



Original Article

Forecasting stock price using
financial and non-financial information:
Application of event study and random forest analysis

Vu Thi Loan*, Nguyen Thi Anh Thu, Nong Thi Huong Ly, Vu Mai Linh

*VNU University of Economics and Business
No. 144 Xuan Thuy Road, Cau Giay District, Hanoi, Vietnam*

Received: March 10, 2023

Revised: May 19, 2023; Accepted: June 25, 2023

Abstract: This study aims to assess the effects of new information on the movements of stock returns in Vietnam. The Event Study approach and the Random Forest algorithm are applied for stock price prediction. The data sample is collected from 84 companies registered on the Hochiminh Stock Exchange (HOSE) between January 2017 and December 2021. The research analysis results disclose the abnormal returns around the dates of information release. Specifically, abnormal returns as a response to financial information are more significant than abnormal returns caused by non-financial information. In addition, the stock price movements can be predicted from 8 variables including information, market capitalization, three financial indicators, and three macroeconomic variables.

Keywords: Information, financial information, non-financial information, Event Study, Random Forest.

* Corresponding author

E-mail address: loanvu.ktt@vnu.edu.vn

<https://doi.org/10.57110/jebvn.v3i1.178>

Copyright © 2023 The author(s)

Licensing: This article is published under a CC BY-NC 4.0 license.

Dự báo giá cổ phiếu trước các thông tin tài chính và phi tài chính: Áp dụng phân tích Event Study và Random Forest

Vũ Thị Loan*, Nguyễn Thị Anh Thu, Nông Thị Hương Ly, Vũ Mai Linh

*Trường Đại học Kinh tế - Đại học Quốc gia Hà Nội
No. 144 Đường Xuân Thủy, Quận Cầu Giấy, Hà Nội, Việt Nam*

Nhận ngày 10 tháng 3 năm 2023

Chỉnh sửa ngày 19 tháng 5 năm 2023; Chấp nhận đăng ngày 25 tháng 6 năm 2023

Tóm tắt: Bài viết đề xuất và xem xét khía cạnh thông tin có ảnh hưởng như thế nào đến biến động giá cổ phiếu bằng phương pháp Event Study (nghiên cứu sự kiện), từ đó dùng thuật toán học máy Random Forest để dự báo biến động tiếp theo của giá cổ phiếu. Với bộ dữ liệu thu thập từ 84 doanh nghiệp niêm yết trên Sở Giao dịch Chứng khoán Thành phố Hồ Chí Minh (HOSE) từ tháng 1/2017 đến tháng 12/2021, kết quả nghiên cứu tìm thấy lợi nhuận bất thường xung quanh ngày công bố thông tin về doanh nghiệp. Đặc biệt, lợi nhuận bất thường xung quanh ngày có thông tin phi tài chính đáng kể hơn so với ngày thông tin tài chính của doanh nghiệp được công bố. Bên cạnh đó, bài viết đề xuất mô hình dự đoán biến động giá cổ phiếu dựa trên 8 yếu tố gồm: thông tin, vốn hóa thị trường, 3 chỉ số tài chính và 3 yếu tố thuộc nền kinh tế vĩ mô.

Từ khóa: Thông tin, thông tin tài chính, thông tin phi tài chính, Event Study, Random Forest.

1. Giới thiệu

Dự đoán thị trường chứng khoán là một trong những bài toán khó do tính chất phi tuyến tính và phức tạp của nó. Trên thực tế có rất nhiều yếu tố trong nền kinh tế tác động lên biến động giá của một cổ phiếu và đã có nhiều cố gắng trong việc định lượng tác động của các yếu tố đó, từ đó dự đoán biến động tiếp theo trên thị trường chứng khoán nhưng vẫn còn vô số mâu thuẫn về độ tin cậy của những dự đoán. Một số phương pháp đã được sử dụng như phân tích cơ bản và kỹ thuật hay các thuật toán học máy SVM, ANN để dự đoán giá cổ phiếu trong tương lai. Tuy nhiên, các phương pháp học máy kết hợp vẫn chưa khai thác nhiều trong lĩnh vực này, đặc biệt là ở thị trường Việt Nam.

Dự đoán thị trường chứng khoán (TTCK) là một trong những bài toán khó do tính chất phi tuyến tính và phức tạp của nó. Trên thực tế có rất

nhiều yếu tố trong nền kinh tế tác động đến biến động giá của một cổ phiếu. Nhiều nghiên cứu đã cố gắng định lượng tác động của các yếu tố đó để dự đoán biến động tiếp theo trên TTCK, tuy nhiên vẫn tồn tại mâu thuẫn về độ tin cậy của các dự đoán. Một số phương pháp như phân tích cơ bản và kỹ thuật hay các thuật toán học máy SVM, ANN được sử dụng để dự đoán giá cổ phiếu trong tương lai. Tuy nhiên, các phương pháp học máy kết hợp vẫn chưa được khai thác nhiều trong lĩnh vực này, đặc biệt là ở thị trường Việt Nam.

Mặt khác, giá cổ phiếu chịu tác động không nhỏ từ các thông tin truyền thông về doanh nghiệp (Son, 2018). Nếu một cổ phiếu trên sàn vướng phải những tin tức xấu thì sẽ khiến tâm lý của nhà đầu tư bị ảnh hưởng, từ đó mất thiện cảm với doanh nghiệp và quyết định bán cổ phiếu để tránh rủi ro rớt giá. Thông tin được phân loại

* Tác giả liên hệ

Địa chỉ email: loanvu.ktt@vnu.edu.vn

<https://doi.org/10.57110/jebvn.v3i1.178>

Bản quyền © 2023 (Các) tác giả

Bài báo này được xuất bản theo CC BY-NC 4.0 license.

thành thông tin tài chính và thông tin phi tài chính, đã được một số tác giả xem xét như Dutescu (2000), Filip và cộng sự (2012), Moreno và Zhang (2021)... Thông tin tài chính là tất cả những thông tin có được thông qua báo cáo tài chính do doanh nghiệp cung cấp, trong đó thông tin về trả cổ tức và lợi nhuận là chỉ số chính. Thông tin phi tài chính là tất cả những thông tin không phải thông tin tài chính, được thu thập từ nhiều nguồn khác nhau (chất lượng đầu vào, hệ thống vận hành, sự hài lòng khách hàng...). Khảo sát của Hiệp hội Kế toán Công chứng Anh Quốc - ACCA (2013) trên 18 quốc gia cho thấy nguồn thông tin phi tài chính quan trọng nhất là từ báo cáo trách nhiệm xã hội/ phát triển bền vững và báo cáo thường niên. Việc xác định các thông tin giúp giảm rủi ro giữa những nhà đầu tư có và không có thông tin (Usman & Tandelilin, 2014). Tuy nhiên, hiện tại ở Việt Nam vẫn chưa có nhiều nghiên cứu làm rõ khía cạnh này.

Với khoảng trống nghiên cứu trên, bài viết này xem xét biến động giá cổ phiếu trước các thông tin tài chính và phi tài chính trên TTCK Việt Nam, từ đó đóng góp kết quả nghiên cứu gồm: (1) Khẳng định sự tồn tại lợi nhuận bất thường tích lũy xung quanh sự kiện thông báo; (2) Mức độ ảnh hưởng của thông tin phi tài chính đáng kể hơn so với thông tin tài chính; (3) Cung cấp mô hình dự báo giá cổ phiếu dựa trên 8 yếu tố gồm: thông tin, vốn hóa thị trường, 3 chỉ số tài chính (P/E, ROS, Beta) và 3 yếu tố thuộc nền kinh tế vĩ mô.

2. Tổng quan nghiên cứu

Tại Việt Nam, một số nghiên cứu đã tiến hành xem xét các yếu tố ảnh hưởng đến giá cổ phiếu niêm yết trên TTCK trong nước. Điển hình, Vo và Dang (2016) phân tích phản ứng của thị trường thông qua giá cổ phiếu đối với thông báo thay đổi cổ phiếu trong tương lai, sử dụng phương pháp nghiên cứu sự kiện trong tài chính kinh tế. Truong (2014) chỉ ra có 4 yếu tố tác động đến sự thay đổi giá cổ phiếu niêm yết trên sàn HOSE bao gồm chuỗi giá, lợi nhuận trên mỗi cổ phiếu (EPS) của 20 cổ phiếu, lãi suất cho vay, Tỷ giá USD/VND, giá vàng và chỉ số giá tiêu dùng (CPI). Kết quả phân tích hồi quy cho thấy EPS và tỷ giá USD/VND suất có tương quan thuận với sự thay đổi giá của cổ phiếu.

Ở nước ngoài, Yu và Huarng (2020) đề xuất một phương pháp nghiên cứu sự kiện mới để dự báo lợi nhuận bất thường cho các doanh nghiệp riêng lẻ nhằm khắc phục hạn chế của các phương pháp tiếp cận liên quan hiện tại. Tác giả xem xét hai tiêu chí để xác định liệu có lợi nhuận bất thường hay không. Hiệu suất cổ phiếu riêng lẻ và hiệu suất tương đối giữa một cổ phiếu riêng lẻ và thị trường. Để khám phá mối quan hệ giữa chủ đề liên quan đến chính trị và biến động của chỉ số FTSE 100, Nisar và Yeung (2018) đã tiến hành nghiên cứu sự kiện trong thời gian ngắn về một sự kiện chính trị có trụ sở tại Anh. Bogdan và cộng sự (2010) dự đoán sự gia tăng đáng kể và mức độ phù hợp tích cực của các mục thông tin và sự không đồng nhất đáng kể của việc điều chỉnh giá, sau đó so sánh với một tiêu chuẩn cụ thể cho một thị trường phát triển hơn - Sở Giao dịch Chứng khoán Madrid. Trên thế giới, các nghiên cứu thực nghiệm xác định các yếu tố tài chính và phi tài chính cho thấy mức độ ảnh hưởng của lĩnh vực hoạt động đến tỷ suất lợi nhuận trên vốn của doanh nghiệp. Một trong những lý thuyết cơ bản về mối quan hệ giữa hai biến này được đề cập trong nghiên cứu của Robu và cộng sự (2013). Cụ thể, tác giả kết luận rằng khi lựa chọn các cổ phiếu mang lại lợi nhuận cao với rủi ro chấp nhận được, các nhà đầu tư chú ý đến các yếu tố chính có thể ảnh hưởng đến hiệu suất của cổ phiếu.

Một điểm chung của các nghiên cứu này đều chưa có sự phân biệt theo ngành hoặc quy mô công ty và chưa có tính cập nhật khi hầu hết đều nằm trong giai đoạn từ 3 năm trước và xa hơn. Ngoài ra, thuật ngữ về thông tin tài chính và thông tin phi tài chính chưa được làm rõ trong các bài nghiên cứu tại Việt Nam. Bên cạnh đó, nhóm nghiên cứu sử dụng hai phương pháp chưa được sử dụng rộng rãi trong các nghiên cứu ở Việt Nam, kết hợp phương pháp sự kiện Event Study và phương pháp Random Forest để có cái nhìn tổng quát hơn về các nhân tố ảnh hưởng đến giá cổ phiếu.

3. Số liệu sử dụng và phương pháp nghiên cứu sự kiện (Event Study)

3.1. Số liệu sử dụng

Nghiên cứu lựa chọn 84 doanh nghiệp có giá trị vốn hóa lớn, trung bình và nhỏ niêm yết trên sàn HOSE từ tháng 1/2017 đến tháng 12/2021.

Các doanh nghiệp này có xu hướng tự đăng tải thông tin tài chính và phi tài chính để truyền tải đến nhà đầu tư kịp thời. Việc công bố đầy đủ và chính xác thông tin tạo được niềm tin cho cổ đông và thu hút nhà đầu tư dù các doanh nghiệp có kích cỡ vốn hóa nhỏ hay lớn. Nghiên cứu lựa chọn giai đoạn 2017-2021 vì TTCK trong giai đoạn này chịu nhiều biến động lớn, đồng thời các doanh nghiệp có chính sách tăng cường công bố thông tin để đáp ứng yêu cầu của thị trường và cổ đông. Các doanh nghiệp niêm yết đã chú ý đến việc công bố thông tin đầy đủ và kịp thời để tạo lòng tin và gia tăng giá trị của cổ phiếu trên TTCK Việt Nam - điều này giúp mẫu nghiên cứu mang tính đa dạng, bao gồm cả thông tin tài chính và phi tài chính.

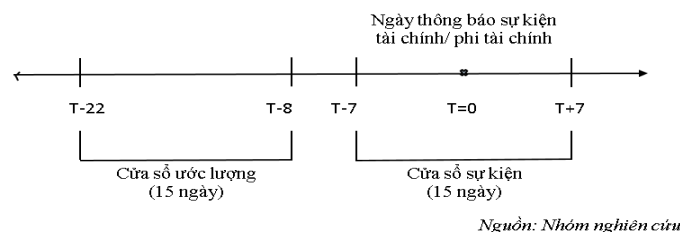
3.2. Phương pháp Event Study

Phương pháp Event Study nghiên cứu sự tác động của một sự kiện, từng được đề cập trong các nghiên cứu của Bonnier và Bruner (1989), MacKinlay (1997), Vo và Dang (2016). Các bước thực hiện trong nghiên cứu như sau: (i) Xác định sự kiện phân tích, (ii) Lựa chọn mô hình đo lường phản ứng giá với sự kiện, (iii) Ước lượng lợi nhuận vượt trội (lợi nhuận bất thường), (iv) Sắp xếp và phân nhóm lợi nhuận vượt trội, (v) Phân tích kết quả. Chính vì gắn với thời điểm công bố của thông tin nên phương pháp Event

Study được lựa chọn và kết hợp với một mô hình dự báo để giúp các nhà đầu tư trên TTCK Việt Nam đo lường ảnh hưởng của việc công bố thông tin đến sự thay đổi giá và đưa ra quyết định đầu tư đúng đắn hơn.

Một cách cụ thể, khung thời gian nghiên cứu gồm hai khái niệm: Khung cửa sổ sự kiện và Khung cửa sổ ước lượng. Trong đó, khung cửa sổ sự kiện nhằm đánh giá rõ hơn sự tác động giữa biến nghiên cứu và mức độ phản ứng của thị trường khi thông tin được công bố, coi $t = 0$ là ngày công bố thông tin. Khung cửa sổ ước lượng giúp ước lượng hệ số của mô hình thị trường trong việc tính toán thu nhập bất thường của cổ phiếu. Độ dài của khung sự kiện là 15 ngày¹, tức là giai đoạn bắt đầu từ 7 ngày trước ($t = -7$) và kéo dài tới 7 ngày sau của sự kiện ($t = +7$).

Độ dài của khung cửa sổ ước lượng nhằm xác định thu nhập kỳ vọng của cổ phiếu được lựa chọn là 15 ngày trước khung sự kiện² (được tính từ ngày $t = -8$ đến $t = -22$). Việc nhóm lựa chọn khung cửa sổ ước lượng đủ dài để giảm thiểu tác động của biến động ngắn hạn và đánh giá mức thu nhập kỳ vọng của cổ phiếu chính xác hơn. Việc áp dụng khung ước lượng cũng giúp giảm thiểu ảnh hưởng của các hoạt động công bố thông tin của doanh nghiệp tới giá cổ phiếu trong thời kỳ quan sát.



Nguồn: Nhóm nghiên cứu

3.3. Mô hình nghiên cứu

Lợi nhuận kỳ vọng ($ER_{i,t}$): được tính toán dựa trên mô hình yếu tố thị trường và rủi ro:

$$E(R_{i,t}) = \alpha_i + \beta_i R_{m,t}$$

Trong đó: $E(R_{i,t})$ là lợi nhuận kỳ vọng của cổ phiếu i ngày t ; $R_{m,t}$ là lợi nhuận của danh mục thị trường ngày t ; α_i là tham số Alpha của cổ phiếu i ; β_i là tham số beta của cổ phiếu i .

Tiếp theo, để xác định lợi nhuận bất thường, cần phải so sánh mức độ thay đổi giữa lợi nhuận kỳ vọng và lợi nhuận thực tế khi sự kiện xảy ra.

Lợi nhuận bất thường ($AR_{i,t}$): Được tính toán dựa trên nghiên cứu của Võ Xuân Vinh và Đặng Bửu Kiêm (2014, 2016a). Cụ thể, AR được tính trong khung cửa sổ sự kiện 31 ngày (-15; +15):

$$AR_{i,t} = R_{i,t} - E(R_{i,t})$$

¹ Độ dài khung sự kiện tham khảo từ các nghiên cứu của Gurgul và cộng sự (2003), Dasilas và Leventis (2011).

² Không có tiêu chuẩn cho thời gian khung cửa sổ ước

lượng, nhưng một số nghiên cứu sử dụng khoảng thời gian dưới 100 ngày (tối thiểu 10 ngày) đề xuất trong nghiên cứu của Puspitaningtyas (2019).

Trong đó: $AR_{i,t}$ là lợi nhuận bất thường của cổ phiếu i ngày t .

Kiểm định sự tồn tại lợi nhuận bất thường và lợi nhuận bất thường trung bình tích lũy: Tổng hợp các kết quả lợi nhuận bất thường của 622 quan sát trong 15 ngày, tính lợi nhuận bất thường trung bình theo từng ngày diễn ra:

$$AAR_t = 1/n \sum_{i=0}^n AR_{i,t}$$

Trong đó: AAR_t là lợi nhuận bất thường trung bình của tất cả cổ phiếu ngày t ; N là số sự kiện tài chính, phi tài chính diễn ra.

Lợi nhuận bất thường trung bình (AAR) được phân tích trong khung cửa sổ sự kiện là $(-3, +3)$, giúp nghiên cứu đánh giá rõ hơn sự tác động của biến nghiên cứu, bao gồm cả những trường hợp bị rò rỉ thông tin trước ngày công bố.

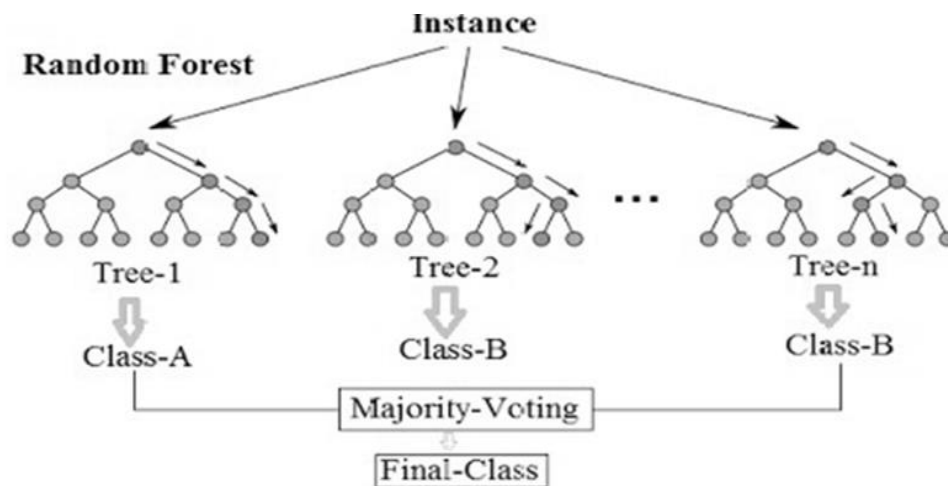
Nếu AAR được tích lũy trong một khoảng thời gian cụ thể (xung quanh mốc ngày công bố thông tin), ta có thể nhận được lợi nhuận bất thường trung bình tích lũy như sau:

$$CAAR_{t1,t2} = \sum_{t=t1}^{t2} AAR_t$$

Trong đó: $CAAR_{t1,t2}$ là lợi nhuận bất thường trung bình tích lũy từ ngày $t1$ đến ngày $t2$.

3.4. Thuật toán random forest

Random forest là một phương pháp thống kê mô hình hóa bằng máy dùng để phục vụ các mục đích phân loại, tính hồi quy và các nhiệm vụ khác bằng cách xây dựng nhiều cây quyết định. Trong đó, cây quyết định là mô hình dự báo cấu trúc hình cây, phân loại dữ liệu bằng cách đi từ gốc của cây, đi theo các nhánh cho tới khi gặp phân lá, đó là các lớp của dữ kiện đang xét.



Hình 1: Mô hình Random Forest

Nguồn: Le (2017).

Mục tiêu của nghiên cứu là xem xét tác động của 8 biến độc lập Thông tin, VHTT, P/E, ROS, Beta và 3 yếu tố thuộc kinh tế vĩ mô đến biến phụ thuộc CAR đã được một số tác giả nghiên cứu (Vo, 2016; Huynh, 2016). Các biến nhóm đều độc lập với nhau để tránh hiện tượng đa cộng tuyến, đảm bảo tính thực tiễn và có khả năng áp dụng khả thi để đảm bảo tính ứng dụng và giá trị của mô hình. Thêm vào đó, các biến này đều có thể dễ dàng tìm kiếm trên các trang phân tích chứng khoán hoặc các trang thống kê của chính phủ, hoặc rất dễ tính toán.

Biến phụ thuộc - Lợi nhuận bất thường tích lũy (CAR): CAR được nghiên cứu trong khung sự kiện là $[-3, 3]$ vì trong khoảng thời gian này,

biến động giá cổ phiếu có ý nghĩa nhất. Để CAR của một sự kiện có ý nghĩa tổng quát nhất, nhóm nghiên cứu lấy giá trị trung bình trong khoảng thời gian từ ngày -3 trước sự kiện đến ngày thứ 3 sau sự kiện.

Biến độc lập - Biến thông tin: Các thông tin tài chính được chứng minh là có ảnh hưởng khá khiêm tốn đến sự biến động giá cổ phiếu (Moreno và Zhang, 2021), trong khi thông tin phi tài chính dường như dẫn đến những thay đổi trong hầu hết các cổ phiếu ở một quốc gia nhất định hoặc thậm chí trên toàn thế giới (Filip & Nistor, 2012). Biến thông tin về cơ bản là biến định tính nên phải lượng hóa chúng bằng biến giả. Theo đó, biến giả 0 cho biết không có thuộc

tính và biến giả 1 là có thuộc tính. Ví dụ, biến giới tính có thể lượng hóa như sau: nữ = 1 và nam = 0, hoặc ngược lại. Cuối cùng, nhóm nghiên cứu quyết định đặt biến thông tin tài chính bằng 0 và thông tin phi tài chính bằng 1.

Thông tin tài chính	0
Thông tin phi tài chính	1

Biến vốn hóa thị trường: Thể hiện quy mô hoạt động, thị giá cổ phiếu doanh nghiệp và cả vị thế trong ngành, tiềm năng tăng trưởng hay sự đánh giá của thị trường. Việc phân loại nhóm cổ phiếu dựa trên quy mô vốn hóa thị trường giúp nhà đầu tư chọn đúng cổ phiếu cho mục tiêu đầu tư và khả năng chấp nhận rủi ro. Với những mã có quy mô vốn hóa lớn, các nhà đầu tư sẽ hạn chế được rủi ro thanh khoản của cổ phiếu. Có 3 nhóm vốn hóa cơ bản trên thị trường:

- Nhóm doanh nghiệp vốn hóa lớn: $MCK > 10.000$ tỷ đồng;
- Nhóm doanh nghiệp vốn hóa trung bình: $1.000 < MCK < 10.000$ tỷ đồng;
- Nhóm doanh nghiệp vốn hóa nhỏ và siêu nhỏ: $MCK < 1.000$ tỷ đồng.

Các biến kinh tế vĩ mô như P/E, ROS, Beta, lãi suất thị trường, GDP, tỷ giá hối đoái: Kết quả phân tích của Truong (2014) cho thấy EPS và tỷ giá USD/VND có tương quan tỷ lệ thuận với sự thay đổi giá của cổ phiếu. Ngược lại, biến động của tỷ lệ lạm phát có mối tương quan nghịch với tỷ suất sinh lời của cổ phiếu. Cùng kết quả với

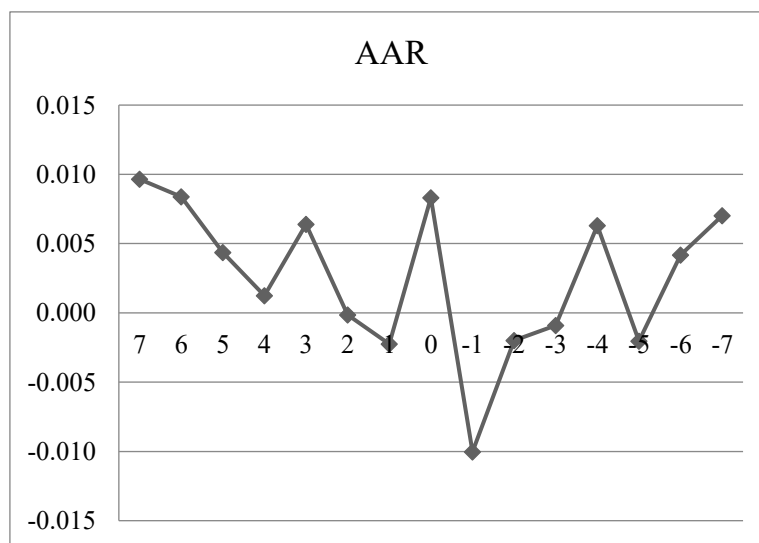
Truong (2014), nghiên cứu của Nguyen (2020) cho thấy, giá cổ phiếu (Sp) bị tác động bởi các yếu tố EPS, tỷ số P/E, tốc độ tăng trưởng GDP và lãi suất (INT).

4. Kết quả và thảo luận

4.1. Kết quả phân tích mô hình Event Study

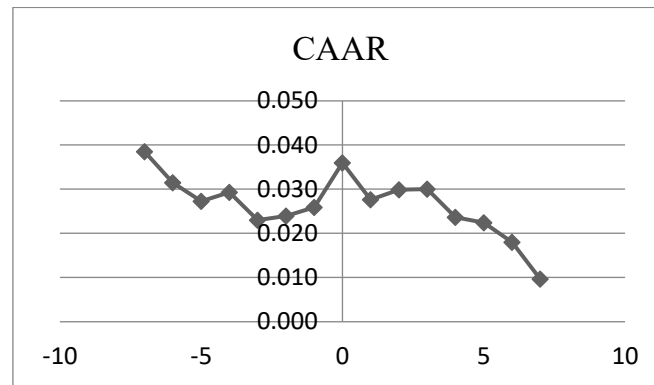
Giá trị AAR của 622 sự kiện đa phần là các giá trị dương cho thấy giá cổ phiếu có phản ứng tích cực đối với thông tin thông báo tài chính và phi tài chính trong khung thời gian (-2; -7) và (+1; +3). Trước ngày công bố thông tin, thị trường đều tiêu cực, AAR đang biến động ở mức khá ổn định nhưng lại giảm mạnh xuống -0,01% vào ngày $T = -1$. Tại ngày công bố thông tin ($T = 0$), thị trường tăng nhẹ với mức tăng bình quân đạt gần 0,008%. Một ngày trước và một ngày sau khi công bố thông tin, thị trường phản ánh khá tiêu cực với thông tin được đưa ra, sau đó từ ngày thứ 3 trở đi thì thị trường phản ánh tích cực.

Mặt khác, CAAR lại có xu hướng tăng từ ngày 0 và các ngày tiếp theo. Các giá trị lợi nhuận trung bình tích lũy giai đoạn từ ngày 0 đến ngày 7 đều mang giá trị dương, trong đó CAAR (-7; 0), CAAR (-6; 0) ghi nhận mức ý nghĩa cao 1%. Đặc biệt, kết quả quan sát CAAR (-5; +5), CAAR (-4; +5) đều ở mức ý nghĩa 5%, thống nhất với kết quả nghiên cứu của Dasilas và Leventis (2011), Hue (2015).



Hình 2: Xu hướng biến động của lợi nhuận bất thường trung bình (AAR)

Nguồn: Kết quả khảo sát.



Hình 3: Xu hướng biến động của lợi nhuận bất thường trung bình tích lũy (CAAR)

Nguồn: Kết quả khảo sát.

Bảng 1: Thống kê mô tả các biến

Tên biến	Mô tả	Giá trị trung bình	Giá trị nhỏ nhất	Giá trị lớn nhất	Độ lệch chuẩn
Y	Lợi nhuận bất thường của cổ phiếu	0,0576	-0,35	37,16	1,49
VHTT	Vốn hóa thị trường	3,64	2,02	5,57	0,92
PE	Hệ số giá trên lợi nhuận một cổ phiếu	20,13	-8,00	408	49,58
ROS	Tỷ suất lợi nhuận trên doanh thu	0,0009	-0,12	0,12	0,012
Beta	Chỉ số rủi ro hệ thống của cổ phiếu	0,6385	-0,98	1,71	0,511
GDP	Tổng sản phẩm quốc nội	1,7586	-2,64	18	1,51
LS	Lãi suất thị trường	4,588	-6,17	7,46	3,39
TG	Tỷ giá hối đoái	0,1765	-1,54	1,52	0,55

Nguồn: Kết quả phân tích SPSS.

4.2. Kết quả dự báo giá cổ phiếu từ thị trường tài chính và phi tài chính sử dụng mô hình Random Forest

Kết quả thống kê mô tả các biến

Bộ dữ liệu gồm 8 biến và 622 quan sát được thu thập trong khoảng thời gian 5 năm từ 01/2017 đến 12/2021. Kết quả thống kê mô tả sử dụng phần mềm SPSS của các biến cho thấy giá trị trung bình, giá trị nhỏ nhất và giá trị lớn nhất của các biến trong mô hình nghiên cứu. Cuối cùng, độ lệch chuẩn thể hiện mức độ dao động của các biến xung quanh giá trị trung bình. Độ lệch chuẩn càng lớn thì giá trị của các số liệu thu thập càng chênh lệch.

Có thể thấy độ lệch chuẩn của các biến khá nhỏ, tức là các dữ liệu thu thập được đồng đều, không có sự chênh lệch lớn trừ độ lệch chuẩn của biến P/E.

P/E cao nhất là 408 lần thuộc mã PNJ, giá trị nhỏ nhất là -8 lần thuộc mã NKG. Độ lệch chuẩn của P/E khá cao, chứng tỏ các giá trị P/E thu thập được có sự chênh lệch lớn. Cụ thể, P/E của PNJ quý 2 năm 2020 đạt mức 408 lần do ảnh

hưởng của COVID-19, PNJ ghi nhận 31,7 tỷ đồng lợi nhuận sau thuế trong quý II/2020, giảm 81% so với cùng kỳ năm trước, khiến EPS cũng giảm mạnh. Ngoài ra, P/E của các ngành nghề khác nhau có chênh lệch rất lớn do đặc điểm thị phần, cấu trúc vốn, mô hình hoạt động, rủi ro và lợi thế cạnh tranh của mỗi doanh nghiệp là hoàn toàn khác nhau. Thông thường, chỉ số P/E cao thể hiện sự kỳ vọng của nhà đầu tư về việc tăng trưởng thu nhập từ cổ phiếu đó sẽ cao hơn trong tương lai, ví dụ như một số doanh nghiệp trong danh mục VN30 - PLX hoặc PNJ. Ta có thể giải thích ngược lại đối với các doanh nghiệp có chỉ số P/E thấp.

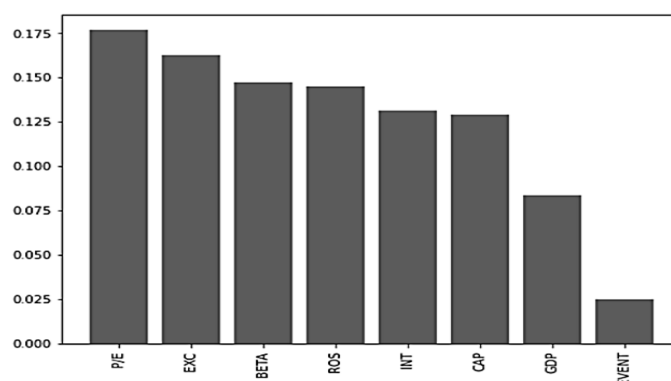
Kết quả của mô hình dự báo

Kết quả Bảng 2 cho thấy q biến nào hệ số càng cao thì khả năng dự báo càng tốt. Biến GDP và biến thông tin xếp cuối về mức độ ảnh hưởng đến biến phụ thuộc. Biến thông tin hay đối tượng nghiên cứu chính trong nghiên cứu có tác động khá khiêm tốn đối với lợi nhuận bất thường của cổ phiếu với kết quả chỉ bằng 0,024. Nói cách khác, không có sự khác biệt nhiều giữa thông tin tài chính và thông tin phi tài chính.

Bảng 2: Kết quả mô hình Random Forest

P/E	0,1765
TG	0,1627
beta	0,1473
ROS	0,1448
LS	0,1314
VHTT	0,1289
GDP	0,0832
Thongtin	0,0248

Nguồn: Kết quả chạy Random Forest.



Hình 4: Mức độ tác động của các biến độc lập đến biến phụ thuộc

Nguồn: Kết quả chạy Random Forest.

Dựa trên kết quả nghiên cứu mô hình Random Forest, mô hình đánh giá mức độ tác động của các biến độc lập đến Y có dạng như sau:

$$Y = 0,024\text{Event} + 0,176\text{PE} + 0,163\text{TG} + 0,147\text{BETA} + 0,145\text{ROS} + 0,131\text{LS} + 0,129\text{VHTT} + 0,083\text{GDP}$$

Kết quả cho thấy biến Thongtin có tác động không nhiều đến biến phụ thuộc Y hay lợi nhuận bất thường trung bình của cổ phiếu có thể được lý giải bởi các nguyên nhân sau:

Thứ nhất, điều kiện kinh tế luôn thay đổi trong khi con người không theo kịp những thay đổi đó. Điều này có thể dẫn đến tình trạng chậm phản ứng với thay đổi và gán nhận định của cá nhân với tình hình chung trong một thời kỳ dài hạn trước đó. Tâm lý bảo thủ của nhà đầu tư khiến họ trở nên kém linh hoạt với các thông tin mới và không có chiến lược đầu tư phản ứng với các thông tin này.

Thứ hai, mức độ hiệu quả của thông tin còn kém, dẫn đến nhà đầu tư phản ứng không đủ. Barberis và cộng sự (1998) cho rằng các nhà đầu tư cũng mắc sai lầm khi họ phản ứng không đủ với một số thông tin tài chính. Thông tin kém

hiệu quả còn thể hiện ở sự chậm trễ trong công bố thông tin, do thông tin bị làm giả hoặc bị thao túng để đạt được mục đích riêng của một số cá nhân hoặc tổ chức. Khi công ty chứng khoán phân tích một mã chứng khoán, thông tin sẽ được lan truyền trong nội bộ công ty một khoảng thời gian trước khi công bố công khai, dẫn đến việc giá cổ phiếu bắt đầu dao động từ trước khi thông tin được công bố.

Thứ ba, có thể do dữ liệu bảng mà nhóm nghiên cứu thu thập chưa bao gồm thông tin trên các trang thông tin phi chính thức như Facebook, Twitter..., trong khi tỷ lệ phần trăm người dùng tiếp nhận thông tin từ các mạng xã hội lên đến 36%. Các trang mạng xã hội này cung cấp một lượng thông tin khổng lồ. Tại Việt Nam, Facebook là mạng xã hội được sử dụng nhiều nhất với 96%, theo sau đó là Zalo (91%) và YouTube (86%) (Reuters, 2020). Như vậy, mạng xã hội góp phần không nhỏ để các nhà đầu tư mới (F0) tiếp cận thị trường TTCK nhanh hơn. Tuy nhiên, không thể không kể đến thông tin nhiễu được tung ra hay các phân tích báo cáo tràn lan trên mạng khiến nhà đầu tư hoang mang.

Kiểm định khả năng dự báo giá cổ phiếu của thông tin tài chính và thông tin phi tài chính sử dụng ngày không sự kiện

Để kiểm tra khả năng dự báo giá cổ phiếu của thông tin tài chính và phi tài chính, nhóm tác giả đã tiến hành thêm hai mô hình sử dụng ngày không sự kiện. Cụ thể, mô hình 2: Tác động của thông tin phi tài chính so với khi không có sự kiện; Mô hình 3: Tác động của thông tin tài chính so với khi không có sự kiện. Cả hai mô hình được thực hiện tương tự mô hình 1 và sau đó so sánh kết quả của hai mô hình.

Kết quả Bảng 3 bổ sung cho kết luận thị trường có phản ứng với những thông tin chứng khoán được công bố nhưng không có nhiều sự khác biệt giữa phản ứng đối với thông tin tài chính hoặc phi tài chính được công bố. Tuy nhiên, có thể thấy thông tin cũng là một yếu tố quan trọng trong việc dự báo biến động giá cổ phiếu khi đem so sánh với những ngày thị trường không công bố thông tin.

Bảng 3: Kết mô mô hình 2 và mô hình 3

	Mức độ tác động đến biến phụ thuộc	Khả năng dự báo của mô hình
Thông tin phi tài chính	6,2%	91,4%
Thông tin tài chính	6,0%	88,8%

Nguồn: Kết quả khảo sát.

5. Kết luận và khuyến nghị

5.1. Kết luận

Nghiên cứu đã sử dụng kết hợp phương pháp Event Study và thuật toán Random Forest để xây dựng mô hình dự báo biến động tiếp theo của giá cổ phiếu, đưa ra kết quả thực sự ấn tượng trong việc dự đoán hướng biến động lợi nhuận bất thường của cổ phiếu trong tương lai với độ chính xác là gần 90%. Nghiên cứu đóng góp 3 kết quả chính sau:

Thứ nhất, khẳng định có tồn tại lợi nhuận bất thường tích lũy xung quanh sự kiện thông báo thông tin tài chính và phi tài chính, hay nói cách khác có sự biến động trong giá cổ phiếu xung quanh ngày diễn ra sự kiện hoặc ngày thông tin được công bố. Kết quả này đồng nhất với nghiên cứu của Barberis và cộng sự (1998), Bogdan và

cộng sự (2010), Robu và cộng sự (2013). Kết quả biến thông tin có tác động không nhiều đến lợi nhuận bất thường trung bình của cổ phiếu là điều dễ hiểu, có thể được lý giải bởi ở Việt Nam, yếu tố tâm lý nhà đầu tư bị tác động rất mạnh bởi thông tin chính thức và cả thông tin gây nhiễu.

Thứ hai, kết quả chạy mô hình hồi quy thông qua thuật toán Random Forest cho thấy biến nào có hệ số càng cao thì khả năng dự báo càng tốt. Theo đó, 2 biến có tác động nhiều nhất đến lợi nhuận bất thường của cổ phiếu là biến P/E và biến tỷ giá với hệ số lần lượt là 0,17 và 0,16. Xếp ngay sau đó là nhóm chỉ số beta và ROS có mức độ tác động xấp xỉ nhau với hệ số khoảng 0,14. Mức độ tác động của lãi suất và biến vốn hóa thị trường xếp thứ ba với hệ số lần lượt là 0,13 và 0,12. Biến GDP và biến thông tin xếp cuối về mức độ ảnh hưởng đối với biến phụ thuộc. Một số nghiên cứu đồng nhất với kết quả các biến kinh tế vĩ mô có tác động đến giá cổ phiếu, điển hình như Nguyen và Nguyen (2013), Dinh và Nguyen (2016), Nguyen và Le (2020).

Thứ ba, nghiên cứu cung cấp mô hình dự báo giá cổ phiếu với 8 yếu tố độc lập gồm thông tin, vốn hóa thị trường, 3 chỉ số tài chính P/E, ROS và Beta và 3 yếu tố kinh tế vĩ mô. Tuy nhiên, mô hình không thể dự đoán xu hướng cổ phiếu trong dài hạn và yếu tố tâm lý thị trường không thể nắm bắt được, do đó nhà đầu tư cần áp dụng song song nhiều mô hình để đạt được kết quả tối ưu.

5.2. Đề xuất khuyến nghị cho các nhà đầu tư chứng khoán

Thứ nhất, để nâng cao hiệu quả các quyết định đầu tư của nhà đầu tư cá nhân trên TTCK, các nhà đầu tư không nên dựa quá nhiều vào thông tin được công bố trên thị trường, cần thận trọng trước các nguồn công bố thông tin.

Thứ hai, các nhà đầu tư cá nhân cần nâng cao trình độ đọc hiểu và phân tích thông tin trên báo cáo tài chính như ROS, P/E, Beta hoặc phân tích lợi nhuận bất thường AR của cổ phiếu – đây có thể là công cụ hữu ích giúp các nhà đầu tư đưa ra quyết định đúng đắn hơn.

Thứ ba, cần cải thiện các yếu tố tâm lý trong đầu tư nhằm nâng cao hiệu quả các quyết định đầu tư. Tâm lý là yếu tố chi phối hành vi nhà đầu tư, có ý nghĩa quan trọng trong việc xác định giá chứng khoán. Quyết định đầu tư không hợp lý thường bị chi phối bởi những cảm xúc như sợ hãi, tham lam, sợ bỏ lỡ, tâm lý đám đông..., từ đó dẫn đến giao dịch dựa trên nỗi sợ hãi.

Thứ tư, để tăng tính hiệu quả của TTCK, các doanh nghiệp niêm yết cần nâng cao chất lượng thông tin báo cáo tài chính và hình ảnh doanh nghiệp, cũng như xem xét việc công bố thông tin đảm bảo.

Thứ năm, đối với cơ quan quản lý nhà nước hoặc Ủy ban Chứng khoán Nhà nước, cần tăng cường giám sát việc công bố thông tin của các doanh nghiệp niêm yết nhằm tăng cường tính minh bạch, đảm bảo quyền lợi cho các đối tượng sử dụng, trong đó có các nhà đầu tư cá nhân.

Tài liệu tham khảo

- Barberis, N., Shleifer, A., Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49, 307-345. [https://doi.org/10.1016/S0304-405X\(98\)00027-0](https://doi.org/10.1016/S0304-405X(98)00027-0)
- Bogdan, V., Dumitrescu, D., & Văcaru, D. (2013). Effects of financial and non-financial information disclosure on prices' mechanisms for emergent markets: The case of Bucharest Stock Exchange. *Accounting and Management Information Systems*, 12(1), 76.
- Bonnier, K. -A., & Bruner, R. F. (1989). An analysis of stock price reaction to management change in distressed firms. *Journal of Accounting and Economic*, 11(1), 95-106.
- Dasilas, A., & Leventis, S. (2011). Stock market reaction to dividend announcements: Evidence from the Greek stock market. *International Review of Economics and Finance*, 20(2), 302-311. <https://doi.org/10.1016/j.iref.2010.06.003>
- Ngoc, D. B., & Cuong, N. C. (2016). Dividend announcement and ex-dividend effects on stock return. *International Journal of Economics and Finance*, 8(7), 207-215. <https://doi.org/10.5539/ijef.v8n7p207>
- Filip, A. C., Raffournier, B., & Ionascu, I. (2012). The impact of non-financial reporting on stock markets in emerging economies. *Procedia Economics and Finance*, 3, 781-785. [https://doi.org/10.1016/S2212-5671\(12\)00232-7](https://doi.org/10.1016/S2212-5671(12)00232-7)
- Gurgul, H., Mestel, R., & Schleicher, C. (2003). Stock market reactions to dividend announcements: empirical evidence from the Austrian stock market. *Financial Markets and Portfolio Management*, 17(3), 332-348. <https://doi.org/10.1007/s11408-003-0302-x>
- Hegazy, O., Soliman, O. S., & Salam, M. A. (2014). A machine learning model for stock market prediction. *arXiv Preprint arXiv:1402.7351*. <https://doi.org/10.48550/arXiv.1402.7351>
- Kieu, M. N., & Nhen, L. T. N. (2020). Factors affecting the market value of commercial bank stocks in Vietnam. *Banking Review*, 7.
- Loc, T. D. (2014). Factors affecting stock price change: Evidence from Ho Chi Minh City Stock Exchange. *Scientific Journal of Can Tho University*, 33, 72-78. <https://sj.ctu.edu.vn/ql/docgia/tacgia-81894/baibao-2961.html>
- MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, 35(1), 13-39. <https://www.jstor.org/stable/2729691>
- Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383-417. <https://doi.org/10.2307/2325486>
- Manh, P. T., & Huong, T. T. (2018). Impact of dividend and profit disclosure on stock price volatility of companies listed on Ho Chi Minh City Stock Exchange (HOSE). *Vietnam Science and Technology Magazine*, 60 (11).
- Mehedințu, C., Avram, N., Medrea, N., & Duțescu, C. (2000). Fluorine values in biological fluids and paraclinical changes in industrial fluorosis of ruminants. *Studies and Researches in Veterinary Medicine*, 8, 89-98.
- Moreno, K. K., & Zhang, Y. (2021). The impact of the Big Fish Effect on investor reactions to financial and nonfinancial disclosure. *Journal of Behavioral Finance*, 22(2), 113-125. <https://doi.org/10.1080/15427560.2020.1835059>
- Nam, B. T., & Anh, N. T. K. (2021). The Event Study in Quantitative Research. *TNU Journal of Science and Technology*, 226(08), 149-156. <https://doi.org/10.34238/tnu-jst.4187>
- Nguyen, M. K. (2013). The relationship between macroeconomic factors and stock market volatility: Empirical evidence from the Vietnam stock market. *Science & Technology Development*, 16.
- Nguyet, P. T. B., & Thao, D. P. (2013). Analyze the impact of macroeconomic factors on Vietnam's stock market. *Journal of Development and Integration*, 8, 34-41.
- Nisar, T. M., & Yeung, M. (2018). Twitter as a tool for forecasting stock market movements: A short-window event study. *The Journal of Finance and Data Science*, 4(2), 101-119. <http://doi.org/10.1016/j.jfds.2017.11.002>
- Polamuri, S. R., Srinivas, K., & Mohan, A. K. (2019). Stock market prices prediction using random forest and extra tree regression. *International Journal of Recent Technology and Engineering*, 8(1), 1224-1228. <http://doi.org/10.35940/ijrte.C4314.098319>
- Puspitaningtyas, Z. (2019). Empirical evidence of market reactions based on signaling theory in Indonesia stock exchange. *Investment Management and Financial Innovations*, 16(2), 66-77. [https://doi.org/10.21511/imfi.16\(2\).2019.07](https://doi.org/10.21511/imfi.16(2).2019.07)
- Reuters Institute Digital News Report (2020). *Digital news report 2020*. Reuters Institute for the Study of Journalism. <https://www.digitalnewsreport.org/survey/2020/>

- Robu, M. A., Sandu, M., & Siminică, M. (2013). The statistical assessment of an emerging capital market using the panel data analysis of the financial information. *Economic Computation and Economic Cybernetics Studies and Research*, 47(2), 21-36.
- Son, N. (2018). Information communication with security of the stock market in the era of digital information technology. *Vietnam Economic and Financial Magazine*, (4), 1–10.
- The Association of Chartered Certified Accountants (2013). *What do investors expect from non-financial reporting?*
<https://www.accaglobal.com/gb/en/technical-activities/technical-resources-search/2013/august/investors-and-non-financial-reporting.html>
- Thu, H. A. (2016). The relationship between abnormal profit and financial statement disclosure time of companies listed on Vietnam stock market. *Lac Hong Scientific Journal*, 5, 1-10.
- Todea, A., & Buglea, A. M. (2017). Individualism and stock price reaction to market-wide information. *Economics Letters*, 160, 4-6.
<https://doi.org/10.1016/j.econlet.2017.08.023>
- Usman, B., & Tandelilin, E. (2014). internet search traffic and its influence on liquidity and returns of indonesia stocks: An empirical study. *Journal of Indonesian Economy and Business: JIEB*, 29(3), 203-215. <http://doi.org/10.22146/jieb.6477>
- Vinh, V. X. (2016). Capital structure and unusual profits accumulated on the vietnamese stock market. *Journal of Economic Studies* 7 (458).
- Vinh, V. X., & Kiem, D. B. (2016). Market reaction to the announcement of stock changes in the VNM ETF portfolio on the Vietnamese stock market. *VNU Journal of Economics and Business*, 32(3), 1-12.
- Yu, T. H. K., & Huarng, K. H. (2020). A new event study method to forecast stock returns: The case of Facebook. *Journal of Business Research*, 115, 317-321.
<https://doi.org/10.1016/j.jbusres.2020.05.048>