việt Nam. Sự sụt giảm sức mạnh dự báo khẳng định thị trường đang chuyển dịch nhanh chóng về cấu trúc vi mô, khiến các mẫu hình động lượng lịch sử (2021-2023) bị triệt tiêu và mất đi tính lặp lại.

BẢNG 1: THỐNG KÊ MÔ TẢ DỮ LIỆU GIÁ ĐÓNG CỦA GIAI ĐOẠN 2021-2025

Đơn vị tính: Đồng			
Chỉ số thống kê	VCB	TCB	MBB
Số lượng phiên	1.100 phiên	1.100 phiên	1.100 phiên
Giá trung bình	88,450	36,200	21,800
Độ lệch chuẩn	9,120	11,450	4,950
Khoảng biến thiên	65,000 - 108,000	19,800 - 58,500	12,500 - 34,200
Hệ số biến thiên	10,3%	31,6%	22,7%

BẢNG 2: CÁC CHỈ SỐ ĐÁNH GIÁ HIỆU SUẤT MÔ HÌNH TRÊN TẬP KIỂM THỬ (NĂM 2025)

Mã cổ phiếu	Lớp mục tiêu	Precision	Recall	F1-Score	Support (Số phiên)	Accuracy	ROC-AUC
VCB	Lớp 0 (Giảm/Đứng)	0,38	0,07	0,11	119	51,00%	0,5067
	Lớp 1 (Tăng)	0,52	0,90	0,66	132		
TCB	Lớp 0 (Giảm/Đứng)	0,52	0,47	0,50	136	48,00%	0,4841
	Lớp 1 (Tăng)	0,44	0,50	0,47	115		
MBB	Lớp 0 (Giảm/Đứng)	0,41	0,46	0,43	120	43,00%	0,4138
	Lớp 1 (Tăng)	0,45	0,40	0,43	131		

Nguồn: Kết quả nghiên cứu

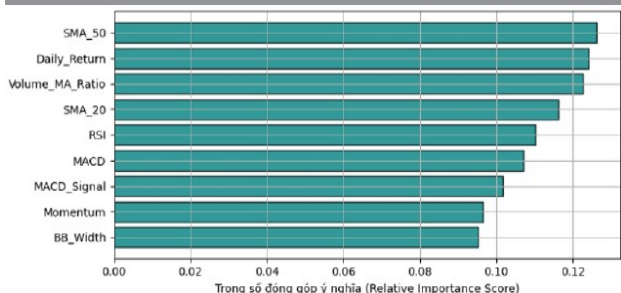
Đi sâu vào hành vi phân loại, kết quả bộc lộ sự thiên lệch cấu trúc gắn liền với đặc thù từng cổ phiếu. Ở VCB, mô hình đạt tỷ lệ Recall chiều tăng lên tới 90% nhưng gần như bỏ qua chiều giảm (Recall 7%). Sự thiên lệch này bắt nguồn từ chức năng điều tiết chỉ số của VCB; việc dòng tiền tổ chức duy trì nền giá dài hạn đã tạo ra sai số hệ thống, khiến thuật toán bị thiên lệch về xác suất tăng giá. Ngược lại, TCB và MBB ghi nhận phân phối dự báo cân bằng hơn, phản ánh chính xác đặc tính giao dịch tự do của thị trường, dù sai số tổng thể vẫn tiềm ẩn rủi ro nếu ứng dụng thực chiến.

Về cấu trúc động lực giá vi mô, biểu đồ tầm quan trọng đặc trưng và biểu đồ SHAP Summary Plot chỉ ra sự phân hóa rõ rệt giữa hai nhóm ngân hàng. Đối với VCB (Hình 1 và Hình 2), nhóm chỉ báo động lượng chiếm ưu thế tuyệt đối, trong đó RSI đóng vai trò dẫn dắt. Việc các giá trị SHAP dương hội tụ tại vùng RSI cao cho thấy khi cổ phiếu tiến sát ngưỡng quá mua, mô hình ghi nhận xác suất điều chỉnh giảm, qua đó khẳng định chức năng điều tiết và duy trì sự ổn định của cổ phiếu quốc doanh vốn hóa lớn.

Trái ngược với tính ổn định của VCB, TCB và MBB thể hiện độ nhạy cảm cao với biến động dải biên và thanh khoản. Tại TCB (Hình 3 và Hình 4), độ rộng dải Bollinger (BB_Width) và tỷ lệ khối lượng giao dịch (Volume_MA_Ratio) bám sát vị trí dẫn đầu của RSI, phản ánh sự phụ thuộc vào dao động biên

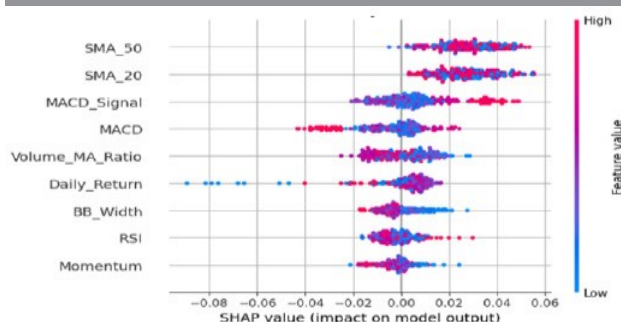


HÌNH 1: BIỂU ĐỒ TẦM QUAN TRỌNG CỦA CÁC ĐẶC TRƯNG - VCB



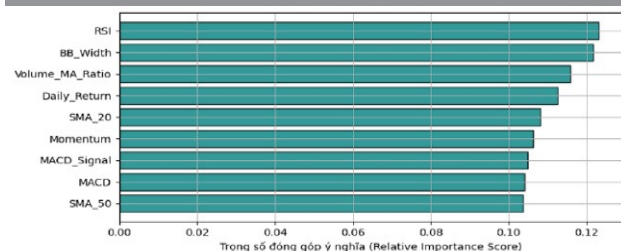
Nguồn: Kết quả nghiên cứu

HÌNH 2: BIỂU ĐỒ TÓM TẮT SHAP (SHAP SUMMARY PLOT) - VCB



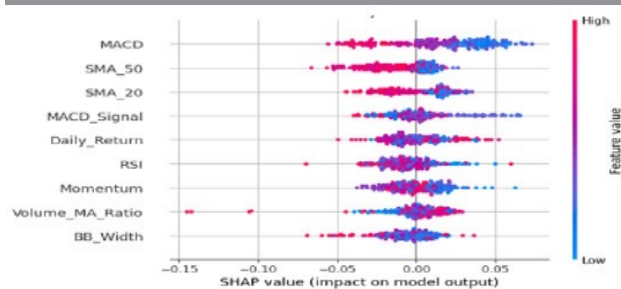
Nguồn: Kết quả nghiên cứu

HÌNH 3: BIỂU ĐỒ TẦM QUAN TRỌNG CỦA CÁC ĐẶC TRƯNG - TCB



Nguồn: Kết quả nghiên cứu

HÌNH 4: BIỂU ĐỒ TÓM TẮT SHAP (SHAP SUMMARY PLOT) - TCB

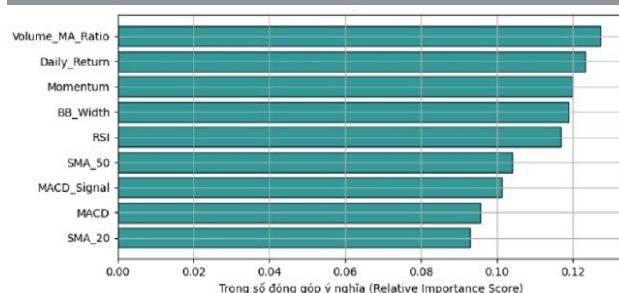


Nguồn: Kết quả nghiên cứu

độ và dòng tiền ngắn hạn. Tương tự, tại MBB (Hình 5 và Hình 6), biến số đột phá thanh khoản chiếm vị trí trọng yếu nhất, đồng nhất với đặc tính của một cổ phiếu có độ đại chúng cao và thu hút lượng lớn giao dịch năng động.

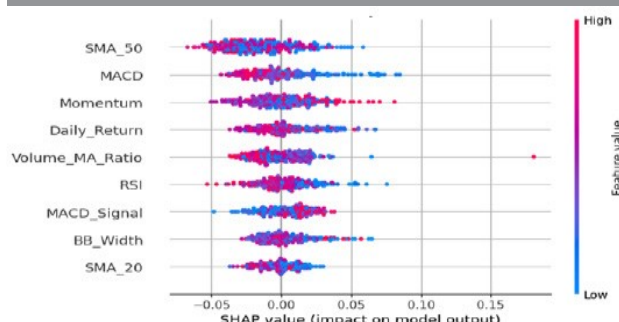
Điểm nhấn khoa học của nghiên cứu được bộc lộ qua biểu đồ SHAP của nhóm tư nhân, minh chứng

HÌNH 5: BIỂU ĐỒ TẦM QUAN TRỌNG CỦA CÁC ĐẶC TRƯNG - MBB



Nguồn: Kết quả nghiên cứu

HÌNH 6: BIỂU ĐỒ TÓM TẮT SHAP (SHAP SUMMARY PLOT) - MBB



Nguồn: Kết quả nghiên cứu

cho sự phân tách giữa biến số rẽ nhánh và biến số quyết định. Dù các chỉ báo ngắn hạn được ưu tiên để chia tách trạng thái, xu hướng trung hạn (SMA_50) và quán tính động lượng (MACD) mới là các biến số nền tảng quyết định biên xác suất đầu ra. Sự hội tụ này chứng tỏ thuật toán đã tự tối ưu hóa, sử dụng xu hướng dài hạn nhằm kiểm soát và triệt tiêu các tín hiệu nhiễu do dòng tiền đầu cơ ngắn hạn gây ra trong môi trường dữ liệu phức tạp.

Thảo luận kết quả nghiên cứu

Kết quả thực chứng từ nghiên cứu củng cố mạnh mẽ thuyết thị trường hiệu quả (EMH) dạng yếu tại Việt Nam. Sự suy giảm hiệu năng dự báo trên tập dữ liệu tương lai (2025) phản ánh sự dịch chuyển cấu trúc vi mô của thị trường; theo đó, các thông tin và mẫu hình động lượng trong lịch sử nhanh chóng bị triệt tiêu, làm gia tăng tính ngẫu nhiên trong biến động giá.

Thông qua phương pháp XAI, nghiên cứu bóc tách được cấu trúc động lực giá gắn liền với đặc thù mô hình kinh doanh của từng ngân hàng. Đối với VCB, sự thiên lệch dự báo chiều tăng xuất phát từ đặc tính điều tiết của cổ phiếu này. Việc dòng tiền tổ chức liên tục duy trì nền giá để ổn định thị trường đã tạo ra sự mất cân bằng dữ liệu, gây ra sai số hệ thống và làm giảm độ nhạy của thuật toán trước các nhịp điều chỉnh giảm. Ngược lại, sự biến động của TCB và MBB phụ thuộc mạnh vào dải Bollinger và quy mô thanh khoản, phản ánh bản chất giao dịch ngắn hạn của khối ngân hàng tư nhân. Đáng chú ý,



việc thuật toán tự động phân bổ trọng số cao cho SMA_50 và MACD, sử dụng chúng làm các biến số nền tảng để định hướng xác suất cho TCB và MBB, cho thấy hệ thống đã tự tối ưu hóa. Cơ chế này khai thác xu hướng trung hạn nhằm triệt tiêu các tín hiệu nhiễu do dòng tiền đầu cơ ngắn hạn gây ra.

Phân tích hồi cứu đồng thời chỉ ra giới hạn cốt lõi của các mô hình học máy thuần kỹ thuật. Năng lực dự báo bị vô hiệu hóa trước các cú sốc ngoại sinh và rủi ro đuôi, điển hình như cuộc khủng hoảng thanh khoản hệ thống tháng 10/2022. Trong bối cảnh rủi ro hệ thống leo thang, các chỉ báo động lượng (RSI) liên tục kích hoạt tín hiệu phục hồi giả tại vùng quá bán, trong khi giá cổ phiếu thực tế sụt giảm sâu do hệ lụy giải chấp chéo. Điều này khẳng định thuật toán không thể dự báo các đứt gãy kinh tế nếu thiếu vắng biến số tâm lý vĩ mô. Thêm vào đó, trong các giai đoạn kiệt quệ thanh khoản, giao dịch cường độ thấp làm khuếch đại nhiễu ngẫu nhiên, gia tăng tỷ lệ báo động giả. Điều này tái khẳng định nguyên lý định lượng: các tín hiệu kỹ thuật chỉ đạt độ tin cậy cao khi được xác nhận bởi quy mô thanh khoản đột phá.

KẾT LUẬN VÀ KHUYẾN NGHỊ GIẢI PHÁP

Nghiên cứu ứng dụng phương pháp SHAP kết hợp thuật toán Random Forest nhằm dự báo xu hướng và giải mã cấu trúc động lực giá của nhóm cổ phiếu ngân hàng dẫn dắt giai đoạn 2021-2025. Kết quả thực chứng khẳng định vai trò định hướng của RSI và SMA_20, đồng thời chỉ ra sự phân hóa vi mô sâu sắc: khối quốc doanh (VCB) mang đặc tính phòng thủ và vận động theo chu kỳ động lượng; trong

khí khối tư nhân (TCB, MBB) nhạy cảm với thanh khoản ngắn hạn nhưng chịu sự chi phối nền tảng bởi xu hướng trung hạn. Bên cạnh đó, sự suy giảm hiệu năng dự báo trên tập kiểm thử độc lập (2025) cung cấp minh chứng rõ nét về sự dịch chuyển của thị trường chứng khoán Việt Nam theo Thuyết thị trường hiệu quả (EMH) dạng yếu.

Từ các phát hiện thực chứng, nghiên cứu đề xuất các hàm ý quản trị trên hai phương diện. Về phía nhà quản lý quỹ và đầu tư, hệ thống học máy nên được định vị là công cụ hỗ trợ quyết định nhằm gia tăng lợi thế thông tin thay vì sử dụng như một mô hình giao dịch độc lập. Theo đó, chiến lược phân bổ vốn cần được thiết lập linh hoạt theo cấu trúc động lực của từng cổ phiếu và đặt trong một khung quản trị rủi ro danh mục nghiêm ngặt. Việc ứng dụng kiểm thử quá khứ chuyên sâu để đo lường tỷ suất sinh lời trên rủi ro và mức sụt giảm tài sản tối đa là yêu cầu tất yếu nhằm đối phó hiệu quả với tính phi tuyến của thị trường. Đồng thời, dưới góc độ vĩ mô đối với các cơ quan hoạch định chính sách, việc các mô hình định lượng bị vô hiệu hóa trước các cú sốc ngoại sinh đã nhấn mạnh vai trò thiết yếu của tính minh bạch thông tin. Các cơ quan quản lý cần đẩy mạnh những giải pháp giảm thiểu bất đối xứng thông tin, đồng thời kiểm soát chặt chẽ rủi ro liên thông từ thị trường trái phiếu và bất động sản. Cùng với đó, các hoạt động truyền thông chính sách kịp thời sẽ đóng vai trò then chốt trong việc ngăn chặn các cú sốc thanh khoản hệ thống cũng như tâm lý hoảng loạn đám đông, từ đó bảo vệ vững chắc cơ chế tự điều tiết bền vững của thị trường tài chính.■

Tài liệu tham khảo:

1. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1).
2. Box, G. E. P., & Jenkins, G. M. (1970). *Time Series Analysis: Forecasting and Control*. Holden-Day.
3. Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>.
4. Chen, W., Hussain, W., Cauteruccio, F., & Zhang, X. (2023). Deep learning for financial time series prediction: A state-of-the-art review of standalone and hybrid models. *CMES-Computer Modeling in Engineering and Sciences*. <https://opus.lib.uts.edu.au/handle/10453/179034>.
5. Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223-236. <https://doi.org/10.1080/713665670>.
6. Deniz, E., Chifumi, N., & Ansgar, S. (2020). Comparison of machine learning methods for financial time series forecasting at the examples of over 10 years of daily and hourly data of DAX 30 and S&P 500. *Journal of Computational Social Science*, 3(1), 103-133.
7. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30. <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>.
8. Trần Thị Mỹ Linh. (2025, May 15). *Dự báo giá cổ phiếu bằng kỹ thuật học máy: Nghiên cứu từ thị trường chứng khoán Việt Nam*. <https://tapchicongthuong.vn/du-bao-gia-co-phieu-bang-ky-thuat-hoc-may--nghiien-cuu-tu-thi-truong-chung-khoan-viet-nam-417103.htm>.
9. Winter, E. (2002). Chapter 53 The shapley value. In *Handbook of Game Theory with Economic Applications* (Vol. 3, pp. 2025-2054). Elsevier. [https://doi.org/10.1016/S1574-0005\(02\)03016-3](https://doi.org/10.1016/S1574-0005(02)03016-3).