



Phản ứng của thị trường chứng khoán trước các bản tin kinh tế vĩ mô

HUYỀN TẤN ĐẠT^{a, *}, TỬ HÀ PHÚC^b, ĐẶNG ANH NHƯ^b

^a University of Technology Sydney

^b Trường Đại học Kinh tế - Luật, Đại học Quốc gia Thành phố Hồ Chí Minh

THÔNG TIN	TÓM TẮT
<p>Ngày nhận: 21/08/2025 Ngày nhận lại: 25/12/2025 Duyệt đăng: 26/12/2025</p> <p>Mã phân loại JEL: G14; G17.</p> <p>Từ khóa: Phản ứng thị trường chứng khoán; Tin tức kinh tế; Phân tích cảm xúc; Mô hình ngôn ngữ lớn; AI Giải thích.</p> <p>Keywords: Stock market reaction; Economic news; Sentiment analysis; Large language models; Explainable AI.</p>	<p>Nghiên cứu này khám phá phản ứng của VN-Index trước các tin tức kinh tế vĩ mô bằng cách áp dụng xử lý ngôn ngữ tự nhiên và máy học tiên tiến. Nhóm tác giả phân tích 18.253 bài báo kinh tế từ trang cafef.vn, thu thập từ ngày 01/09/2021 đến 31/08/2024, để đánh giá mức độ tác động của 12 yếu tố kinh tế quan trọng và cảm xúc tổng thể. Sử dụng Mô hình ngôn ngữ lớn như GPT-4o-mini (dùng để đánh giá mức độ tác động chính) và Gemini-1.5-flash (đối chiếu với độ tương đồng 70%), mức độ tác động được định lượng và tích hợp với dữ liệu lịch sử VN-Index để xây dựng mô hình XGBoost. Mô hình dự đoán giá đóng cửa VN-Index với R^2 đạt 0,9879, thể hiện khả năng dự báo vượt trội. Các công cụ trí tuệ nhân tạo giải thích như SHAP và LIME được sử dụng để phân tích, cho thấy "Bảo cáo thu nhập và cổ tức", "Chính sách đầu tư nước ngoài" và "Ổn định chính trị" là yếu tố chính tác động đến biến động VN-Index. Nghiên cứu xác nhận ảnh hưởng mạnh mẽ của tin tức kinh tế lên thị trường, bổ sung lý thuyết nhà đầu tư bất hợp lý bằng cách chứng minh phản ứng nhà đầu tư thay đổi theo loại tin tức, đồng thời nâng cao hiểu biết về động lực thị trường tại các nước mới nổi như Việt Nam và cung cấp khung AI minh bạch hỗ trợ nhà đầu tư ra quyết định.</p> <p>Abstract</p> <p>This study examines the VN-Index's response to macroeconomic news through advanced natural language processing (NLP) and machine learning techniques. The authors analyzed 18,253 economic news articles from Cafef.vn, collected from September 1, 2021, to August 31,</p>

* Tác giả liên hệ.

Biên tập viên: Nguyễn Lương Tâm.

Email: tandat.huynh@student.uts.edu.au (Huỳnh Tấn Đạt); phuchtt22406@st.uel.edu.vn (Tử Hà Phúc); nhuda22416c@st.uel.edu.vn (Đặng Anh Như).

Trích dẫn bài viết: Huỳnh Tấn Đạt, Tử Hà Phúc, & Đặng Anh Như. (2025). Phản ứng của thị trường chứng khoán trước các bản tin kinh tế vĩ mô. *Tạp chí Nghiên cứu Kinh tế và Kinh doanh Châu Á*, 36(11), 04-22, <https://doi.org/10.24311/jabes/2025.36.11.01>

2024, to evaluate the impact of 12 key economic factors and overall sentiment. Utilizing Large Language Models (LLMs) such as GPT-4o-mini (for primary sentiment scoring) and Gemini-1.5-flash (for validation with 70% correlation), the authors quantified sentiment and integrated it with VN-Index historical data to build an XGBoost model. The model predicts VN-Index closing prices with an R^2 of 0.9879, demonstrating superior forecasting accuracy. Explainable AI (XAI) tools like SHAP and LIME were applied to interpret results, identifying “Earnings Reports and Dividends”, “Foreign Investment Policies”, and “Political Stability” as the main drivers of VN-Index volatility. This research confirms the strong impact of macroeconomic news on stock market dynamics, enriches the “noise trader behavior” theory by illustrating category-specific investor reactions, and provides actionable insights for market dynamics in emerging economies like Vietnam, offering investors a transparent AI framework for decision-making.

1. Giới thiệu

Thị trường chứng khoán (TTCK) đóng vai trò quan trọng trong nền kinh tế, huy động vốn hiệu quả cho doanh nghiệp và phản ánh kỳ vọng của nhà đầu tư về triển vọng kinh tế. Tại Việt Nam, chỉ số đại diện cho thị trường chứng khoán Việt Nam (VN-Index) thường xuyên biến động dưới tác động của các bản tin kinh tế vĩ mô trong và ngoài nước, chẳng hạn: tăng trưởng GDP, lãi suất, lạm phát, và chính sách tài chính. Những thông tin này ảnh hưởng đến tâm lý nhà đầu tư, định hình xu hướng thị trường, đặc biệt ở một thị trường mới nổi nhạy cảm với yếu tố kinh tế - chính trị. Tuy nhiên, định lượng mối quan hệ giữa tin tức và VN-Index vẫn là thách thức, đòi hỏi phương pháp phân tích tiên tiến phù hợp với đặc thù địa phương. Các nghiên cứu trước thường tập trung rời rạc vào từng yếu tố vĩ mô, thiếu cách tiếp cận tích hợp để đánh giá sự xuất hiện và tương tác của chúng trong dữ liệu văn bản lớn.

Trên thế giới, mối liên hệ giữa tin tức kinh tế và biến động chứng khoán đã được chứng minh qua các công trình như: Tetlock (2007) phát hiện cảm xúc tiêu cực từ báo chí dự đoán giảm chỉ số trung bình công nghiệp Dow Jones, Li và cộng sự (2020) chứng minh rằng việc tích hợp phân tích cảm xúc từ tin tức với các chỉ báo kỹ thuật giúp cải thiện đáng kể độ chính xác dự đoán giá cổ phiếu tại thị trường Hong Kong – một thị trường mới nổi có nhiều đặc điểm tương đồng với Việt Nam. Gần đây, sự phát triển của mô hình ngôn ngữ lớn (Large Language Models – LLMs) như GPT và Gemini đã nâng cao khả năng khai thác văn bản, với Kirtac và Germano (2024) đạt độ chính xác 74,4% dự đoán lợi nhuận Mỹ, hay Kwon và cộng sự (2025) sử dụng LLM phân tích cảm xúc với bản tin vĩ mô thời gian thực cho chính sách ở các nước mới nổi. Tại Việt Nam, các nghiên cứu còn hạn chế, với Phuoc và cộng sự (2024) và Khang và cộng sự (2021) chủ yếu dựa vào dữ liệu lịch sử để dự đoán giá hoặc tính thanh khoản, chưa khai thác triệt để tin tức hay công nghệ trí tuệ nhân tạo (Artificial Intelligence – AI) hiện đại như LLMs và AI Giải thích (Explainable AI – XAI). Trong bối cảnh thị trường Việt Nam hội nhập toàn cầu, khoảng trống này cần được lấp đầy bằng cách tiếp cận toàn diện hơn.

Để khắc phục, nghiên cứu này đánh giá tác động của tin tức kinh tế vĩ mô lên VN-Index bằng cách sử dụng LLMs làm công cụ phân tích sự xuất hiện và mức độ tác động (tích cực/tiêu cực) của 12 yếu tố vĩ mô (từ GDP đến ổn định chính trị) trong 18.253 bài báo từ trang *cafef.vn* (giai đoạn 2021–2024). Phương pháp kết hợp LLMs (GPT và Gemini) để định lượng các yếu tố kinh tế, mô hình XGBoost dự đoán giá đóng cửa, và XAI với SHapley Additive exPlanations (SHAP) (Lundberg & Lee, 2017) và Local Interpretable Model-agnostic Explanations (LIME) (Ribeiro và cộng sự, 2016) để giải thích kết quả dự đoán. Nghiên cứu này làm phong phú lý thuyết “nhà đầu tư bất hợp lý” (Noise Trader Behavior) của De Long và cộng sự (1990), chứng minh bản tin tích cực (như GDP tăng) thúc đẩy mua vào không lý tính dẫn đến giá tăng, trong khi tin tiêu cực (như lãi suất tăng) gây bán tháo, đặc biệt ở thị trường mới nổi như Việt Nam. Kết quả nhấn mạnh vai trò của "Báo cáo thu nhập và cổ tức" cùng "Chính sách đầu tư nước ngoài".

Về lý thuyết, nghiên cứu này mở rộng khung phân tích yếu tố kinh tế trong tài chính hành vi, cung cấp bằng chứng thực nghiệm từ thị trường mới nổi tương tự (Ónozó và cộng sự, 2024). Về thực tiễn, nghiên cứu hiện tại mang lại khung AI minh bạch, hỗ trợ nhà đầu tư ra quyết định trong môi trường biến động, góp phần phát triển thị trường chứng khoán Việt Nam.

2. Tổng quan các tài liệu liên quan

2.1. Các nghiên cứu liên quan

Các nghiên cứu quốc tế đã chứng minh mối quan hệ chặt chẽ giữa tin tức kinh tế và biến động chứng khoán. Tetlock (2007) phát hiện cảm xúc tiêu cực từ Nhật báo kinh doanh và tài chính (Wall Street Journal) dự đoán giảm Dow Jones, trong khi Li và cộng sự (2020) áp dụng phân tích bản tin để cải thiện dự đoán lợi nhuận tại Trung Quốc – thị trường mới nổi tương đồng với Việt Nam. Gần đây, mô hình ngôn ngữ lớn (LLMs) như GPT-4o-mini và Gemini-1.5-flash nâng cao độ chính xác xử lý văn bản phức tạp, với Kirtac và Germano (2024) đạt 74,4% dự đoán lợi nhuận Mỹ, Kwon và cộng sự (2025) sử dụng LLMs để phân tích cảm xúc vĩ mô thời gian thực từ tin tức kinh tế Mỹ, và Chen và cộng sự (2025) áp dụng mô hình ngôn ngữ tổng quát hóa (Generalized Language Model – GLM) và câu lệnh Chain-of-Thought (CoT)¹ để giảm thiên kiến trong phân tích cảm xúc thị trường chứng khoán. Những tiến bộ này mở ra cơ hội ứng dụng LLMs tại các thị trường khác, nhờ khả năng định lượng cảm xúc từ dữ liệu lớn mà không cần huấn luyện thủ công.

Nghiên cứu tại Việt Nam còn hạn chế, với Phuoc và cộng sự (2024) dùng mô hình mạng nơ-ron hồi quy bộ nhớ dài ngắn (Long Short-term Memory – LSTM) dự đoán giá cổ phiếu đạt 93% chính xác nhưng chỉ dựa vào dữ liệu lịch sử, Khang và cộng sự (2021) dự đoán thanh khoản giai đoạn 2011–2019 mà không khai thác tin tức hoặc phân tích cảm xúc từ chúng. Khoảng trống này nổi bật khi thị trường Việt Nam chịu ảnh hưởng mạnh từ thông tin vĩ mô, đòi hỏi cách tiếp cận toàn diện hơn. Nghiên cứu hiện tại lấp đầy khoảng trống này bằng cách sử dụng GPT-4o-mini với câu lệnh Chain-of-Thought để phân tích cảm xúc từ bản tin, và sử dụng thêm Gemini-1.5-flash để đối chiếu cách đánh giá giữa hai mô hình, đạt độ tương đồng 70% nhằm xác nhận độ tin cậy và giảm sai lệch, sau đó tích hợp vào

¹ Chain-of-Thought (CoT) là một kỹ thuật trong mô hình ngôn ngữ lớn (LLMs), nơi các bước suy luận được thực hiện theo một chuỗi suy nghĩ để giải quyết các vấn đề phức tạp. Thay vì đưa ra kết quả trực tiếp, mô hình sẽ mô tả các bước tư duy của mình, giúp cải thiện khả năng hiểu và ra quyết định, đặc biệt trong các tác vụ yêu cầu phân tích hoặc lý luận chi tiết.

XGBoost dự đoán giá (với Hyperparameter Tuning² như Grid Search³), và XAI (SHAP, LIME) giải thích đóng góp yếu tố. Điều này không chỉ minh bạch hóa cơ chế từ bản tin đến mô hình mà còn bổ sung lý thuyết “nhà đầu tư bất hợp lý” (De Long và cộng sự, 1990) thông qua bằng chứng thực nghiệm tại thị trường mới nổi.

2.2. Các lý thuyết kinh tế và lựa chọn các yếu tố vĩ mô

Các yếu tố vĩ mô được lựa chọn nhằm phân tích tác động kinh tế đến thị trường chứng khoán, đặc biệt tại các nền kinh tế mới nổi như Việt Nam. Tổng cộng 12 yếu tố được xác định từ lý thuyết định giá cổ phiếu (Santow & Gordon, 1962) và bằng chứng thực nghiệm (Chikwira & Mohammed, 2023), cũng như các nghiên cứu địa phương (Thanh & Linh, 2016). Các yếu tố này thường được xem xét rời rạc, nhưng nghiên cứu này tích hợp các yếu tố qua các bản tin để đánh giá ảnh hưởng tổng hợp đến VN-Index, nhấn mạnh vai trò kinh tế trong biến động giá.

Tăng trưởng GDP, chỉ số cốt lõi phản ánh sản lượng quốc gia, thường thúc đẩy giá cổ phiếu nhờ nâng cao lợi nhuận kỳ vọng doanh nghiệp. Bunjaku (2024) chứng minh mối quan hệ tích cực với β đạt 0,911 cho nhóm S&P 500 tại Mỹ (1990–2019), tương tự, Chikwira và Mohammed (2023) tại Zimbabwe (2013–2022) với ý nghĩa thống kê 10%. Tuy nhiên, Ritter (2004) ghi nhận tương quan âm ở thị trường phát triển do lợi ích không trực tiếp đến cổ đông. Tại Việt Nam, dự báo tăng trưởng GDP đạt khoảng 6,3 % (World Bank, 2023) liên kết với tăng nhẹ VN-Index sau dữ liệu quý III/2023 (Biswas, 2023), khẳng định tác động tích cực đến tăng trưởng kinh tế. Bên cạnh đó, lãi suất tăng ảnh hưởng đáng kể đến chi phí vốn, thường giảm giá cổ phiếu dài hạn tại Việt Nam (Thanh & Linh, 2016; Nguyen và cộng sự, 2023), với rủi ro hệ thống lan tỏa gián tiếp. Đối với quốc tế, tác động phụ thuộc chu kỳ kinh tế, rõ nét hơn trong suy thoái (Wang & Mayes, 2012; Ngo và cộng sự, 2024), dù một số ngành như xuất khẩu Hàn Quốc phản ứng tích cực với lãi suất của Cục Dự trữ Liên bang Hoa Kỳ (Federal Reserve System – FED) nhờ điều chỉnh tỷ giá (Kim, 2023).

Quy mô thị trường chứng khoán, đo lường qua vốn hóa hoặc số niêm yết, thúc đẩy tăng trưởng kinh tế bằng huy động vốn hiệu quả và quản lý rủi ro (Demirgüç-Kunt & Levine, 1996; Levine, 1997). Tại Việt Nam, mối liên hệ tích cực được xác nhận qua nghiên cứu của Su và cộng sự (2017), hay của Nguyen và Nguyen (2022) khi so sánh với Trung Quốc, và nghiên cứu của Thanh và Linh (2016) nhấn mạnh sự đóng góp vào GDP. Chính sách thuế tác động đến lợi nhuận sau thuế và hành vi đầu tư (Downs & Hendershott, 1986), với thuế tối thiểu toàn cầu ảnh hưởng gián tiếp qua vốn đầu tư nước ngoài (Foreign Direct Investment – FDI) tại Việt Nam (Hoi và cộng sự, 2024); tuy nhiên, ít biến động trong giai đoạn 2021–2024 giải thích tác động hạn chế.

Lạm phát thường gây tác động tiêu cực bằng cách làm giảm sức mua và tăng chi phí sản xuất (Boudoukh & Richardson, 1993; Lee, 2010). Tác động này rõ nét hơn tại Việt Nam trong các giai đoạn lạm phát cao và kinh tế biến động mạnh (2001–2022), khi lạm phát làm gia tăng đáng kể biến động của VN-Index do hành vi bày đàn của nhà đầu tư (Nguyen & Le, 2024). Tuy nhiên, trong giai đoạn ổn định tương đối 2021–2024, ảnh hưởng của lạm phát đã giảm mạnh. Chính sách đầu tư nước

² Hyperparameter Tuning là quá trình tối ưu hóa các siêu tham số của mô hình học máy để cải thiện hiệu suất, thường được thực hiện qua các kỹ thuật như Grid Search, Random Search, hay Bayesian Optimization.

³ Grid Search là phương pháp tìm kiếm tối ưu các giá trị của siêu tham số bằng cách thử tất cả các kết hợp có thể có trong một không gian đã được xác định trước, giúp xác định cấu hình mô hình tốt nhất dựa trên các chỉ số hiệu suất.

ngoài định hình hiệu suất thị trường ở nước đang phát triển nhờ dòng vốn ổn định (Doukas & Lang, 2003).

Ổn định chính trị là nền tảng giảm rủi ro hệ thống, đặc biệt ở thị trường mới nổi, với bất ổn dẫn đến giảm giá cổ phiếu (Boutchkova và cộng sự, 2012; Pástor & Veronesi, 2013). Tại Việt Nam, các sự kiện chính trị gây biến động VN-Index (Cao & Vo, 2025) khẳng định vai trò quan trọng trong tăng trưởng kinh tế bền vững.

Báo cáo thu nhập và cổ tức ảnh hưởng trực tiếp đến định giá cổ phiếu, với thị trường phản ứng nhanh chóng theo lý thuyết thị trường hiệu quả (Fama và cộng sự, 1969). Tại Việt Nam, tác động tức thì được ghi nhận (Phan & Tran, 2019), phản ánh sức khỏe tài chính doanh nghiệp là yếu tố cốt lõi định hình xu hướng VN-Index.

Thay đổi quy định pháp lý tạo bất ổn hoặc cơ hội đầu tư (Bhattacharya & Daouk, 2002; Coates, 2007), nhưng ít tác động tại Việt Nam gần đây do ít cải cách. Hoạt động M&A thay đổi cấu trúc thị trường và kỳ vọng lợi nhuận (Jensen & Ruback, 1983; Moeller và cộng sự, 2004), với tiềm năng tại Việt Nam dù quy mô nhỏ (Vuong và cộng sự, 2009). Yếu tố tình cảm mạng xã hội dự đoán xu hướng qua tâm lý nhà đầu tư (Bollen và cộng sự, 2011; Chen và cộng sự, 2014), tích cực nhưng yếu tại Việt Nam do phát triển sơ khai (Van và cộng sự, 2024). Tin tức kinh tế toàn cầu ảnh hưởng đến nền kinh tế mở (Andersen và cộng sự, 2007; Ehrmann & Fratzscher, 2005), với tương quan dương đến VN-Index, phản ánh mức độ hội nhập.

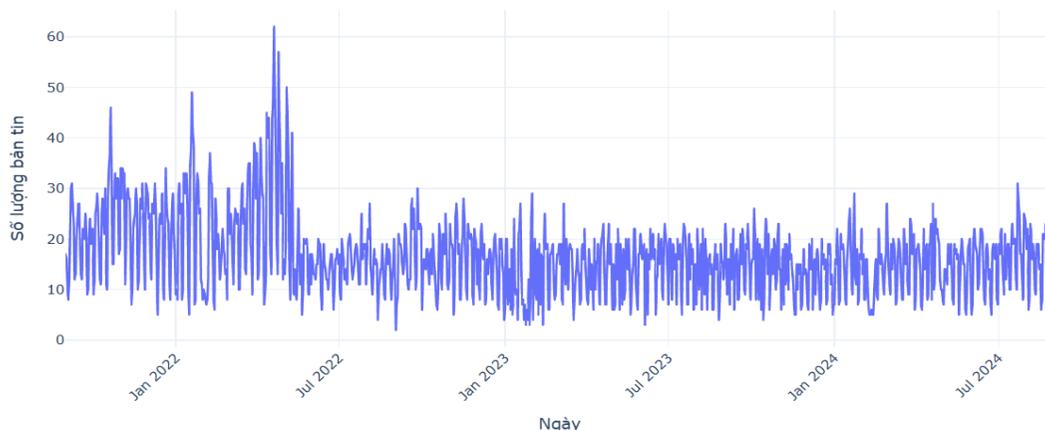
3. Dữ liệu và Phương pháp nghiên cứu

3.1. Dữ liệu

Nghiên cứu này phân tích phản ứng của VN-Index trước tin tức kinh tế vĩ mô từ ngày 01/09/2021 đến 31/08/2024, sử dụng hai nguồn dữ liệu chính. Bản tin lấy từ trang thông tin cafef.vn (mục "Thị trường chứng khoán"⁴, tổng 18.253 bài báo thu thập theo ngày, đề cập yếu tố vĩ mô (GDP, lạm phát, và lãi suất), chính sách tài chính và hoạt động doanh nghiệp (báo cáo tài chính, cổ tức, và sáp nhập). Trung bình 25 bài/ngày trong 748 ngày giao dịch, độ dài trung bình 730 từ, dao động theo sự kiện (Hình 1). Thống kê tần suất yếu tố chi tiết ở Bảng 1, phản ánh trọng tâm kinh tế trong bản tin. Dữ liệu VN-Index từ trang Refinitiv bao gồm giá mở cửa, đóng cửa, cao nhất, thấp nhất hàng ngày, dùng tính biến động và đánh giá tác động bản tin lên TTCK Việt Nam.

⁴ Xem thêm tại <https://cafef.vn/thi-truong-chung-khoan.chn>

Số lượng bản tin kinh tế theo ngày



Hình 1. Số lượng bản tin theo ngày

Bảng 1.

Thông kê mô tả dữ liệu văn bản và tần suất 12 yếu tố vĩ mô.

Yếu tố vĩ mô	Số bài đề cập	Tỷ lệ (%)
Báo cáo thu nhập và cổ tức	8.062	44,17
Tình cảm trên mạng xã hội	2.226	12,20
Lãi suất	2.092	11,46
Quy mô thị trường	1.949	10,68
Thay đổi quy định	1.486	8,14
Ổn định chính trị	1.350	7,40
Chính sách đầu tư nước ngoài	1.097	6,01
Lạm phát	1.022	5,60
Sáp nhập và mua lại (M&A)	796	4,36
Tăng trưởng GDP	779	4,27
Tỷ giá hối đoái	455	2,49
Chính sách thuế	386	2,11

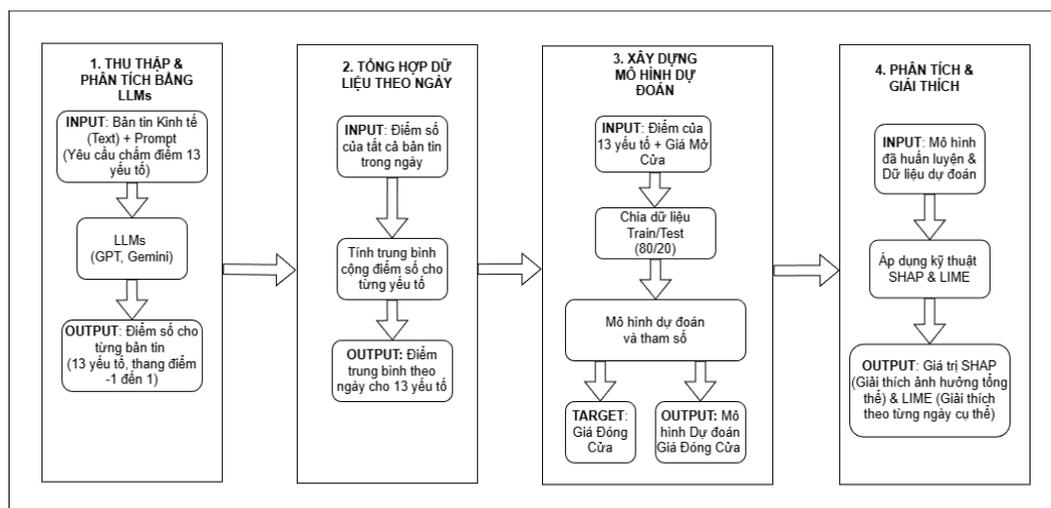
3.2. Phương pháp nghiên cứu lý thuyết

Phương pháp nghiên cứu lý thuyết được xây dựng nhằm xác định các yếu tố kinh tế vĩ mô quan trọng ảnh hưởng đến thị trường chứng khoán Việt Nam, đặc biệt là VN-Index. Quá trình này được thực hiện qua ba bước chính gồm: (1) nghiên cứu tài liệu học thuật để xác định các yếu tố tiềm năng dựa trên lý thuyết và bằng chứng thực nghiệm từ các nghiên cứu trước đó; (2) kiểm tra và đối chiếu các yếu tố với các nghiên cứu khác (dựa trên phần 2) để đảm bảo tính toàn diện, dẫn đến 12 yếu tố

kinh tế vĩ mô được đề xuất: “Tăng trưởng GDP”, “Lãi suất”, “Quy mô thị trường”, “Chính sách thuế”, “Lạm phát”, “Chính sách đầu tư nước ngoài”, “Ổn định chính trị”, “Báo cáo thu nhập và cổ tức”, “Thay đổi quy định”, “Sáp nhập và mua lại (M&A)”, “Tình cảm trên mạng xã hội”, và “Tin tức kinh tế toàn cầu”; và (3) đánh giá tính thực tiễn của các yếu tố thông qua sự xuất hiện trong bản tin kinh tế, dựa trên tần suất và mức độ quan trọng để đảm bảo ứng dụng cao. Tóm lại, phương pháp được thực hiện có hệ thống, đảm bảo danh sách 12 yếu tố kinh tế vĩ mô có cơ sở khoa học chặt chẽ và phù hợp với thị trường chứng khoán Việt Nam.

3.3. Phương pháp nghiên cứu thực nghiệm

Phương pháp nghiên cứu thực nghiệm được thiết kế nhằm phân tích và dự đoán phản ứng của thị trường chứng khoán Việt Nam trước các bản tin kinh tế vĩ mô thông qua bốn giai đoạn chính: thu thập và phân tích bản tin bằng Mô hình ngôn ngữ lớn, tổng hợp dữ liệu theo ngày, xây dựng mô hình dự đoán giá đóng cửa của VN-Index, và phân tích kết quả bằng kỹ thuật AI Giải thích (XAI). Quy trình tổng thể được minh họa ở Hình 2, với cơ chế bốn bước chính: (1) Input bản tin (tiêu đề và nội dung từ trang cafe.vn, giai đoạn 2021–2024 đại diện phục hồi sau COVID và biến động vĩ mô, sử dụng phương pháp CoT (Wei và cộng sự, 2022) để thiết kế câu lệnh nhằm đánh giá mức độ tác động (tích cực/tiêu cực) của 12 yếu tố; (2) tính điểm trung bình theo ngày cho từng yếu tố và tổng hợp với giá mở cửa; (3) xây dựng mô hình dự đoán bằng XGBoost (Chen & Guestrin, 2016), với chia dữ liệu train/test (80/20) để huấn luyện và kiểm tra; (4) áp dụng SHAP và LIME để giải thích đóng góp của từng yếu tố đến kết quả dự đoán giá đóng cửa. Quá trình này tập trung vào việc định lượng tác động của các yếu tố kinh tế từ bản tin lên biến động giá cổ phiếu, đồng thời đảm bảo tính minh bạch và khả năng giải thích của kết quả dự đoán, mang lại giá trị thực tiễn cho nhà đầu tư và nhà phân tích.



Hình 2. Quy trình nghiên cứu tổng thể: Xử lý dữ liệu, tổng hợp mức độ tác động (tích cực/tiêu cực) theo từng yếu tố, xây dựng mô hình dự đoán và giải thích kết quả bằng XAI.

3.3.1. Phân tích bản tin kinh tế bằng Mô hình ngôn ngữ lớn

Giai đoạn đầu tiên sử dụng hai Mô hình ngôn ngữ lớn là GPT-4o-mini và Gemini-1.5-flash với các câu lệnh cụ thể⁵ để xử lý và đánh giá các bản tin kinh tế. Trong đó, GPT-4o-mini được lựa chọn làm mô hình chính nhằm trích xuất điểm số đánh giá và làm đầu vào cho mô hình dự báo, trong khi Gemini-1.5-flash được sử dụng như một mô hình đối chứng độc lập để kiểm tra tính nhất quán và độ tin cậy của quá trình đánh giá. Các mô hình này được lựa chọn nhờ khả năng phân tích ngôn ngữ tự nhiên và độ chính xác cao trong việc xử lý thông tin tài chính. Để đảm bảo tính chi tiết và nhất quán trong quá trình đánh giá, kỹ thuật sử dụng câu lệnh CoT được tích hợp vào thiết kế câu lệnh, hướng dẫn mô hình phân tích tác động của từng yếu tố kinh tế lên thị trường chứng khoán. Cụ thể, mỗi bản tin được đánh giá dựa trên 12 yếu tố kinh tế và đánh giá tổng quan toàn bộ bản tin. Kết quả đánh giá được định lượng trên thang điểm từ -1 (tác động tiêu cực) đến 1 (tác động tích cực) cho mỗi yếu tố đến chỉ số VN-Index, lý do chọn thang này là để chuẩn hóa dữ liệu để tích hợp vào mô hình và phản ánh tác động (tiêu cực giảm giá, tích cực tăng giá). Chỉ các điểm số được sinh ra từ mô hình GPT-4o-mini mới được sử dụng làm dữ liệu đầu vào cho mô hình dự báo, trong khi kết quả từ Gemini-1.5-flash được dùng để so sánh và đánh giá mức độ tương đồng trong nhận định giữa hai mô hình ngôn ngữ lớn. Kết quả thực nghiệm cho thấy mức độ tương đồng trong đánh giá đạt khoảng 70%, cho thấy tính ổn định và độ tin cậy của phương pháp tiếp cận dựa trên LLM.

Giải thích Gemini:

```
{
  "Tăng trưởng GDP": 0.0,
  "Lãi suất": 0.0,
  "Quy mô thị trường": -0.2,
  "Chính sách thuế": 0.0,
  "Lạm phát": 0.0,
  "Chính sách đầu tư nước ngoài": -0.6,
  "Ổn định chính trị": 0.6,
  "Báo cáo thu nhập và cổ tức": 0.6,
  "Thay đổi quy định": 0.6,
  "Sáp nhập và mua lại (M&A)": 0.0,
  "Tinh cảm trên mạng xã hội": -0.3,
  "Tín tức kinh tế toàn cầu": -0.2,
  "Sentiment tổng thể": 0.2
}
```

Giải thích:

- **Tăng trưởng GDP (0.0):** Bản tin không đề cập đến số liệu GDP cụ thể.
- **Lãi suất (0.0):** Bản tin không đề cập đến biến động lãi suất.
- **Quy mô thị trường (-0.2):** Mặc dù thanh khoản được cải thiện, nhưng mức tăng của VN-Index chỉ 3.16% là thấp nhất trong 3 năm, cho thấy sự tăng trưởng của thị trường không mạnh mẽ.
- **Chính sách thuế (0.0):** Không đề cập đến chính sách thuế.
- **Lạm phát (0.0):** Không đề cập đến lạm phát.
- **Chính sách đầu tư nước ngoài (-0.6):** Khối ngoại bán ròng gần 26.000 tỷ đồng là một tín hiệu tiêu cực rất rõ rệt.
- **Ổn định chính trị (0.6):** Chính phủ có những chính sách thúc đẩy kinh tế tích cực, tạo niềm tin.
- **Báo cáo thu nhập và cổ tức (0.6):** Dự báo tăng trưởng thu nhập gần 20% cho các công ty Việt

Hình 3. Kết quả đánh giá bản tin kinh tế bằng LLMs

⁵ Câu lệnh chi tiết, bao gồm hướng dẫn suy luận cho từng yếu tố, được sử dụng trong quá trình nghiên cứu và có thể được cung cấp khi có yêu cầu thông qua tác giả liên hệ.

3.3.2. Xây dựng mô hình dự đoán giá đóng cửa VN-Index

Giai đoạn thứ hai triển khai hai mô hình máy học là XGBoost (Chen & Guestrin, 2016; Lundberg & Lee, 2017) và Random Forest (Breiman, 2001) nhằm dự đoán giá đóng cửa VN-Index và so sánh hiệu suất. Dữ liệu đầu vào gồm giá trị định lượng của 12 yếu tố kinh tế từ bản tin, mức độ tác động (tích cực/tiêu cực) tổng hợp, và giá mở cửa VN-Index. Dữ liệu được chuẩn bị bằng cách liên kết bản tin với VN-Index theo ngày, sử dụng trung bình các yếu tố cho ngày có nhiều bản tin để đảm bảo đại diện. Siêu tham số được điều chỉnh qua GridSearchCV⁶ trên tập huấn luyện cho cả hai mô hình.

XGBoost dựa trên Gradient Boosting Tree⁷, tối ưu sai số bình phương trung (Mean Squared Error – MSE) thông qua các cơ chế điều chuẩn hóa (Regularization) nhằm ngăn chặn hiện tượng quá khớp (Overfitting), đồng thời xử lý hiệu quả các quan hệ phi tuyến tính trong dữ liệu bảng và chuỗi thời gian tài chính. Siêu tham số: `n_estimators=126`, `learning_rate=0.11`, `max_depth=5`, `subsample=1.0`, `colsample_bytree=1.0`, `reg_lambda=1.0`, `gamma=0`, `random_state=42`. Trong khi đó, Random Forest dựa trên cây quyết định ngẫu nhiên, giảm thiểu phương sai thông qua phương pháp tổng hợp lấy mẫu (Bootstrap Aggregating – Bagging) kết hợp với việc đặc trưng ngẫu nhiên, phù hợp dữ liệu tương tác phức tạp. Siêu tham số: `n_estimators=50`, `max_depth=7`, `min_samples_split=10`, `min_samples_leaf=5`, `max_features='sqrt'`, `random_state=42`.

3.3.3. Áp dụng kỹ thuật AI Giải thích (XAI)

Cuối cùng, nghiên cứu này áp dụng hai kỹ thuật XAI: SHAP (Lundberg & Lee, 2017) và LIME để phân tích dự đoán của mô hình XGBoost. SHAP cung cấp tổng quan về mức độ quan trọng trung bình của từng yếu tố kinh tế đối với dự đoán giá đóng cửa trên toàn bộ dữ liệu, giúp xác định yếu tố ảnh hưởng lớn nhất đến biến động thị trường. LIME tập trung giải thích dự đoán cụ thể cho từng điểm dữ liệu, làm rõ vai trò của các yếu tố kinh tế trong từng trường hợp riêng lẻ. Sự kết hợp SHAP và LIME không chỉ tăng cường hiểu biết về cơ chế mô hình mà còn hỗ trợ nhà đầu tư và phân tích đưa ra quyết định dựa trên yếu tố kinh tế quan trọng.

4. Kết quả và thảo luận

4.1. Kết quả

Nghiên cứu này phân tích tác động của các bản tin kinh tế vĩ mô đến VN-Index, sử dụng Mô hình ngôn ngữ lớn (LLMs) như GPT và Gemini, cùng mô hình máy học XGBoost. Kết quả được chia thành các phần chính: độ tương đồng giữa đánh giá của GPT và Gemini, tương quan đơn biến giữa yếu tố kinh tế và biến động VN-Index, hiệu suất dự đoán của XGBoost, và phân tích giải thích từ SHAP cùng LIME. Hai mô hình GPT và Gemini đánh giá bản tin kinh tế cho thấy độ tương đồng trung bình 70%, thể hiện sự nhất quán tương đối cao trong phân tích tác động tin tức lên thị trường chứng khoán, dù khác biệt về thuật toán vẫn tồn tại.

⁶ GridSearchCV là một phương pháp tìm kiếm tối ưu các giá trị của siêu tham số trong mô hình học máy. Nó thực hiện thử tất cả các kết hợp có thể của siêu tham số đã được xác định trước, giúp tìm ra bộ tham số tốt nhất cho mô hình.

⁷ Gradient Boosting Tree là một phương pháp học máy mạnh mẽ, trong đó nhiều cây quyết định yếu (Weak Learners) được xây dựng theo chuỗi, mỗi cây mới cố gắng sửa chữa sai lầm của cây trước đó, nhằm tối ưu hóa độ chính xác dự đoán.

Phân tích tương quan đơn biến đánh giá mối quan hệ tuyến tính giữa các yếu tố kinh tế và VN-Index. Theo Bảng 2, "Tin tức kinh tế toàn cầu" (0,161), "Lãi suất" (0,136), và "Chính sách đầu tư nước ngoài" (0,115) có tương quan dương đáng kể, thúc đẩy VN-Index tăng. Ngược lại, "Tăng trưởng GDP" (-0,114) và "Sáp nhập và mua lại (M&A)" (-0,05) có tương quan âm, tác động tiêu cực đến chỉ số. Các yếu tố như "Báo cáo thu nhập và cổ tức" (0,001) và "Chính sách thuế" (0,001) cho thấy tương quan rất thấp, hầu như không ảnh hưởng tuyến tính. Tuy nhiên, các yếu tố này được giữ trong mô hình vì XGBoost xử lý tốt quan hệ phi tuyến tính và tương tác kinh tế gián tiếp (ví dụ: báo cáo thu nhập có thể kết hợp với ổn định chính trị để ảnh hưởng mạnh, như SHAP cho thấy +1,87), không phụ thuộc vào tương quan đơn biến.

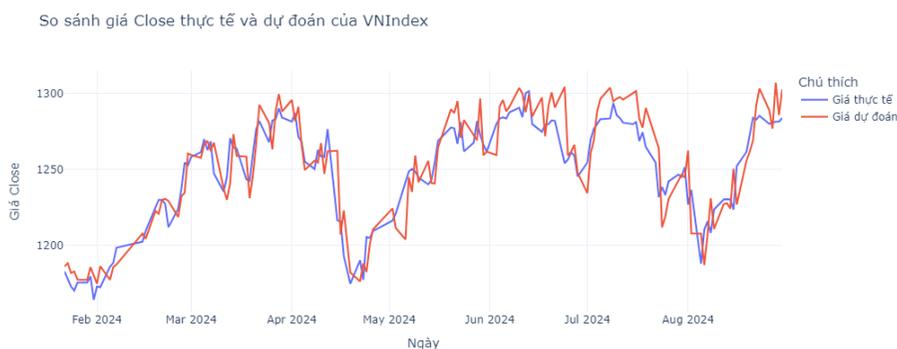
Bảng 2.

Tương quan Pearson giữa các yếu tố kinh tế và biến động VN-Index.

Các yếu tố vĩ mô	Pearson
Tăng trưởng GDP	-0,114***
Lãi suất	0,136***
Quy mô thị trường	0,048
Chính sách thuế	0,001
Lạm phát	0,062
Chính sách đầu tư nước ngoài	0,115***
Ổn định chính trị	0,102***
Báo cáo thu nhập và cổ tức	0,001
Thay đổi quy định	0,043
Sáp nhập và mua lại (M&A)	-0,05
Tình cảm trên mạng xã hội	0,075*
Tin tức kinh tế toàn cầu	0,161***
Đánh giá chung tổng thể mức độ tác động của bản tin	0,064

Ghi chú: * và *** lần lượt tương đương mức ý nghĩa thống kê 5% và 1%.

Mô hình XGBoost được áp dụng để dự đoán giá đóng cửa của VN-Index dựa trên các yếu tố kinh tế từ bản tin và giá mở cửa. Kết quả dự đoán được thể hiện trong Hình 4 với đường đồ biểu thị giá thực tế và đường xanh lam biểu thị giá dự đoán. Mô hình cho thấy khả năng theo dõi sát xu hướng biến động của VN-Index, đặc biệt trong các giai đoạn thị trường biến động mạnh.



Hình 4. So sánh giá đóng cửa thực tế và dự đoán của VN-Index bằng XGBoost.

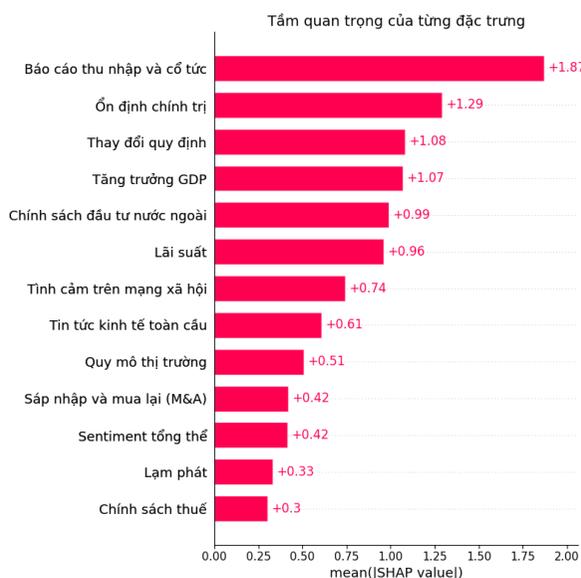
Hiệu suất định lượng của mô hình được đánh giá qua các chỉ số sai số tuyệt đối trung bình (Mean Absolute Error – MAE), sai số bình phương trung bình (Mean Squared Error – MSE), căn bậc hai của sai số bình phương trung bình (Root Mean Squared Error – RMSE), và hệ số xác định (Coefficient of Determination – R²) (Bảng 3). Kết quả cho thấy XGBoost có độ chính xác cao hơn so với Random Forest.

Bảng 3.

Hiệu suất của mô hình XGBoost.

Mô hình	MAE	RMSE	R2
XGBoost	10,88	15,71	0,9879
Random Forest	35,74	45,22	0,8998

Kỹ thuật SHAP (Lundberg & Lee, 2017) được sử dụng để đánh giá mức độ ảnh hưởng của từng yếu tố kinh tế đến dự đoán của mô hình XGBoost. Kết quả được thể hiện trong Hình 5, xếp hạng các yếu tố theo giá trị SHAP trung bình. "Bảo cáo thu nhập và cổ tức" dẫn đầu với giá trị SHAP trung bình là +1,87, tiếp theo là "Ổn định chính trị" (+1,29), "Thay đổi quy định" (+1,08), và "Tăng trưởng GDP" (+1,07), cho thấy đây là những yếu tố quan trọng nhất. Ngược lại, "Chính sách thuế" (+0,3) và "Lạm phát" (+0,33) có tác động thấp nhất. Các giá trị SHAP dương chỉ ra rằng sự gia tăng của các yếu tố này thường làm tăng giá đóng cửa dự đoán.



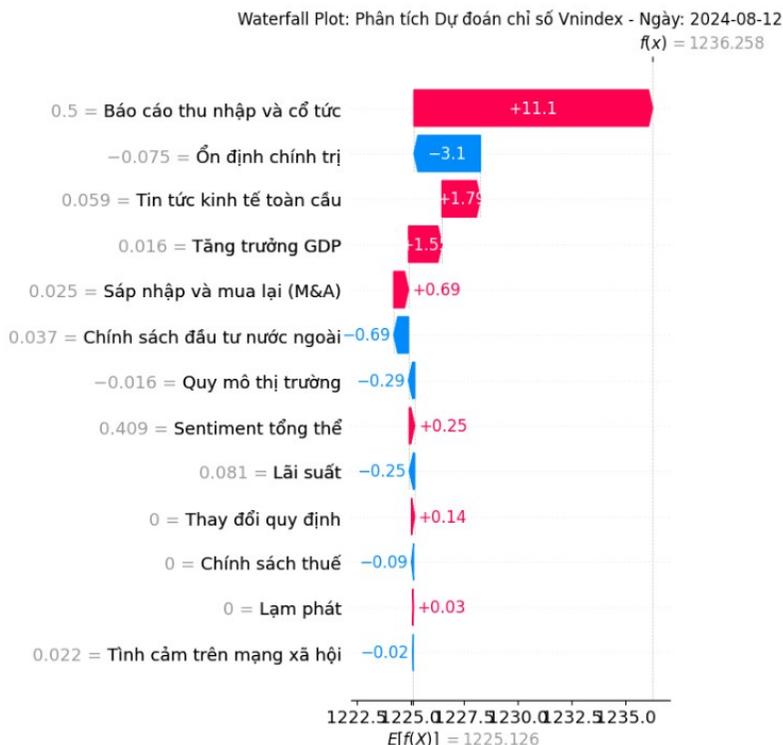
Hình 5. Biểu đồ SHAP summary plot cho mô hình XGBoost.

Kỹ thuật LIME (Ribeiro và cộng sự, 2016) được áp dụng để giải thích dự đoán của mô hình trong hai ngày giao dịch cụ thể: ngày giá tăng (12/08/2024) và ngày giá giảm sâu (03/04/2025).

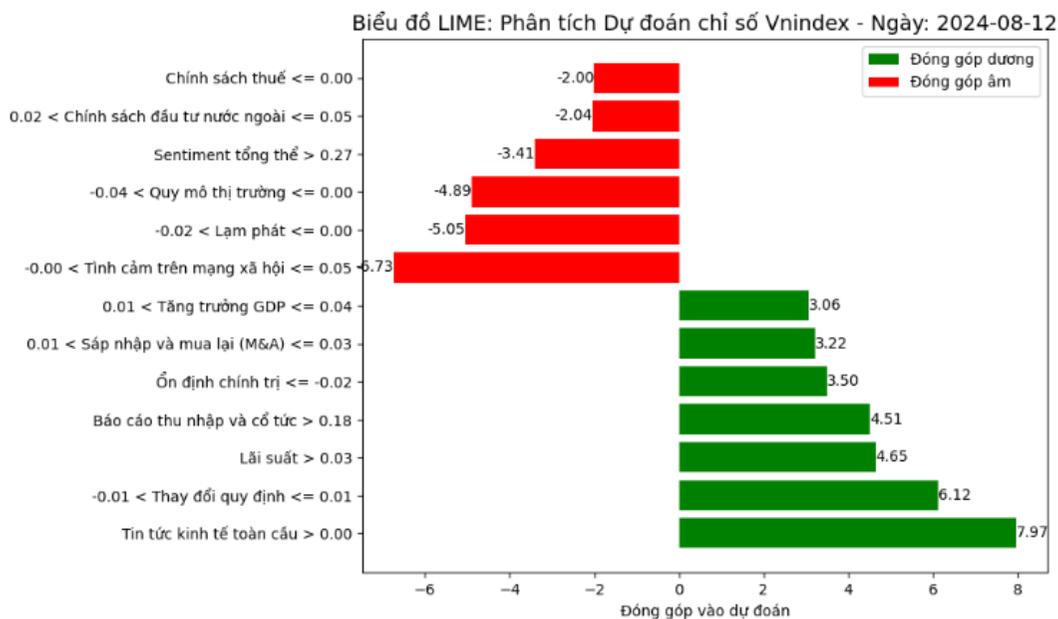
Trường hợp 1: Ngày 12/08/2024 – Giá tăng

Vào ngày 12/08/2024, VN-Index đóng cửa tại 1.230,28 điểm, tăng 6,64 điểm. Kết quả phân tích LIME được trình bày trong Hình 6 và Hình 7. Hình 6 cho thấy “Bảo cáo thu nhập và cổ tức” và “Tin

tức kinh tế toàn cầu” là hai yếu tố chính thúc đẩy dự đoán giá tăng, trong khi Hình 7 cung cấp chi tiết về giá trị và tác động của từng yếu tố.



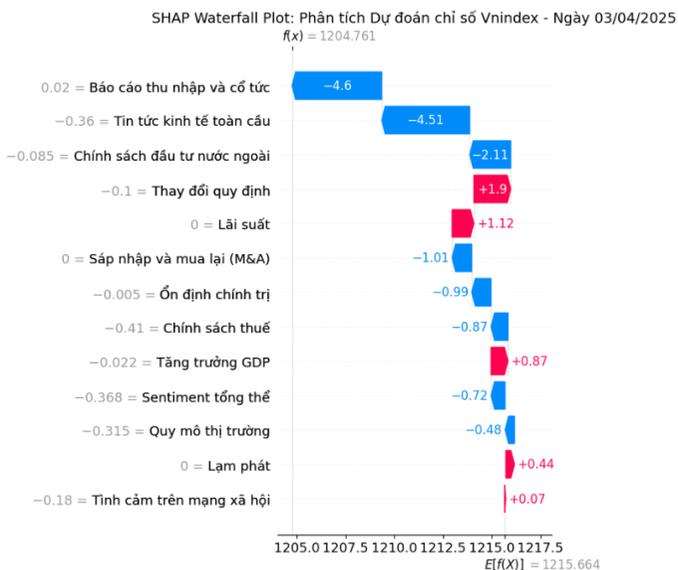
Hình 6. SHAP Waterfall Plot: Phân tích dự đoán chỉ số VN-Index cho ngày 12/08/2024



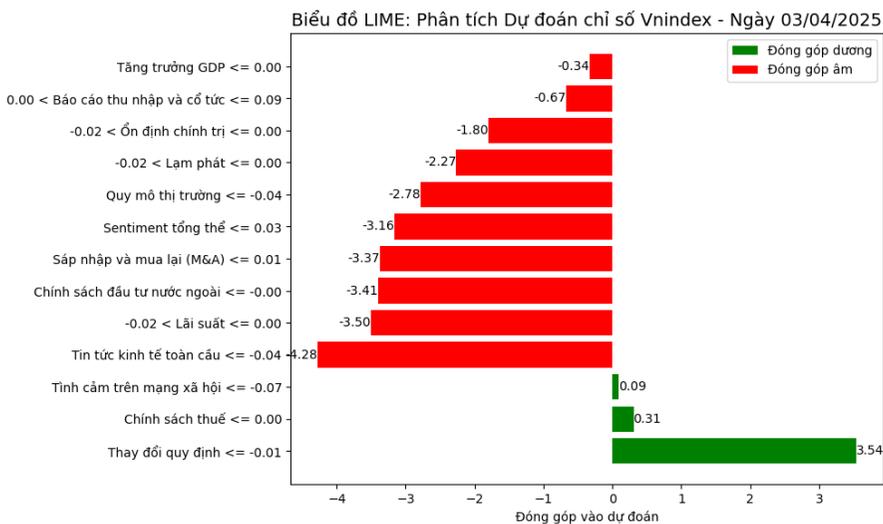
Hình 7. Biểu đồ chi tiết LIME cho ngày 12/08/2024

Trường hợp 2: Ngày 03/04/2025 – Giá giảm sâu (ví dụ minh họa ngoài mẫu)

Ngày 03/04/2025, sau sự kiện tổng thống Trump áp thuế đối ứng 46% lên Việt Nam, VN-Index giảm 87,99 điểm (-6,68%), đóng cửa tại 1.229,84 điểm. Để minh họa khả năng khái quát hóa của mô hình, chúng tôi áp dụng mô hình XGBoost đã được huấn luyện trên dữ liệu từ năm 2021 đến 2024 để dự báo và giải thích biến động giá đóng cửa của ngày này thông qua các công cụ trí tuệ nhân tạo giải thích được. Kết quả phân tích LIME được thể hiện trong Hình 8 và Hình 9. Hình 8 chỉ ra “Báo cáo thu nhập và cổ tức”, “Tin tức kinh tế toàn cầu” và “Chính sách đầu tư nước ngoài” có tác động tiêu cực lớn, phản ánh lo ngại của thị trường trước chính sách thuế quan. Hình 9 minh họa chi tiết mức độ ảnh hưởng của từng yếu tố.



Hình 8. SHAP Waterfall Plot: Phân tích dự đoán chỉ số VN-Index cho ngày 03/04/2025



Hình 9. Biểu đồ chi tiết LIME cho ngày 03/04/2025

4.2. Thảo luận

Kết quả nghiên cứu khẳng định tác động mạnh mẽ của tin tức kinh tế vĩ mô đến biến động VN-Index, đồng thời chứng minh hiệu quả kết hợp LLMs và mô hình máy học trong phân tích, dự đoán chứng khoán. Nghiên cứu hiện tại không chỉ tái khẳng định vai trò của các yếu tố kinh tế quan trọng mà còn đóng góp mới qua phương pháp tiên tiến, mở rộng khung lý thuyết tài chính. Nổi bật là áp dụng SHAP lượng hóa ảnh hưởng chủ đề báo chí, xác định chính xác tầm quan trọng của "Báo cáo thu nhập và cổ tức", "Chính sách đầu tư nước ngoài", và "Ổn định chính trị" trong định hình xu hướng VN-Index. Kết hợp SHAP với dữ liệu báo chí và XAI là tiếp cận chưa phổ biến trong nghiên cứu tài chính tại Việt Nam, đánh dấu bước tiến ứng dụng AI hiện đại, nâng cao độ chính xác dự đoán và minh bạch giải thích – vượt trội so với nghiên cứu truyền thống.

Nghiên cứu này cũng mở rộng khung phân tích mức độ tác động của các yếu tố vĩ mô từ các bài báo kinh tế trong mô hình tài chính qua LLMs (GPT, Gemini) trích xuất từ hơn 18.000 bài báo kinh tế (giai đoạn 2021–2024). Kết quả cho thấy mức độ tác động tổng thể và yếu tố kinh tế cụ thể ảnh hưởng rõ rệt đến VN-Index, được củng cố bởi XAI (SHAP, LIME) giải thích cơ chế. Tiếp cận này định lượng ảnh hưởng của tin tức, cung cấp cái nhìn sâu về định hình hành vi thị trường, nâng cao ứng dụng thực tiễn trong tài chính Việt Nam.

Quan trọng hơn, nghiên cứu này làm giàu lý thuyết “nhà đầu tư bất hợp lý” của De Long và cộng sự (1990), vốn cho rằng các nhà đầu tư không lý trí (Noise Traders) ra quyết định dựa trên thông tin không liên quan hoặc cảm xúc, gây ra biến động giá không phản ánh giá trị cơ bản của tài sản. Trong ngữ cảnh này, các bản tin kinh tế vĩ mô được xem là nguồn thông tin mà các nhà đầu tư không lý trí phản ứng, dẫn đến sự biến động của VN-Index. Kết quả thực nghiệm từ nghiên cứu cung cấp bằng chứng tin cậy rằng nhà đầu tư trên thị trường chứng khoán Việt Nam chịu sự dẫn dắt bởi tin tức, nhưng phản ứng của họ không đồng nhất mà phụ thuộc vào chủ đề cụ thể của bản tin. Chẳng hạn, "Báo cáo thu nhập và cổ tức" tạo ra tác động mạnh mẽ hơn so với "Lạm phát" hay "Chính sách thuế", cho thấy nhà đầu tư không phản ứng chung chung mà phân biệt rõ ràng giữa các loại tin tức khác nhau. Phát hiện này làm phong phú thêm lý thuyết “nhà đầu tư bất hợp lý” bằng cách chỉ ra rằng mức độ ảnh hưởng của tin tức không chỉ phụ thuộc vào sự hiện diện của nó mà còn vào nội dung cụ thể, đồng thời phù hợp với nghiên cứu của De Long và cộng sự (1990) về vai trò của thông tin nhiễu trong biến động thị trường.

Cụ thể, "Báo cáo thu nhập và cổ tức" được xác định là yếu tố quan trọng nhất ảnh hưởng đến VN-Index, phù hợp với lý thuyết tài chính cơ bản của Fama và cộng sự (1969). Trong bối cảnh thị trường mới nổi như Việt Nam, sức khỏe tài chính của các công ty niêm yết đóng vai trò cốt lõi trong việc định hình niềm tin nhà đầu tư, từ đó thúc đẩy chỉ số thị trường tăng trưởng. Các yếu tố "Chính sách đầu tư nước ngoài" và "Ổn định chính trị" cũng cho thấy tác động đáng kể, phản ánh sự nhạy cảm của TTCK Việt Nam với dòng vốn quốc tế và bối cảnh chính trị nội địa. Kết quả này được củng cố bởi nghiên cứu của Doukas và Lang (2003) và Boutchkova và cộng sự (2012), nhấn mạnh tầm quan trọng của các yếu tố ngoại lai trong các thị trường mới nổi. Ngược lại, các yếu tố như "Lạm phát", "Sáp nhập và mua lại" (M&A), và "Chính sách thuế" chỉ có ảnh hưởng hạn chế trong giai đoạn 2021–2024. Điều này có thể được giải thích bởi sự ổn định kinh tế vĩ mô của Việt Nam, với lạm phát được kiểm soát và quy mô hoạt động M&A chưa đủ lớn để tạo ra tác động mạnh, như đã được ghi nhận trong các nghiên cứu của Vương và cộng sự (2009), Su và cộng sự (2017).

Nghiên cứu này thành công trong việc đánh giá tác động tin tức bằng LLMs và dự đoán máy học (độ chính xác cao), với XAI (đặc biệt SHAP) làm rõ vai trò yếu tố chính như "Báo cáo thu nhập và cổ tức", "Chính sách đầu tư nước ngoài", "Ổn định chính trị". Về mặt lý thuyết, nghiên cứu hiện tại giúp mở rộng phân tích mức độ tác động của báo chí, giàu hóa lý thuyết "nhà đầu tư bất hợp lý" cho tài chính hành vi. Về mặt thực tiễn, nghiên cứu này cung cấp công cụ hỗ trợ quyết định nhà đầu tư, nâng cao phân tích dự đoán TTCK Việt Nam, tạo nền tảng cho nghiên cứu tương lai.

5. Kết luận

Nghiên cứu này đã xác nhận tác động mạnh mẽ của tin tức kinh tế vĩ mô đến biến động VN-Index, thông qua phân tích mức độ tác động từ 18.253 bài báo trên trang cafe.vn giai đoạn 2021–2024. Bằng cách sử dụng LLMs (GPT-4o-mini và Gemini-1.5-flash) để định lượng mức độ tác động của 12 yếu tố kinh tế chính và tích hợp vào mô hình XGBoost, nghiên cứu đạt $R^2=0,9879$, chứng minh khả năng dự báo vượt trội. Các công cụ XAI (SHAP và LIME) làm nổi bật “Báo cáo thu nhập và cổ tức”, “Chính sách đầu tư nước ngoài” và “Ổn định chính trị” là các yếu tố kinh tế chính thúc đẩy biến động thị trường, bổ sung bằng chứng thực nghiệm cho lý thuyết “nhà đầu tư bất hợp lý” (De Long và cộng sự, 1990) bằng cách chứng minh phản ứng nhà đầu tư không lý tính thay đổi theo loại tin tức, đặc biệt ở thị trường mới nổi như Việt Nam.

Về đóng góp lý thuyết, nghiên cứu này giúp mở rộng khung tài chính hành vi bằng cách tích hợp phân tích mức độ tác động của các bản tin vĩ mô, làm rõ mối liên hệ giữa tin tức kinh tế và biến động giá cổ phiếu. Về thực tiễn, nghiên cứu này cung cấp khung AI minh bạch hỗ trợ nhà đầu tư ra quyết định kinh tế kịp thời, như điều chỉnh danh mục dựa trên mức độ tác động từ chính sách đầu tư nước ngoài, góp phần phát triển thị trường chứng khoán Việt Nam trong bối cảnh hội nhập toàn cầu.

5.1. Hạn chế của nghiên cứu

Mặc dù đạt kết quả cao, nghiên cứu này vẫn còn tồn tại một số hạn chế liên quan đến quy mô và nguồn dữ liệu. *Thứ nhất*, việc dữ liệu chỉ từ trang cafe.vn có thể mang sai lệch từ nguồn, chưa bao quát đầy đủ các góc nhìn kinh tế từ các trang khác như vnexpress.net hay thanhnien.vn, dẫn đến thiếu đa dạng trong phân tích. *Thứ hai*, giai đoạn 2021–2024 chủ yếu đại diện cho chu kỳ phục hồi hậu COVID-19 với tăng trưởng kinh tế ổn định, nhưng chưa bao quát đầy đủ các chu kỳ kinh tế đa dạng như suy thoái hoặc lạm phát cao, có thể ảnh hưởng đến tính tổng quát hóa của mô hình. *Thứ ba*, mặc dù phương pháp đánh giá bản tin dựa trên mô hình ngôn ngữ lớn (LLM) cho thấy mức độ tương đồng khoảng 70% trong nhận định giữa mô hình chính và mô hình đối chứng; tuy nhiên, kết quả này chỉ mang ý nghĩa kiểm chứng tính nhất quán trong đánh giá, chứ không được sử dụng để kết hợp hay làm trung bình điểm số đầu vào. Do đó, vẫn tồn tại dư địa để cải thiện thiết kế câu lệnh và cơ chế đánh giá nhằm giảm thiểu thiên kiến chủ quan của LLM trong quá trình phân tích bản tin kinh tế. Những hạn chế này nhấn mạnh nhu cầu mở rộng để nâng cao tính đại diện kinh tế của nghiên cứu.

5.2. Hướng nghiên cứu tiếp theo

Để khắc phục hạn chế và làm phong phú thêm mục tiêu phân tích tác động kinh tế của tin tức vĩ mô đến thị trường chứng khoán, các hướng nghiên cứu tương lai có thể tập trung vào các yếu tố sau: *Thứ nhất*, mở rộng nguồn dữ liệu bằng cách tích hợp thêm các trang tin kinh tế đa dạng như

vnexpress.net hoặc dantri.com.vn, nhằm đánh giá các yếu tố toàn diện hơn. *Thứ hai*, các nghiên cứu tương lai có thể áp dụng khung phân tích cho các thị trường mới nổi khác như Trung Quốc hoặc Indonesia, để so sánh tác động kinh tế vĩ mô giữa các nền kinh tế hội nhập. *Thứ ba*, tích hợp dữ liệu thời gian thực để dự báo động thái kinh tế ngắn hạn, chẳng hạn như ảnh hưởng của chính sách tiền tệ toàn cầu đến dòng vốn đầu tư nước ngoài. Cuối cùng là khám phá tương tác sâu hơn giữa các yếu tố vĩ mô (ví dụ: GDP và lạm phát) qua mô hình kinh tế lượng nâng cao, nhằm bổ sung lý thuyết “nhà đầu tư bất hợp lý” với bằng chứng từ các chu kỳ kinh tế đầy đủ. Những hướng này sẽ nâng cao giá trị kinh tế của các nghiên cứu tiếp theo, hỗ trợ chính sách phát triển thị trường chứng khoán bền vững tại Việt Nam và khu vực.

Tài liệu tham khảo

- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Vega, C. (2007). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics*, 73(2), 251-277. <https://doi.org/10.1016/j.jinteco.2007.02.004>
- Bhattacharya, U., & Daouk, H. (2002). The world price of insider trading. *The Journal of Finance*, 57(1), 75-108. <http://www.jstor.org/stable/2697834>
- Biswas, R. (2023). Vietnam GDP growth improves in third quarter of 2023. *S&P Global Market Intelligence*. <https://www.spglobal.com/marketintelligence/en/mi/research-analysis/vietnam-gdp-growth-improves-in-third-quarter-of-2023-oct23.html>
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8. <https://doi.org/10.1016/j.jocs.2010.12.007>
- Boudoukh, J., & Richardson, M. (1993). Stock returns and inflation: A long-horizon perspective. *The American Economic Review*, 83(5), 1346-1355. <http://www.jstor.org/stable/2117566>
- Boutchkova, M., Doshi, H., Durnev, A., & Molchanov, A. (2012). Precarious politics and return volatility. *The Review of Financial Studies*, 25(4), 1111-1154. <https://doi.org/10.1093/rfs/hhr100>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32.
- Bunjaku, F. (2024). Decoding the stock market and GDP relationship over the long term: Implications for index fund investments. *Studies in Business and Economics*, 19, 49-59. <https://doi.org/10.2478/sbe-2024-0024>
- Cao, P. T. H., & Vo, D. H. (2025). Market responses to geopolitical risk and economic policy uncertainty: Evidence from Vietnam. *Heliyon*, 11(4), e42703. <https://doi.org/10.1016/j.heliyon.2025.e42703>
- Chen, C., Dongxing, W., Chunyan, H., & Xiaojie, Y. (2014). Exploiting social media for stock market prediction with factorization machine. *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, 2, 142-149. <https://doi.org/10.1109/WI-IAT.2014.91>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13–17 August 2016*, 785-794. <https://doi.org/10.1145/2939672.2939785>

- Chen, W., Liu, W., Zheng, J., & Zhang, X. (2025). Leveraging large language model as news sentiment predictor in stock markets: A knowledge-enhanced strategy. *Discover Computing*, 28(1), 74. <https://doi.org/10.1007/s10791-025-09573-7>
- Chikwira, C., & Mohammed, J. I. (2023). The impact of the stock market on liquidity and economic growth: Evidence of volatile market. *Economies*, 11(6). <https://doi.org/10.3390/economies11060155>
- Coates, J. (2007). The goals and promise of the Sarbanes-Oxley Act. *Journal of Economic Perspectives*, 21, 91-116. <https://doi.org/10.1257/jep.21.1.91>
- De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98(4), 703-738. <https://doi.org/10.1086/261703>
- Demirgüç-Kunt, A., & Levine, R. (1996). Stock markets, corporate finance, and economic growth: An overview. *The World Bank Economic Review*, 10(2), 223-239. <https://doi.org/10.1093/wber/10.2.223>
- Dinh, T. S., Bui, T., Bui, T. M. H., & Nguyen, V. B. (2017). Determinants of stock market development: The case of developing countries and Vietnam. *Journal of Economic Development*, 24, 32-53. <https://doi.org/10.24311/jed/2017.24.1.05>
- Doukas, J., & Lang, H. (2003). Foreign direct investment, diversification and firm performance. *Journal of International Business Studies*, 34, 153-172. <https://doi.org/10.1057/palgrave.jibs.8400014>
- Downs, T., & Hendershott, P. H. (1987). Tax policy and stock prices. *National Tax Journal*, 40(2), 183-190. <https://doi.org/10.1086/NTJ41788656>
- Ehrmann, M., & Fratzscher, M. (2005). Equal size, equal role? Interest rate interdependence between the euro area and the United States. *The Economic Journal*, 115(506), 928-948. <http://www.jstor.org/stable/3590356>
- Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review*, 10(1), 1-21. <https://doi.org/10.2307/2525569>
- Hoi, L. Q., Thu, N. T. H., Hung, N. X., Uyen, P. T., Huong, T. T., Minh, T. T. H., & Anh, H. T. P. (2024). The impact of the global minimum tax on Vietnam's foreign direct investment attraction. *Asia and the Global Economy*, 4(2), 100090. <https://doi.org/10.1016/j.aglobe.2024.100090>
- Nguyen, H. D., & Le, V. L. (2024). Impacts of inflation on the Vietnamese stock market in economic turbulence. *The VMOST Journal of Social Sciences and Humanities*, 66(1), 16-20. [https://doi.org/10.31276/VMOSTJOSSH.66\(1\).16-20](https://doi.org/10.31276/VMOSTJOSSH.66(1).16-20)
- Jensen, M. C., & Ruback, R. S. (1983). The market for corporate control: The scientific evidence. *Journal of Financial Economics*, 11(1), 5-50. [https://doi.org/10.1016/0304-405X\(83\)90004-1](https://doi.org/10.1016/0304-405X(83)90004-1)
- Khang, P. Q., Kaczmarczyk, K., Tutak, P., Golec, P., Kuziak, K., Depczyński, R., Hernes, M., & Rot, A. (2021). Machine learning for liquidity prediction on Vietnamese stock market. *Procedia Computer Science*, 192, 3590-3597. <https://doi.org/10.1016/j.procs.2021.09.132>
- Kim, J. (2023). Stock market reaction to US interest rate hike: Evidence from an emerging market. *Heliyon*, 9(5), e15758. <https://doi.org/10.1016/j.heliyon.2023.e15758>

- Kirtac, K., & Germano, G. (2024). Sentiment trading with large language models. *Finance Research Letters*, 62, 105227. <https://doi.org/10.1016/j.frl.2024.105227>
- Kwon, B., Park, T., Rungcharoenkitkul, P., & Smets, F. (2025). Parsing the pulse: Decomposing macroeconomic sentiment with LLMs. BIS Working Papers (Issue 1294). Bank for International Settlements. <https://EconPapers.repec.org/RePEc:bis:biswps:1294>
- Lee, B. S. (2010). Stock returns and inflation revisited: An evaluation of the inflation illusion hypothesis. *Journal of Banking & Finance*, 34(6), 1257-1273. <https://doi.org/10.1016/j.jbankfin.2009.11.023>
- Levine, R. (1997). Financial development and economic growth: Views and agenda. *Journal of Economic Literature*, 35(2), 688-726. <http://www.jstor.org/stable/2729790>
- Li, X., Wu, P., & Wang, W. (2020). Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong. *Information Processing & Management*, 57(5), 102212. <https://doi.org/10.1016/j.ipm.2020.102212>
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
- Moeller, S. B., Schlingemann, F. P., & Stulz, R. M. (2004). Firm size and the gains from acquisitions. *Journal of Financial Economics*, 73(2), 201-228. <https://doi.org/10.1016/j.jfineco.2003.07.002>
- Ngo, N., Nguyen, H., Nguyen, Y., & Le, S. (2024). How does the Vietnamese stock market react when the Fed gives an announcement in time at the zero lower bound?. *Heliyon*, 10, e40047. <https://doi.org/10.1016/j.heliyon.2024.e40047>
- Nguyen, H. T. T., Tram, H. T. X., & Nguyen, L. T. T. (2023). Interest rates and systemic risk: Evidence from the Vietnamese economy. *The Journal of Economic Asymmetries*, 27, e00294. <https://doi.org/10.1016/j.jeca.2023.e00294>
- Nguyen, V. C., & Nguyen, T. T. (2022). Dependence between Chinese stock market and Vietnamese stock market during the Covid-19 pandemic. *Heliyon*, 8(10), e11090. <https://doi.org/10.1016/j.heliyon.2022.e11090>
- Pástor, L., & Veronesi, P. (2013). Political uncertainty and risk premia. *Journal of Financial Economics*, 110(3), 520–545. <https://doi.org/10.1016/j.jfineco.2013.08.007>
- Phan, T. K. H., Hoai, N., & Tran. (2019). Dividend policy and stock price volatility in an emerging market: Does ownership structure matter? *Cogent Economics & Finance*, 7(1), 1637051. <https://doi.org/10.1080/23322039.2019.1637051>
- Phuoc, T., Anh, P. T. K., Tam, P. H., & Nguyen, C. V. (2024). Applying machine learning algorithms to predict the stock price trend in the stock market – The case of Vietnam. *Humanities and Social Sciences Communications*, 11(1), 393. <https://doi.org/10.1057/s41599-024-02807-x>
- Ónozó, L. R., Arthur, F. V., & Gyires-Tóth, B. (2024). Leveraging LLMs for financial news analysis and macroeconomic indicator nowcasting. *IEEE Access*, 12, 160529–160547. <https://doi.org/10.1109/ACCESS.2024.3488363>

- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, 1135-1144. Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939778>
- Ritter, J. (2004). Economic growth and equity returns. *Pacific-Basin Finance Journal*, 13, 489-503. <https://doi.org/10.1016/j.pacfin.2005.07.001>
- Santow, L. J., & Gordon, M. J. (1962). *The investment, financing, and valuation of the corporation*. <https://api.semanticscholar.org/CorpusID:203555679>
- Su, D. T., Bui, T., Bui, T. M. H., & Nguyen, V. B. (2017). Determinants of stock market development: The case of developing countries and Vietnam. *Journal of Economic Development*, 24, 32-53.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139-1168. <https://doi.org/10.1111/j.1540-6261.2007.01232.x>
- Thanh, N. T., & Linh, D. T. (2016). Impacts of monetary policy on Vietnam stock price. *Proceedings of the International Conference on Electronics, Mechanics, Culture and Medicine*, 136-142. <https://doi.org/10.2991/emcm-15.2016.26>
- Van, C. B., Cao, T. D., Thi, T. N., Dung, H. P., Minh, H. D., An, L. N. H., & Hong, T. P. (2024). Studying the influence of Vietnamese social media on Vietnamese stock market to forecast market trends. In R. Silhavy & P. Silhavy (Eds.), *Artificial Intelligence Algorithm Design for Systems* (pp. 437-452). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-70518-2_39
- Vuong, Q.-H., Tran, T. D., & Nguyen, T. T. H. (2009). M&A market in Vietnam's transition economy. *Corporate Governance & Finance EJournal*. <https://api.semanticscholar.org/CorpusID:166249757>
- Wang, S., & Mayes, D. G. (2012). Monetary policy announcements and stock reactions: An international comparison. *The North American Journal of Economics and Finance*, 23(2), 145-164. <https://doi.org/10.1016/j.najef.2012.02.002>
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Richter, B., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, & A. Oh (Eds.), *Advances in Neural Information Processing Systems* (Vol. 35, pp. 24824-24837). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf
- World Bank. (2023). *Vietnam's economy forecast to grow 6.3 percent in 2023*. World Bank. <https://www.worldbank.org/en/news/press-release/2023/03/13/vietnam-s-economy-forecast-to-grow-by-6-3-in-2023-world-bank-report-says>