

DỰ BÁO BIẾN ĐỘNG TRÊN THỊ TRƯỜNG CHỨNG KHOÁN VIỆT NAM BẰNG PHƯƠNG PHÁP HỌC MÁY KẾT HỢP ADAPTIVE LASSO CHO MÔ HÌNH GARCH-X

Ths. Nguyễn Thị Hiền*

Thị trường chứng khoán Việt Nam đang ngày càng trở nên quan trọng trong bối cảnh kinh tế phát triển. Bài viết nghiên cứu độ biến động trên thị trường chứng khoán Việt Nam dựa trên hai chỉ số chính trên thị trường gồm chỉ số VN-Index và chỉ số HNX-Index. Trong nghiên cứu này, phương pháp học máy kết hợp với phương pháp Adaptive Lasso cho mô hình GARCH-X được sử dụng để dự báo. Trong đó, phương pháp học máy sử dụng thuật toán tối ưu Gradient descent. Nghiên cứu xem xét ảnh hưởng của các yếu tố: giá vàng, giá dầu thô, tỷ giá đồng USD/VND và giá tiền điện tử Bitcoin đến độ biến động trên thị trường chứng khoán Việt Nam. Để đánh giá hiệu quả của mô hình GARCH-X, nghiên cứu so sánh kết quả dự báo thu được của các mô hình. Kết quả, chỉ ra độ biến động trên thị trường chứng khoán Việt Nam chịu ảnh hưởng của sự tăng trưởng của giá dầu thô, giá vàng nhưng không chịu ảnh hưởng của sự tăng trưởng tỷ giá đồng USD/VND và tăng trưởng của giá tiền điện tử Bitcoin. Kết quả nghiên cứu cũng cho thấy mô hình GARCH(1,1)-X Adaptive Lasso có hiệu quả dự báo tốt nhất trong các mô hình đề xuất.

• Từ khóa: biến ngoại sinh, độ biến động, mô hình garch-x, phương pháp lasso, phương pháp gradient descent, học máy.

The Vietnamese stock market is becoming increasingly important in the context of economic development. This article studies the volatility in the Vietnamese stock market based on two main indices: the VN-Index and the HNX-Index. In this research, machine learning methods combined with the Adaptive Lasso method for the GARCH-X model are used for forecasting. The machine learning approach utilizes the Gradient Descent optimization algorithm. The study examines the impact of factors such as gold prices, crude oil prices, the USD/VND exchange rate, and Bitcoin prices on volatility in the Vietnamese stock market. To evaluate the effectiveness of the GARCH-X model, the study compares the forecasting results obtained from different models. The findings suggest that the volatility in the Vietnamese stock market is affected by the increases in crude oil and gold prices, while the changes in the USD/VND exchange rate and Bitcoin prices do not have an impact. The research findings also show that the GARCH(1,1)-X Adaptive Lasso model has the best forecasting performance among the proposed models.

• Key words: exogenous variables, volatility, garch-x model, lasso method, gradient descent method, machine learning.

Ngày gửi bài: 02/01/2025

Ngày gửi phản biện: 03/02/2025

Ngày nhận kết quả và sửa phản biện: 10/02/2025

Ngày chấp nhận đăng: 28/02/2025

DOI: <https://doi.org/10.71374/jfar.v25.i285.05>

1. Đặt vấn đề

Thị trường chứng khoán (TTCK) đang thay đổi liên tục với nhiều bất ổn. Sự lan truyền nhanh chóng của thông tin và dòng vốn nhanh sẽ dẫn đến sự biến động giá cổ phiếu và giá cả dao động sẽ ảnh hưởng trở lại đến thị trường. Hơn nữa, sự tác động của nền kinh tế thể hiện ở sự biến động của các chỉ số kinh tế vĩ mô cũng có ảnh hưởng ít nhiều đến sự biến động trên thị trường chứng khoán của mỗi nước. Thị trường chứng khoán Việt Nam đã biến động mạnh mẽ từ những ngày đầu và thường xuất hiện những đợt tăng giảm đột ngột. Độ biến động trên thị trường chứng khoán Việt Nam luôn thu hút sự quan tâm của các nhà đầu tư tài chính. Trong các nghiên cứu, độ biến động của giá cổ phiếu thường được đo lường bằng phương sai của dãy chỉ số giá đóng cửa cổ phiếu hàng ngày. Nhiều nghiên cứu đã thảo luận về hiệu quả của các mô hình GARCH trong việc giải thích độ biến động của TTCK, như nghiên cứu của Taufiq Choudhry (1997) về sự biến động trên thị trường chứng khoán Canada, Đan Mạch, Thụy Điển, Thụy Sĩ, Vương quốc Anh và Hoa Kỳ, và tác động của những thay đổi ngắn hạn giữa các chỉ số chứng khoán của các thị trường này đối với sự biến động trong giai đoạn từ tháng 1 năm 1926 đến tháng 12 năm 1944. Nghiên cứu thực nghiệm được tiến hành bằng các mô hình GARCH(1,1) và GARCH(1,1)-X. Kết quả hai mô hình chỉ ra tác động đáng kể của những thay đổi ngắn hạn giữa các

* Trường Đại học Thương mại; email: hiennguyen@tmu.edu.vn

chỉ số chứng khoán của các thị trường trên đối với biến động. Mô hình GARCH (1,1)-X cho kết quả phân tích tốt hơn mô hình GARCH(1,1). Karmakar (2005) về sự biến động của TTCK Ấn Độ. Nghiên cứu này chỉ ra mô hình GARCH(1,1) cung cấp dự báo biến động thị trường khá tốt và sử dụng mô hình GARCH(1,1) đưa ra các dự báo cho 50 mã cổ phiếu. Tương tự, nghiên cứu của Sohail Chand và cộng sự (2012) cũng đề cập đến vấn đề này, nhóm tác giả chỉ ra mô hình GARCH(1,1) là mô hình dự báo tốt nhất. Trần Mạnh Cường và Đỗ Khắc Hường (2013) đo lường sự dao động của chỉ số chứng khoán VN-Index thông qua mô hình GARCH với bộ số liệu giá đóng cửa hàng ngày giai đoạn 2000 - 2011. Nhóm tác giả sử dụng mô hình GARCH(p,q) với các trường hợp sai số tuân theo các quy luật phân phối chuẩn, phân phối Student, phân phối sai số tổng quát, phân phối Student chệch. Kết quả chỉ ra mô hình GARCH(1,1) với sai số tuân theo quy luật phân phối chuẩn có hiệu quả dự báo tốt nhất. Năm 2014, Erginbay Ugurlu và cộng sự nghiên cứu độ biến động của TTCK châu Âu bằng các mô hình GARCH với bộ dữ liệu gồm giá đóng cửa hàng ngày trên các TTCK mới nổi như Bulgaria, Cộng hòa Séc, Hungary, Ba Lan và Thổ Nhĩ Kỳ trong giai đoạn 2001 - 2012, dữ liệu thu thập được từ Reuters. Kết quả bài báo chỉ ra các cú sốc có ảnh hưởng dai dẳng đến độ biến động của tỷ suất sinh lợi, và ảnh hưởng của các tin tức là bất cân xứng. Năm 2017, tác giả Phạm Chí Khoa cũng đưa ra kết quả dự báo biến động giá chứng khoán qua mô hình GARCH, dựa trên bộ số liệu thu thập trong giai đoạn 2006 - 2016. Kết quả chỉ ra tỷ suất sinh lợi trong quá khứ có vai trò quyết định tỷ suất sinh lợi hiện tại. Hasan và Zaman (2017) đã kiểm tra sự biến động của lợi nhuận thị trường chứng khoán Bangladesh đáp ứng với sự biến động của các biến kinh tế vĩ mô, sử dụng dữ liệu chỉ số hàng tháng của Sở Giao dịch Chứng khoán Dhaka (DSE) và bốn biến kinh tế vĩ mô (Lãi suất tiền gọi, giá dầu thô, tỷ giá hối đoái và SENSEX từ Sở Giao dịch Chứng khoán Bombay) trong giai đoạn từ tháng 1 năm 2001 đến tháng 12 năm 2015. Kết quả của mô hình GARCH-S cho thấy sự biến động lợi nhuận của DSE bị ảnh hưởng đáng kể bởi các biến kinh tế vĩ mô, chẳng hạn như tỷ giá hối đoái và SENSEX. Robiyanto và cộng sự (2019) đã khảo sát ảnh hưởng của các biến kinh tế vĩ mô: Chỉ số Dow Jones Industrial Average, tỷ giá USD/IDR và giá dầu thế giới đến chỉ số giá cổ phiếu Tổng hợp (IHSG) trong giai đoạn 2005-2016, sử dụng mô hình GARCH (1,1), cho thấy rằng Chỉ số Dow

Jones Industrial Average và giá dầu thô thế giới có ảnh hưởng tích cực đến IHSG, trong khi tỷ giá USD/IDR có ảnh hưởng tiêu cực đến IHSG.

Do vậy, trong nghiên cứu này mô hình GARCH-X kết hợp với phương pháp học máy sử dụng thuật toán tối ưu Gradient descent và phương pháp Lasso, Adaptive Lasso sẽ được ứng dụng để dự báo độ biến động của VN-Index và HNX-Index dựa trên bốn chỉ số kinh tế vĩ mô gồm: giá dầu thô, tỷ giá USD/VND, giá Bitcoin và giá vàng thế giới (USD).

2. Phương pháp nghiên cứu

2.1. Dữ liệu nghiên cứu

Nghiên cứu sử dụng hai chỉ số giá chứng khoán chính trên thị trường chứng khoán Việt Nam: VN-Index và HNX-Index. VN-Index thể hiện biến động giá cổ phiếu niêm yết tại TTCK Thành phố Hồ Chí Minh (sàn HOSE). HNX-Index là chỉ số phản ánh biến động giá của các mã cổ phiếu được niêm yết trên Sở Giao dịch Chứng khoán Hà Nội. Nghiên cứu còn thu thập dữ liệu của 4 chỉ số kinh tế vĩ mô gồm: giá dầu thô, tỷ giá USD/VND, giá Bitcoin và giá vàng thế giới (USD) đưa vào mô hình GARCH-X là các biến ngoại sinh. Dữ liệu của các chỉ số được thu thập từ tháng 1/2016 cho đến tháng 1/2025 gồm 2262 quan sát. Dữ liệu được chia thành hai tập, tập huấn luyện gồm 1500 quan sát ban đầu của chuỗi và tập dữ liệu kiểm thử gồm các quan sát còn lại. Dữ liệu thu thập từ trang web: <https://vn.investing.com>. Các dữ liệu được phân tích trên phần mềm Excel, Eviews 12 và Matlab 2024.

Đối với dữ liệu chuỗi thời gian, các chuỗi dữ liệu dừng để dự báo phải là chuỗi dừng. Do vậy, để dự báo biến động của thị trường chứng khoán, nghiên cứu kiểm định tính dừng của các chuỗi dữ liệu và biến đổi các chuỗi không dừng thành chuỗi dừng. Chuỗi chỉ số giá chứng khoán VN-Index và HNX-Index ban đầu biến đổi thành các chuỗi tỷ suất sinh lợi kí hiệu là r_t và được tính theo công thức sau:

$$r_t = 100 \log\left(\frac{P_t}{P_{t-1}}\right)$$

Trong đó: P_t là giá đóng cửa của chuỗi giá tại thời điểm t .

Các chuỗi chỉ số kinh tế vĩ mô biến đổi thành các chuỗi tăng trưởng theo công thức:

$$100.(P_t - P_{t-1})/(P_{t-1})$$

Sử dụng kiểm định nghiệm đơn vị, kết quả cho thấy các chuỗi sau biến đổi đều là các chuỗi dừng với mức nghĩa 1%. Giá trị trung bình của tỷ suất sinh lợi r_t kí hiệu là: $\mu_t = E(r_t / F_{t-1})$, được giả thiết

bằng 0, với F_{t-1} là tập hợp thông tin có được tính đến thời điểm $t-1$. Do vậy chuỗi được biến đổi để có kỳ vọng bằng 0 bằng cách trừ đi giá trị trung bình và chia cho độ lệch tiêu chuẩn. Phương sai có điều kiện của tỷ suất sinh lợi r_t kí hiệu là $\sigma_t^2 = \text{Var}(r_t / F_{t-1})$.

2.2. Mô hình nghiên cứu

Mô hình GARCH(1,1)

Mô hình GARCH được Bollerslev (1986) xây dựng nhằm đo lường độ biến động của một tài sản tài chính (chứng khoán, vàng, dầu thô,...). Mô hình GARCH(1,1) có dạng như sau:

$$\begin{aligned} r_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned} \quad (1)$$

Trong đó $\alpha_0 > 0$; $\alpha_1 \geq 0$, $\beta_1 \geq 0$; $\alpha_1 + \beta_1 < 1$; $E(r_t / F_{t-1})$ giả thiết bằng 0; ε_t là các sai số ngẫu nhiên độc lập cùng có phân phối chuẩn với $E(\varepsilon_t) = 0$; $\text{Var}(\varepsilon_t) = 1$. Do vậy, r_t có phân phối chuẩn $E(r_t) = 0$; $\text{Var}(r_t) = \sigma_t^2$.

Mô hình GARCH-X(1,1)

Khi kết hợp các biến ngoại sinh vào mô hình GARCH(1,1) ta được mô hình GARCH-X(1,1) có dạng:

$$\begin{aligned} r_t &= \sigma_t \varepsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + f(X_{t-1}) \end{aligned} \quad (2)$$

X là một vecto gồm các biến giải thích đưa vào mô hình (2). Nếu X_{t-1} luôn dương thì $f(X_{t-1}) = \pi X_{t-1}$; tổng quát hơn khi X_{t-1} có giá trị âm $f(X_{t-1}) = \pi X_{t-1}^2$ (Han và Kristensen, 2014). Việc ước lượng mô hình (2) trở nên phức tạp hơn nếu f là một hàm bất kỳ. Do X_{t-1} có các giá trị âm nên trong nghiên cứu này ta xét f là một hàm tuyến tính, f có dạng $f(X_{t-1}) = \pi X_{t-1}^2$, các biến độc lập đưa vào mô hình là các chuỗi dừng.

Để ước lượng các tham số trong mô hình GARCH và GARCH-X, nghiên cứu sử dụng phương pháp học máy với thuật toán học là thuật toán tối ưu Gradient descent, và để hiệu chỉnh trọng số phù hợp cho các biến ngoại sinh nghiên cứu sử dụng kết hợp phương pháp Lasso và Adaptive Lasso. Khi tiến hành hồi quy mô hình Lasso trên một bộ dữ liệu mà có các biến đầu vào đa cộng tuyến thì mô hình hồi quy Lasso sẽ có xu hướng lựa chọn ra một biến trong nhóm các biến đa cộng tuyến và bỏ qua những biến còn lại.

Thuật toán Gradient descent

Giả sử ta cần tìm cực trị toàn cục cho hàm $f(\theta)$ trong đó θ (θ là một vector dùng để ký hiệu tập hợp các tham số của một mô hình cần tối ưu (trong mô hình hồi quy thì các tham số chính là các hệ số của mô hình)). Đạo hàm của hàm số đó tại một

điểm θ bất kỳ được ký hiệu là $\nabla f(\theta)$. Tương tự như hàm 1 biến, thuật toán Gradient descent cho hàm nhiều biến cũng bắt đầu bằng một điểm dự đoán θ_0 , sau đó ở vòng lặp thứ t , quy tắc cập nhật là:

$$\hat{\theta} = \theta_t - \tau \nabla_{\theta} f(\theta_t) \quad (3)$$

Quy tắc là luôn luôn đi ngược hướng với đạo hàm.

Với các mô hình hồi quy sử dụng trong nghiên cứu, ta sẽ tối ưu hàm tổn thất bằng thuật toán Gradient descent. Với giả thiết sai số ngẫu nhiên có phân phối chuẩn, hàm tổn thất của mô hình GARCH, GARCH-X có dạng như sau:

$$\log llh(\theta_t) = -\frac{1}{2} \sum_{t=2}^T [\log(2\pi) + \log(\sigma_t^2) + \frac{y_t^2}{\sigma_t^2}] \quad (4)$$

Hàm tổn thất của mô hình GARCH-X Lasso có dạng như sau:

$$\log llh(\theta_t) = -\frac{1}{2} \sum_{t=2}^T [\log(2\pi) + \log(\sigma_t^2) + \frac{y_t^2}{\sigma_t^2}] - \lambda \sum_{j=1}^J |\theta_j| \quad (5)$$

Hàm tổn thất của mô hình GARCH-X Adaptive Lasso có dạng như sau:

$$\log llh(\theta_t) = -\frac{1}{2} \sum_{t=2}^T [\log(2\pi) + \log(\sigma_t^2) + \frac{y_t^2}{\sigma_t^2}] - \lambda \sum_{j=1}^J \hat{\omega}_j |\theta_j| \quad (6)$$

Trong đó: j là số biến ngoại sinh đưa vào mô hình GARCH-X, $\lambda > 0$ là tham số điều chỉnh, $\hat{\omega}_j$ là trọng số thích ứng (Adaptive weight) cho hệ số θ_j . Để có được trọng số $\hat{\omega}_j$ trước tiên ta hồi quy mô hình GARCH-X thu được các hệ số ước lượng $\hat{\theta}_j$ bằng phương pháp cực đại hàm hợp lý (4). Sau đó tính trọng số thích ứng:

$$\hat{\omega}_j = 1/|\hat{\theta}_j|$$

Các hệ số của mô hình cần ước lượng kí hiệu là θ^* và được tính theo công thức sau:

$$\hat{\theta} = \theta_t - \tau \nabla_{\theta} (-\log llh(\theta_t)) \quad (8)$$

Trong đó: τ là bước thay đổi của mỗi lần học (huấn luyện), τ thường có giá trị nhỏ.

3. Kết quả nghiên cứu

Thống kê mô tả tỷ suất sinh lời của hai chỉ số VNIndex và HNXIndex được tổng hợp trong bảng sau:

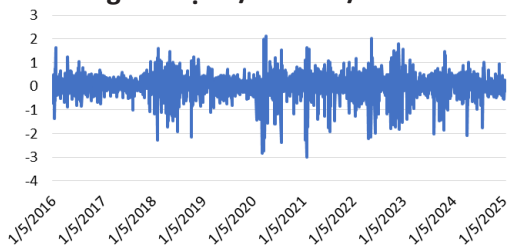
Bảng 1. Thống kê mô tả chuỗi tỷ suất sinh lời của VN-Index và HNX-Index theo ngày

	VN-Index	HNX-Index
Trung bình	0.01492	0.01970
Lớn nhất	2.11067	2.34156
Nhỏ nhất	-2.99999	-3.63764
Độ lệch chuẩn	0.49785	0.55717
Hệ số bất đối xứng	-0.98706	-0.98513
Hệ số nhọn	7.73344	7.91409
Cỡ mẫu	2262	2262

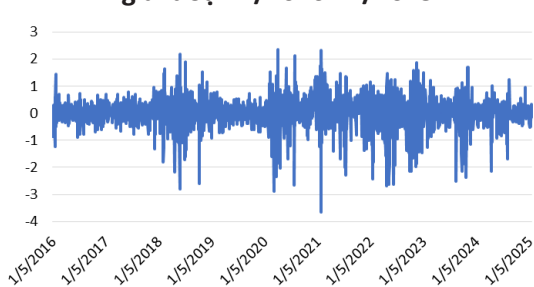
Nguồn: kết quả phân tích số liệu của tác giả

Giá trị trung bình của chuỗi tỷ suất sinh lời của VNIndex và HNXIndex tương ứng là 0.01492% và 0.0197% dương, điều này cho thấy trung bình chỉ số giá chứng khoán trong giai đoạn nghiên cứu (dài hạn) trên thị trường chứng khoán Việt Nam tăng. Hệ số bất cân xứng khác 0 và độ nhọn lớn hơn 3, điều này ngụ ý các chuỗi tỷ suất sinh lời của hai chỉ số không tuân theo quy luật phân phối chuẩn.

Hình 1. Đồ thị chuỗi tỷ suất sinh lời VNIndex giai đoạn 1/2016 - 1/2025



Hình 2. Đồ thị chuỗi tỷ suất sinh lời HNXIndex giai đoạn 1/2016 - 1/2025



Nguồn: Kết quả phân tích số liệu của tác giả

Đồ thị chỉ ra chuỗi tỷ suất sinh lời của hai chỉ số chứng khoán chính trên TTCK Việt Nam có các biên độ khác nhau tùy từng thời điểm, xung quanh lợi suất trung bình. Không chỉ thế những biến động ấy dường như kéo dài qua một giai đoạn nhất định, khoảng thời gian biến động cao thấp có xu hướng nối tiếp, tức là có biến động theo cụm và những biến động này có vẻ tự tương quan với nhau. Do đó, việc chọn mô hình GARCH, GARCH-X để đánh giá độ biến động của TTCK Việt Nam qua thời gian là phù hợp. Tính dừng của chuỗi tỷ suất sinh lợi được thực hiện bằng kiểm định nghiệm đơn vị và thu được kết quả:

Bảng 2. Kết quả kiểm định tính dừng của chuỗi VN-Index và HNX-Index bằng kiểm định nghiệm đơn vị ADF

Thống kê Augmented Dickey-Fuller (ADF)	t-Statistic	P_giá trị
VN-Index	-45.54197	0.0001
HNX-Index	44.73732	0.0001

Nguồn: Kết quả phân tích số liệu của tác giả

Vậy chuỗi tỷ suất sinh lợi của VN-Index và HNX-Index là các chuỗi dừng với mức ý nghĩa 1%. Nghiên

cứ sử dụng phương pháp học máy với thuật toán Gradient descent ước lượng mô hình GARCH(1,1) với số lần huấn luyện (Max iteration) là 1000, mỗi bước dịch chuyển (Step size) là 10^{-6} , mô hình GARCH(1,1)_X, mô hình GARCH(1,1)_X Lasso và mô hình GARCH(1,1)_X Adaptive Lasso. Mô hình GARCH(1,1)_X bổ sung bốn biến độc lập gồm tỷ suất sinh lợi của: vàng (XAU), dầu thô, (WTI) tỷ giá USD/VND(EXR) và tiền điện tử Bitcoin (BIT).

Trước tiên thực hiện ước lượng mô hình GARCH(1,1) để thu được ước lượng của σ_t^2 là σ_t^2 cho VN-Index (σ_t^2 - VNI) và HNX-Index (σ_t^2 - HNXI). Sau đó ta xét mối quan hệ tương quan giữa quả độ biến động của TTCK với các biến ngoại sinh thu được:

Bảng 3: Hệ số tương quan giữa các biến trong mô hình GARCH_X

	σ_t^2 - VNI	σ_t^2 - HNXI	XAU ²	WTI ²	EXR ²	BIT ²
σ_t^2 - VNI	1					
σ_t^2 - HNXI	-0.01747	1				
XAU ²	0.037126	-0.08702	1			
WTI ²	0.01695	-0.02392	0.03184	1		
EXR ²	-0.007458	-0.00347	0.031818	0.003644	1	
BIT ²	-0.011257	-0.00591	0.164144	0.01395	0.004676	1

Kết quả bảng 3 cho thấy, các biến ngoại sinh đưa vào mô hình GARCH_X không có dấu hiệu đa cộng tuyến. Do dấu của các biến ngoại sinh là tỷ suất sinh lợi của 4 chỉ số có mang dấu âm nên các biến đưa vào hàm f có dạng hàm bậc 2. Bảng hệ số tương quan cho thấy sự biến động trên các thị trường vàng, dầu thô, tiền điện tử Bitcoin và tỷ giá đều có ảnh hưởng ít nhiều đến sự biến động trên TTCK Việt Nam. Tuy nhiên, tăng trưởng của tỷ giá và tăng trưởng của giá tiền điện tử Bitcoin có mức tương quan rất thấp tới độ biến động trên TTCK.

Bảng 4: Kết quả ước lượng các mô hình

Hệ số	GARCH(1,1)	GARCH(1,1)_X	GARCH(1,1)_X Lasso	GARCH(1,1)_X Adaptive Lasso
VNIndex				
C	0.0202	0.0118	0.0235**	0.0038***
r^2_{t-1}	0.1141***	0.1233***	0.1286***	0.0966***
σ^2_{t-1}	0.8718***	0.8249***	0.7980***	0.8658***
XAU ² _{t-1}		0.1340***	0.1045**	0.0717*
WTI ² _{t-1}		0.2504	0.2791*	0.2477*
EXR ² _{t-1}		0.0001	0.0002	0
BIT ² _{t-1}		0.0030	0.0008	0
HNXIndex				
C	0.0223	0.0124	0.0126***	0.0070
r^2_{t-1}	0.1497***	0.1681***	0.1668***	0.1334***
σ^2_{t-1}	0.8396***	0.8011***	0.8044***	0.8469***
XAU ² _{t-1}		0.1169	0.1027*	0.0631*
WTI ² _{t-1}		0.2599*	0.2447	0.2408*
EXR ² _{t-1}		0.0013	0.0012*	0
BIT ² _{t-1}		-0.0008	-0.0009	0

Hệ số	GARCH(1,1)	GARCH(1,1)_X	GARCH(1,1)_X Lasso	GARCH(1,1)_X Adaptive Lasso
Số lần học tối đa	1000	1000	1000	1000
Bước chuyển (t)				

***: mức ý nghĩa 1%; **: mức ý nghĩa 5%; *: mức ý nghĩa 10%

Nguồn: Kết quả phân tích số liệu của tác giả

Bảng 4 cho thấy, sự biến động của TTCK Việt Nam chịu ảnh hưởng bởi sự biến động của giá dầu và giá vàng ở kì trước. Kết quả cũng chỉ ra, sự biến động trên thị trường tiền điện tử Bitcoin và tỷ giá USD/VND không có ảnh hưởng đến sự biến động của TTCK Việt Nam thể hiện các hệ số gần với các biến ngoại sinh này bằng 0 hoặc rất gần 0 trong các mô hình GARCH-X. Các phần dư thu được từ hai mô hình GARCH và GARCH-X được dùng để kiểm định lại sự phù hợp của mô hình.

Bảng 5: Kiểm định nghiệm đơn vị cho phần dư của các mô hình

	Thống kê T (VN-Index)	P_giá trị	Thống kê T (HNX-Index)	P_giá trị
Thống kê ADF				
GARCH(1,1)	-3.9491	0.0001	-27.39323	0.0000
GARCH(1,1)-X	-5.4110	0.0000	-28.2353	0.0000
GARCH(1,1)-X Lasso	-5.9132	0.0000	-28.1842	0.0000
GARCH(1,1)-X Adaptive Lasso	-6.2686	0.0000	-27.1802	0.0000

Nguồn: Kết quả phân tích số liệu của tác giả

Các chuỗi phần dư đều là các chuỗi dừng. Để đánh giá hiệu quả dự báo của các mô hình, nghiên cứu thực hiện dự báo trên tập dữ liệu kiểm thử và kết quả thu được như sau:

Bảng 6: Kết quả đánh giá sai số dự báo của các mô hình dựa trên tập dữ liệu kiểm thử

	GARCH(1,1)	GARCH(1,1)_X	GARCH(1,1)_X Lasso	GARCH(1,1)_X Adaptive Lasso
VN-Index				
MAE	1.3576	1.3901	1.3560	1.3419
RMSE	2.4389	2.4581	2.4352	2.4338
Logllh	1030.8	1053.9	1050.7	1054.3
HNX-Index				
MAE	1.7207	1.8174	1.8058	1.7206
RMSE	3.4136	3.4149	3.4128	3.4036
Logllh	1104.6	1114.1	1113.1	1114.7

Nguồn: Kết quả phân tích số liệu của tác giả

Bảng 6 cho thấy, trong 4 mô hình dự báo độ biến động trên TTCK Việt Nam, mô hình GARCH(1,1)-X Adaptive Lasso có hiệu quả dự báo tốt nhất, với sai số dự báo trên tập huấn luyện gồm 762 quan sát là nhỏ nhất theo cả hai tiêu chí sai số trung bình tuyệt đối MAE (Mean absolute error) và sai số RMSE (Root mean square error). Bộ tham số ước lượng của mô hình GARCH(1,1) -X Adaptive Lasso cũng cho ra giá trị của hàm hợp lý lớn nhất. Mô hình GARCH(1,1)-X Lasso cho kết quả dự báo tốt thứ hai theo chỉ tiêu RMSE. Mô hình GARCH(1,1)-X với cả hai chỉ số chứng khoán đều có kết quả dự

báo không tốt bằng mô hình GARCH(1,1), điều này phản ánh khi đưa thêm các biến ngoại sinh vào mô hình GARCH nên kết hợp phương pháp Lasso để gán trọng số phù hợp cho các biến giúp mô hình GARCH-X đạt kết quả dự báo tốt hơn.

Kết luận

Mô hình GARCH (Generalized Autoregressive Conditional Heteroscedastics) đã trở nên phổ biến kể từ khi ra đời do khả năng dự báo biến động. Thông thường, mô hình GARCH nắm bắt biến động dựa trên biến động trong quá khứ và các phần dư bình phương trong quá khứ, nhưng không xem xét tác động của biến ngoại sinh lên quá trình biến động do bản chất đơn biến của nó. Trong lĩnh vực mô hình kinh tế lượng, nơi các biến ngoại sinh đóng vai trò quan trọng, mô hình GARCH với sự bổ sung của các biến ngoại sinh được quan tâm nghiên cứu. Do đó, bài viết này đã giới thiệu và triển khai một mô hình GARCH-X cải tiến có thể tính đến tác động của các yếu tố ảnh hưởng (X) lên phương sai của mô hình GARCH kết hợp sử dụng phương pháp Lasso, Adaptive Lasso và học máy để ước lượng các tham số của mô hình. Nghiên cứu khảo sát ảnh hưởng của các biến kinh tế vĩ mô gồm giá vàng, giá dầu thô, tỷ giá hối đoái và giá tiền điện tử Bitcoin đến sự biến động của TTCK Việt Nam thông qua hai chỉ số giá thị trường chính là VN-Index và HNX-Index dựa trên dữ liệu chuỗi thời gian trong giai đoạn từ tháng 1 năm 2016 đến tháng 1 năm 2025. Kết quả cho thấy tăng trưởng của giá vàng, giá dầu thô có ảnh hưởng đến biến động trên TTCK Việt Nam, trong khi tăng trưởng của giá tiền điện tử Bitcoin và tỷ giá USD/VND thì gần như không ảnh hưởng. Phương pháp ước lượng học máy cho mô hình GARCH-X(1,1) Adaptive Lasso đã cải thiện hiệu quả dự báo, sai số dự báo của mô hình này nhỏ nhất so với ba mô hình còn lại sử dụng trong nghiên cứu.

Tài liệu tham khảo:

- Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity, *Journal of econometrics*, 31, 307-327.
- Engle, R. F. (1982), Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Journal of the Econometric Society*, 987-1007.
- Erginbay Ugurlu và cộng sự (2014), Modeling Volatility in the Stock Markets using GARCH Models: European Emerging Economies and Turkey, *International Journal in Economics and Business Administration Volume II, Issue 3*.
- Karmakar, M. (2005), Modeling conditional volatility of the Indian stock markets, *Vikalpa*, 30, 21.
- Khalid và Khan (2017), 'Effects of macroeconomic variables on the stock market volatility: the Pakistan experience', *Global Journal of Management and Business Research*, 5(2), 42-59.
- Phạm Chí Khoa (2017), Dự báo biến động giá chứng khoán qua mô hình Arch - Garch, *Tạp chí Tài chính*, Kỳ 2, số 6, tr38-39.
- Sohail Chand, Shahid Kamal & Imran Ali (2012), Modelling and volatility analysis of share prices using ARCH and GARCH models, *World Applied Sciences Journal*, 19, 77-82.
- Taufiq Choudhry (1997), Stock Return Volatility and World War II: Evidence from Garch and Garch-X Models, *International journal of Finance and economics*, volume 2, issue 1, 17 - 28.
- Trần Sỹ Mạnh và Đỗ Khắc Hường (2013), Đo lường sự dao động chỉ số chứng khoán VN Index thông qua mô hình Garch, *Tạp chí Khoa học và đào tạo ngân hàng*, số 130, tr42.
- Vũ Duy Thăng (2011), Các mô hình chuỗi thời gian tài chính, *Luận văn thạc sĩ khoa học, Đại học Khoa học tự nhiên - Đại học Quốc gia Hà Nội*.