BỘ GIÁO DỤC VÀ ĐÀO TẠO

**TRƯỜNG ĐẠI HỌC CMC**



**CÔNG TRÌNH THAM DỰ**

**HỘI NGHỊ SINH VIÊN NGHIÊN CỨU KHOA HỌC CẤP TRƯỜNG**

**ĐỀ TÀI: WEB AI DỰ ĐOÁN**

**GIÁ CỔ PHIẾU SỬ DỤNG MÔ HÌNH CHUỖI THỜI GIAN**

**NGÀNH CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

|  |  |
| --- | --- |
| **Sinh viên thực hiện:** | **Ngô Đức Thuận ( BIT220149)**  **Hoàng Minh Hải (BIT220049)** |
| **Giảng viên hướng dẫn:** | **TS. Ngô Hoàng Huy**  **TS. Phạm Thị Kim Dung** |

Hà Nội, ngày 18 tháng 9 năm 2024

**Tóm Tắt Công Trình**

Công trình này phát triển một hệ thống dự đoán giá cổ phiếu dựa trên phân tích dữ liệu lịch sử từ thị trường chứng khoán. Dữ liệu bao gồm giá mở cửa, giá đóng cửa, khối lượng giao dịch hàng ngày, và các chỉ số kinh tế liên quan.

Hệ thống tập trung vào việc xử lý chuỗi thời gian để phát hiện xu hướng và dự đoán biến động giá trong tương lai. Thách thức của bài toán bao gồm dữ liệu nhiễu, sự biến động khó lường của thị trường, và yêu cầu về hiệu suất xử lý cao. Ngoài ra, hệ thống cần cung cấp các dự đoán đáng tin cậy và nhanh chóng để hỗ trợ quyết định đầu tư kịp thời.

Mục tiêu chính là tạo ra mô hình dự đoán giúp nhà đầu tư tối ưu hóa danh mục đầu tư và quản lý rủi ro hiệu quả. Các mô hình dự đoán như ARIMA, LSTM và hồi quy tuyến tính sẽ được sử dụng để phân tích và ước lượng giá cổ phiếu.

Hệ thống không chỉ chú trọng đến tính chính xác và tốc độ, mà còn cần có giao diện dễ sử dụng, thân thiện, giúp người dùng khai thác thông tin một cách thuận tiện.

*Chúng tôi xin cam đoan rằng công trình nghiên cứu này là do chính nhóm thực hiện. Các số liệu, kết quả trong báo cáo này là trung thực và chưa được công bố trong bất kỳ công trình nào khác. Chúng tôi cũng cam kết rằng mọi sự giúp đỡ cho việc thực hiện nghiên cứu đều đã được cảm ơn, và các nguồn thông tin tham khảo đều đã được trích dẫn đầy đủ theo đúng quy định.*

**KÝ TÊN**

**GIẢNG VIÊN HƯỚNG DẪN SINH VIÊN THỰC HIỆN**

***(Ký và ghi rõ họ tên) (Ký và ghi rõ họ tên)***

**Ngô Hoàng Huy Phạm Thị Kim Dung Ngô Đức Thuận Hoàng Minh Hải**

**Thành viên dự án**

1. **Ngô Đức Thuận**



Mail: ngoducthuan24@gmail.com

Linh vực nhiên cứu: Web, AI, Security.

1. **Hoàng Minh Hải**

****

Mail: abuhmhai@gmail.com

Lĩnh vực nghiên cứu: AI, Web, Security.

**Giảng viên hướng dẫn**

1. **TS. Ngô Hoàng Huy**

****

Email: nhhuy@cmc-u.edu.vn

Mobile: 0904140022

Lĩnh vực nghiên cứu: AI, Deep learning(Speech recognition, Computer vision, DocumentAI, LLM, NLP, Time series...)

Lĩnh vực SE(phát triển phần mềm): Web MIS, Web AI, IoT.

----------------------

34 năm kinh nghiệm lập trình

2 books về lập trình

28 năm ở viện CNTT, Viện Hàn lâm KHCN VN

4.5 năm là dạy ở ĐH Điện lực, vị trí khoa CNTT

2 năm - nay ở Trường Đại học CMC.

Có 1 giải ba đồng đội nhân tài đất Việt 2019.

1. **TS. Phạm Thị Kim Dung**



Email: phamthikimzung@gmail.com

Mobile: 0969123899

Lĩnh vực nghiên cứu: AI, Deep learning(Computer vision, Time series...), Khai phá mạng xã hội.

Mục lục

Catalog

[Mục lục hình ảnh](#_Toc11863)

[Mục lục bảng biểu](#_Toc904)

[Bảng chú giải 1](#_Toc24986)

[Lời mở đầu 2](#_Toc30052)

[Chương 1. Giới thiệu tổng quan về bài toán 3](#_Toc6965)

[1.1 Mô tả bài toán 3](#_Toc1195)

[1.2. Mục tiêu của đề tài 3](#_Toc31803)

[1.3. Phát biểu bài toán dự đoán giá cổ phiếu 3](#_Toc27552)

[1.3.1. Khái niệm 3](#_Toc1379)

[1.3.2. Mục tiêu dự báo 3](#_Toc436)

[1.3.3. Dữ liệu đầu vào 4](#_Toc7729)

[1.3.4. Dự báo kết quả 4](#_Toc20146)

[1.3.5. Mô hình phổ biến 4](#_Toc22791)

[1.4. Phương pháp dự báo 4](#_Toc18734)

[1.4.1. Phương pháp hồi quy tuyến tính (Linear Regression) 4](#_Toc25401)

[1.4.2.Phương pháp ARIMA (AutoRegressive Integrated Moving Average) 4](#_Toc19911)

[1.4.3. Phương pháp Prophet 4](#_Toc15230)

[1.4.4. Mô hình mạng nơ-ron hồi quy sâu (RNN) 5](#_Toc9187)

[1.4.5. Mô hình LSTM (Long Short-Term Memory) 5](#_Toc26892)

[1.4.6. Mô hình GRU (Gated Recurrent Unit) 5](#_Toc9655)

[1.4.7. Mô hình Transformer 5](#_Toc30128)

[1.4.8. Kết hợp các phương pháp 5](#_Toc13451)

[Chương 2. Ứng dụng LSTM 6](#_Toc27029)

[2.1. Mô hình LSTM 6](#_Toc15316)

[2.1.1. Giới thiệu mô hình LSTM 6](#_Toc24179)

[2.1.2 Tại Sao Sử Dụng Mô Hình LSTM 6](#_Toc8759)

[2.2 . Cấu trúc cơ bản của LSTM 6](#_Toc11510)

[2.3. Kết quả thực nghiệm mô hình trên 2 mô hình LSTM và ARIMA 7](#_Toc16326)

[Chương 3. Phân tích thiết kế 15](#_Toc3743)

[3.1. Các sơ đồ nghiệp vụ 15](#_Toc27371)

[3.2. Hệ thống quản lý thông tin Web-based 15](#_Toc17559)

[3.3. Thiết kế kiến trúc phần mềm 16](#_Toc20155)

[3.2.1. Client Application 16](#_Toc2139)

[3.2.2. API Gateway 17](#_Toc2464)

[3.2.3. Microservices 17](#_Toc8276)

[3.2.4. Web Hook 17](#_Toc24032)

[3.2.5. Cơ sở Dữ liệu 17](#_Toc27828)

[3.4. Thiết kế cơ sở dữ liệu 17](#_Toc17540)

[Chương 4. Xây dựng website 20](#_Toc17323)

[4.1. Giới thiệu về công cụ sử dụng 20](#_Toc24388)

[4.1.1.Cơ sở dữ liệu: Microsoft SQL Server 20](#_Toc11981)

[4.1.2. Framework Argular 20](#_Toc4576)

[4.1.3. Ngôn ngữ Python 20](#_Toc2450)

[4.1.4. Môi trường phát triển: Visual Studio 20](#_Toc15855)

[4.1.5. Môi trường phát triển: Visual Studio Code 20](#_Toc5102)

[4.1.6. Google Colab 20](#_Toc10328)

[4.1.7. VietQR API 20](#_Toc6356)

[4.1.8. Casso API 20](#_Toc11110)

[4.2. Xây dựng các chức năng 20](#_Toc16405)

[Chương 5: Kết quả đạt được 26](#_Toc29573)

[Chương 6: Kết luận và hướng phát triển 27](#_Toc16555)

[Chương 7: Tài liệu tham khảo 28](#_Toc12042)

# Mục lục hình ảnh

[*Hình* 1 *. Biểu đồ so sánh sự mất mát của tập huấn luyện và tập kiểm tra*](#_bookmark6) 9

[*Hình* 2 *. Biểu đồ so sánh giá dự đoán trên tập kiểm tra so với giá gốc*](#_bookmark7) 10

[*Hình 3* *. Phân tích kết quả huấn luyện và kiểm định mô hình*](#_bookmark7) 11

[*Hình 4* *.Kết quả dự đoán giá cổ phiếu trên 5 mã cổ phiếu*](#_bookmark7) 14

[*Hình 5* *. Sơ đồ UseCasse*](#_bookmark10) 16

[*Hình 6* *. Kiến trúc phần mềm*](#_bookmark10) 17

[*Hình 7* *. Database Diagram*](#_bookmark30) *..*19

[*Hình 8* *. Màn hình đăng nhập*](#_bookmark30) *..*22

*Hình 9 .* *Màn hình hiển thị tổng quan các mã cổ phiếu* ..22

[*Hình* 10 *. Hiển thị thông tin chi tiết từng mã cổ phiếu*](#_bookmark32) ..23

[*Hình* 11 *. File excel sau khi xuất*](#_bookmark34) ..23

[*Hình* 12 *. So sánh hai mã cổ phiếu*](#_bookmark35) ..24

[*Hình 13* *. Màn hình hiển thị thông tin người dùng*](#_bookmark35) ..24

[*Hình 14* *. Nạp tiền vào tài khoản*](#_bookmark35) ..25

[*Hình 15* *. Màn hình thông báo khi bấm vào “Predict Stock” ..*25](#_bookmark35)

[*Hình 16* *. Dự đoán giá cổ phiếu ..*](#_bookmark35)..26

[*Hình 17* *. Nhập mã cổ phiếu tìm kiếm với “stb” ..*](#_bookmark35)..26

# Bảng chú giải

|  |  |
| --- | --- |
| **Ký hiệu** | **Định nghĩa** |
| LSTM | Long short term memory |
| ARIMA | Autoregressive Integrated Moving Average |
| MAE | Mean Absolute Error |
| SEM | Standard Error of the Mean |
| MAPE | Mean Absolute Percentage Error |
| RMSE | Root Mean Square Error |

# Lời mở đầu

Trong bối cảnh toàn cầu hóa và sự phát triển nhanh chóng của công nghệ, thị trường tài chính đang trải qua những biến động chưa từng có. Những thay đổi về kinh tế, chính trị, và các sự kiện xã hội trên thế giới có thể tác động mạnh mẽ đến giá trị của các tài sản tài chính, khiến việc dự đoán trở nên phức tạp và đầy rủi ro. Thị trường chứng khoán, vốn được biết đến với tính chất nhạy cảm và dễ biến động, là một minh chứng rõ ràng cho sự khó lường của các yếu tố này. Các nhà đầu tư phải đối mặt với thách thức không nhỏ khi cố gắng hiểu và dự đoán xu hướng của thị trường, nhằm tối ưu hóa lợi nhuận và bảo vệ danh mục đầu tư của mình trước những biến động bất ngờ.

Trước bối cảnh đó, sự ra đời và phát triển của khoa học dữ liệu đã mở ra những cơ hội mới trong việc phân tích và dự đoán thị trường tài chính. Các mô hình dự đoán chuỗi thời gian, kết hợp với sức mạnh của trí tuệ nhân tạo và học máy, đã trở thành những công cụ không thể thiếu đối với các nhà đầu tư. Chúng không chỉ giúp phân tích dữ liệu lịch sử mà còn có khả năng nhận diện các mẫu hành vi lặp lại, từ đó đưa ra những dự báo về xu hướng và biến động của giá cổ phiếu.

Việc áp dụng các phương pháp dự đoán chuỗi thời gian trong lĩnh vực tài chính đã và đang thay đổi cách tiếp cận của các nhà đầu tư đối với thị trường. Thay vì dựa vào trực giác hay kinh nghiệm cá nhân, họ có thể sử dụng các công cụ phân tích mạnh mẽ để đánh giá và dự báo một cách chính xác hơn. Trong một môi trường kinh doanh ngày càng cạnh tranh, khả năng này không chỉ giúp tối ưu hóa lợi nhuận mà còn đóng vai trò quan trọng trong việc quản lý rủi ro, giúp nhà đầu tư bảo vệ tài sản của mình trước những biến động khó lường của thị trường.

Với sự phát triển liên tục của công nghệ và sự gia tăng về khả năng xử lý dữ liệu, dự đoán chuỗi thời gian trong thị trường hứa hẹn sẽ tiếp tục tiến bộ, mang lại những lợi thế cạnh tranh vượt trội cho các nhà đầu tư có khả năng tận dụng hiệu quả các công cụ phân tích hiện đại này.

# Chương 1. Giới thiệu tổng quan về bài toán

## 1.1 Mô tả bài toán

Bài toán đặt ra là xây dựng một hệ thống dự đoán giá cổ phiếu dựa trên việc phân tích dữ liệu lịch sử của thị trường chứng khoán. Dữ liệu bao gồm các thông số như giá mở cửa, giá đóng cửa, khối lượng giao dịch hàng ngày, các chỉ số kinh tế và tài chính liên quan. Hệ thống cần phải xử lý và phân tích chuỗi thời gian của những dữ liệu này để nhận diện các xu hướng và mẫu hành vi, từ đó dự đoán biến động giá cổ phiếu trong tương lai.

Bên cạnh đó, hệ thống cũng cần phải đối mặt với những thách thức về dữ liệu nhiễu, sự bất ổn của thị trường, yêu cầu về hiệu suất xử lý cao để cung cấp các dự báo kịp thời và đáng tin cậy cho người dùng. Đồng thời, hệ thống cần có một giao diện thân thiện với người dùng, dễ sử dụng, giúp người dùng dễ dàng truy cập và hiểu rõ các thông tin dự đoán.

## 1.2. Mục tiêu của đề tài

Mục tiêu là phát triển mô hình dự đoán giá cổ phiếu giúp nhà đầu tư ra quyết định hiệu quả, tối ưu hóa danh mục và quản lý rủi ro. Hệ thống cần đảm bảo tính chính xác cao, xử lý nhanh chóng để đáp ứng yêu cầu đầu tư kịp thời.

Đồng thời, giao diện người dùng phải thân thiện, dễ sử dụng, cho phép người dùng truy cập và khai thác thông tin dự đoán một cách tối ưu.

## 1.3. Phát biểu bài toán dự đoán giá cổ phiếu

Dự đoán giá cổ phiếu là quá trình sử dụng dữ liệu lịch sử và các mô hình phân tích để ước lượng giá trị của cổ phiếu trong tương lai. Các mô hình thường được sử dụng để thực hiện dự đoán bao gồm ARIMA, LSTM và hồi quy tuyến tính. Những mô hình này giúp phân tích các yếu tố ảnh hưởng đến giá cổ phiếu, từ đó đưa ra dự báo về giá tại các thời điểm sắp tới.

*1.3.1. Khái niệm*

Dự báo giá cổ phiếu dựa trên dữ liệu lịch sử là việc sử dụng các chuỗi dữ liệu đã được thu thập từ trước để ước lượng giá trị cổ phiếu tại các thời điểm tương lai. Việc này có thể thực hiện theo chu kỳ ngày, tuần hoặc tháng, tùy thuộc vào yêu cầu của nhà đầu tư hoặc mô hình dự đoán.

*1.3.2. Mục tiêu dự báo*

Mục tiêu chính của bài toán dự đoán giá cổ phiếu là ước lượng được giá cổ phiếu tại thời điểm t tương lai, dựa trên các yếu tố đã biết từ quá khứ. Việc dự báo có thể được thực hiện theo chu kỳ ngày, tức là dự đoán giá đóng cửa của cổ phiếu vào các ngày t+1, t+2,... đến t+10, giúp nhà đầu tư hiểu rõ hơn về xu hướng giá trong ngắn hạn.

*1.3.3. Dữ liệu đầu vào*

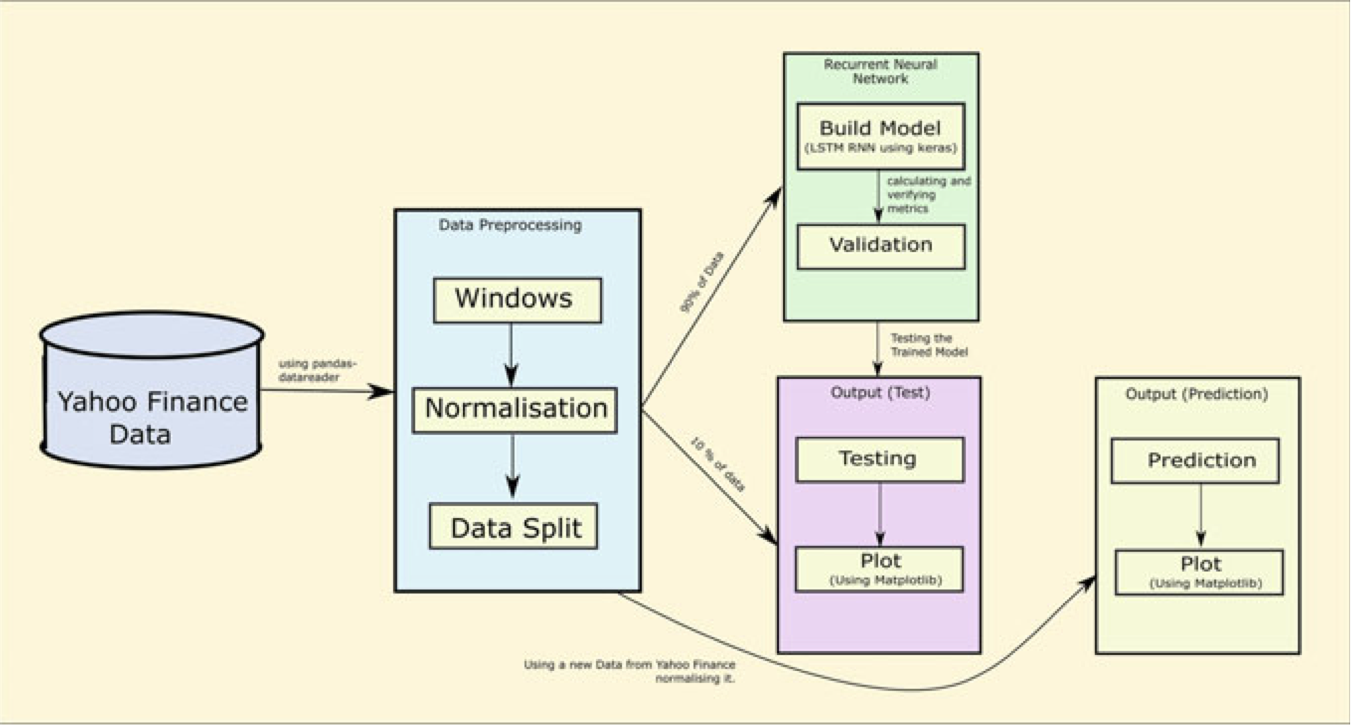
Dữ liệu đầu vào cho bài toán bao gồm giá cổ phiếu hàng ngày, cụ thể là các giá mở cửa, giá đóng cửa, giá cao nhất, giá thấp nhất và khối lượng giao dịch.

*1.3.4. Dự báo kết quả*

Kết quả dự báo sẽ là giá đóng cửa của cổ phiếu cho các ngày tiếp theo, từ t+1 cho đến t+10. Việc dự báo theo chu kỳ này giúp cung cấp thông tin kịp thời cho nhà đầu tư, cho phép họ điều chỉnh chiến lược đầu tư phù hợp với xu hướng thị trường.

*1.3.5. Mô hình phổ biến*

Những mô hình phổ biến trong dự báo giá cổ phiếu bao gồm ARIMA (AutoRegressive Integrated Moving Average), LSTM (Long Short-Term Memory), và hồi quy tuyến tính.



## 1.4. Phương pháp dự báo

*1.4.1. Phương pháp hồi quy tuyến tính (Linear Regression)*

Đây là một trong những phương pháp đơn giản nhất, sử dụng mối quan hệ tuyến tính giữa các biến đầu vào (các yếu tố ảnh hưởng đến giá cổ phiếu) và biến đầu ra (giá cổ phiếu). Phương pháp này phù hợp khi dữ liệu có xu hướng tuyến tính. Tuy nhiên, vì thị trường - thường có tính biến động và không tuyến tính, hồi quy tuyến tính thường không phù hợp với các bài toán phức tạp.

*1.4.2.Phương pháp ARIMA (AutoRegressive Integrated Moving Average)*

ARIMA là một mô hình dự đoán chuỗi thời gian rất phổ biến trong các bài toán tài chính. Mô hình này kết hợp giữa hồi quy tự hồi quy (AR), sai phân (I), và trung bình trượt (MA) để dự đoán chuỗi thời gian. ARIMA phù hợp với các chuỗi thời gian có tính tuần hoàn, nhưng khó khăn khi xử lý những biến động ngẫu nhiên mạnh mẽ từ thị trường chứng khoán.

*1.4.3. Phương pháp Prophet*

Prophet là một công cụ do Facebook phát triển, có khả năng xử lý các xu hướng, tính mùa vụ và những bất thường trong chuỗi thời gian. Prophet rất mạnh mẽ trong việc xử lý dữ liệu có độ biến động không ổn định và mất mát, điều này làm cho nó hữu ích trong dự đoán giá cổ phiếu.

*1.4.4. Mô hình mạng nơ-ron hồi quy sâu (RNN)*

Mạng nơ-ron hồi quy sâu (RNN) là một mô hình mạnh mẽ để xử lý các bài toán chuỗi thời gian. RNN có khả năng ghi nhớ các thông tin từ quá khứ và sử dụng chúng để dự đoán các giá trị trong tương lai. Tuy nhiên, RNN truyền thống gặp phải vấn đề về biến mất độ dốc khi chuỗi thời gian quá dài.

*1.4.5. Mô hình LSTM (Long Short-Term Memory)*

LTM là một phiên bản cải tiến của RNN, có khả năng học được các mẫu từ các chuỗi thời gian dài mà không gặp vấn đề về biến mất hoặc bùng nổ độ dốc. LSTM có thể xử lý tốt các chuỗi thời gian phức tạp, không chỉ giúp dự đoán chính xác mà còn thích hợp với dữ liệu có độ biến động cao, như giá cổ phiếu. LSTM sử dụng các cổng điều khiển để duy trì hoặc lãng quên thông tin, đảm bảo mô hình có thể học từ những dữ liệu quan trọng nhất.

*1.4.6. Mô hình GRU (Gated Recurrent Unit)*

GU là một biến thể khác của RNN, tương tự như LSTM nhưng với cấu trúc đơn giản hơn. GRU cũng có khả năng xử lý chuỗi thời gian dài, nhưng tốc độ huấn luyện nhanh hơn LSTM nhờ vào số lượng tham số ít hơn.

*1.4.7. Mô hình Transformer*

Mô hình Transformer, ban đầu phát triển cho các bài toán xử lý ngôn ngữ tự nhiên, cũng đang được áp dụng cho dự báo chuỗi thời gian. Transformer sử dụng cơ chế Attention để học các mẫu trong chuỗi thời gian mà không cần phải nhớ quá nhiều bước trước đó, làm cho nó trở thành một công cụ mạnh mẽ cho việc dự đoán chính xác trên chuỗi thời gian dài.

*1.4.8. Kết hợp các phương pháp*

Tổng thực tế, các mô hình dự báo giá cổ phiếu thường không chỉ sử dụng một mô hình duy nhất. Việc kết hợp nhiều mô hình (ensemble learning) như sử dụng cả ARIMA và LSTM có thể tăng cường độ chính xác của dự báo bằng cách tận dụng những điểm mạnh của từng phương pháp. Ví dụ, ARIMA có thể dự đoán tốt xu hướng dài hạn trong khi LSTM có thể xử lý các biến động ngắn hạn.

# Chương 2. Ứng dụng LSTM

## Mô hình LSTM

*2.1.1. Giới thiệu mô hình LSTM*

LSTM là một loại mạng nơ-ron tái phát (RNN). RNN là một loại mạng thần kinh nhân tạo mạnh mẽ có thể duy trì nội bộ bộ nhớ của đầu vào. Điều này làm cho chúng đặc biệt phù hợp để giải quyết các vấn đề liên quan đến dữ liệu tuần tự như chuỗi thời gian. Tuy nhiên, RNN thường gặp phải một vấn đề gọi là gradient biến mất dẫn đến việc học mô hình trở nên quá chậm hoặc dừng hoàn toàn. LSTM được tạo ra vào những năm 1990 để giải quyết vấn đề này.

*2.1.2 Tại Sao Sử Dụng Mô Hình LSTM*

LSTM được chọn vì nó có khả năng xử lý các chuỗi dữ liệu dài và giữ thông tin theo thời gian, điều này rất quan trọng trong dự đoán giá cổ phiếu nơi mà các mẫu và xu hướng trong quá khứ ảnh hưởng đến giá trong tương lai. LSTM giúp mô hình hóa sự phụ thuộc lâu dài trong dữ liệu, cải thiện độ chính xác của dự đoán.

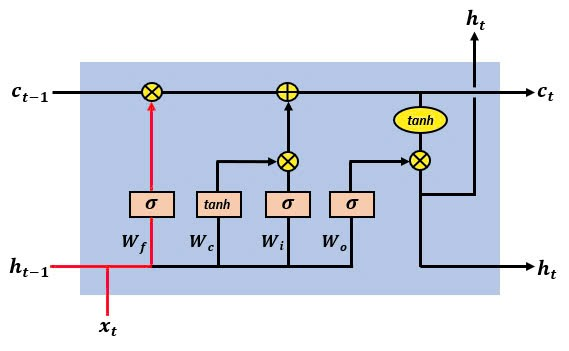
Giải quyết vấn đề biến mất hoặc bùng nổ độ dốc (vanishing/exploding gradient): Một vấn đề lớn với các mô hình RNN truyền thống là khi số lượng bước thời gian tăng lên, thông tin từ các bước đầu có thể bị mất đi do sự biến mất hoặc bùng nổ độ dốc khi huấn luyện. LSTM, với cấu trúc cổng nhớ và khả năng kiểm soát dòng thông tin, giúp giảm thiểu vấn đề này, cho phép mô hình học từ các chuỗi thời gian dài mà không mất thông tin quan trọng.

Khả năng xử lý sự không ổn định của thị trường:Thị trường chứng khoán thường có những biến động khó lường. LSTM, với khả năng học các mẫu phức tạp, có thể giúp mô hình học được những mẫu biến động tiềm ẩn và dự đoán chính xác hơn, ngay cả trong các giai đoạn thị trường có biến động mạnh.

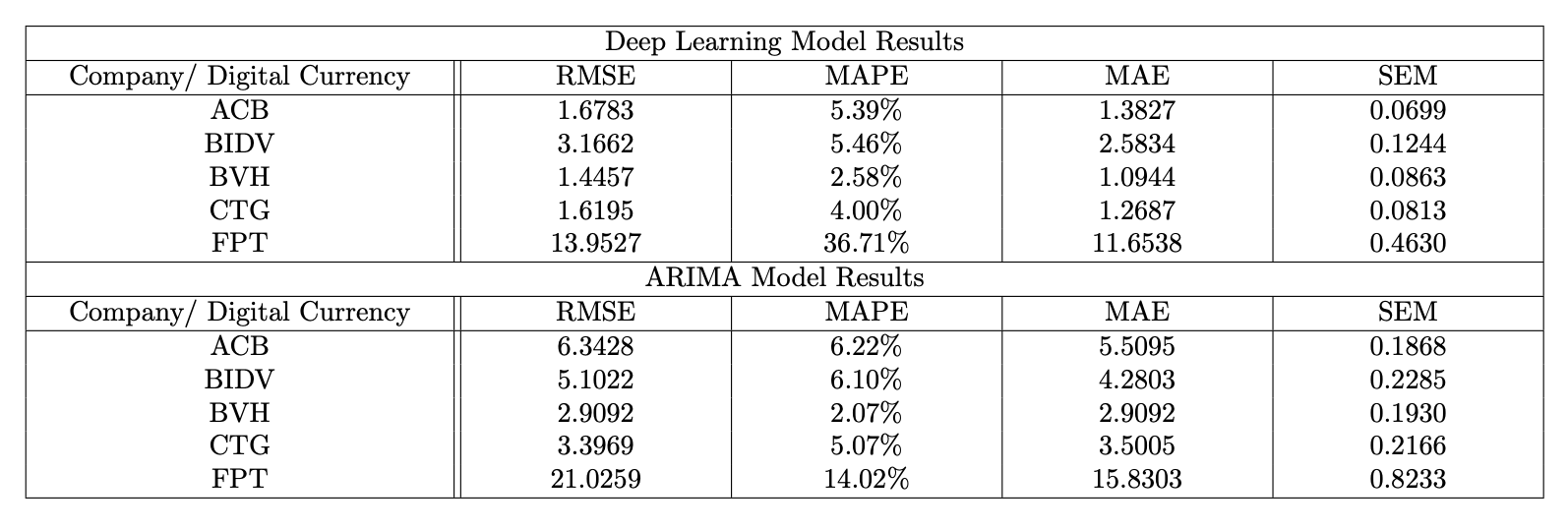
## . Cấu trúc cơ bản của LSTM

TM có bộ nhớ dài hơn và có thể học hỏi từ các đầu vào được tách biệt với nhau do độ trễ thời gian dài. LSTM có một cấu trúc đặc biệt với ba cổng chính và một ô nhớ:

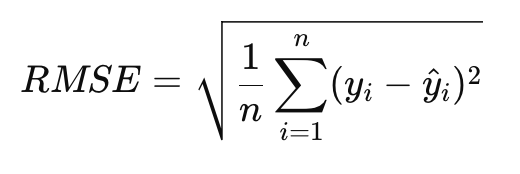
* + - **Cổng Nhập (Input Gate)**: Quyết định thông tin nào sẽ được thêm vào ô nhớ. Nó gồm hai phần:
      * **Cổng Sigmoid**: Xác định các giá trị nào nên được cập nhật (0 cho không cập nhật, 1 cho cập nhật hoàn toàn).
      * **Cổng Tanh**: Tạo ra các giá trị candidate mới, có thể được thêm vào ô nhớ.
    - **Cổng Quên (Forget Gate)**: Quyết định thông tin nào sẽ được loại bỏ khỏi ô nhớ. Cổng này sử dụng hàm sigmoid để chọn các giá trị cần quên.
    - **Cổng Đầu ra (Output Gate)**: Quyết định thông tin nào sẽ được đưa ra ngoài ô nhớ. Cổng này bao gồm một hàm sigmoid và một hàm tanh để điều chỉnh giá trị đầu ra dựa trên ô nhớ.



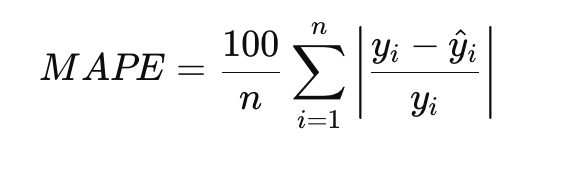
### *2.3. Kết quả thực nghiệm mô hình trên 2 mô hình LSTM và ARIMA*



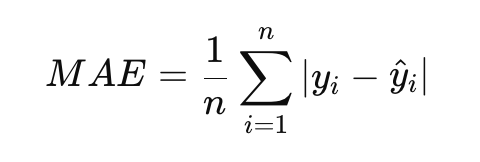
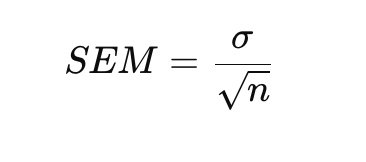
**RMSE (Root Mean Square Error)**: Là đại lượng đo độ lệch giữa giá trị thực tế và giá trị dự đoán, càng nhỏ thì mô hình dự đoán càng tốt.



**MAPE (Mean Absolute Percentage Error)**: Đo độ lệch phần trăm trung bình tuyệt đối giữa dự đoán và giá trị thực. Tương tự, giá trị càng nhỏ thì mô hình càng chính xác.



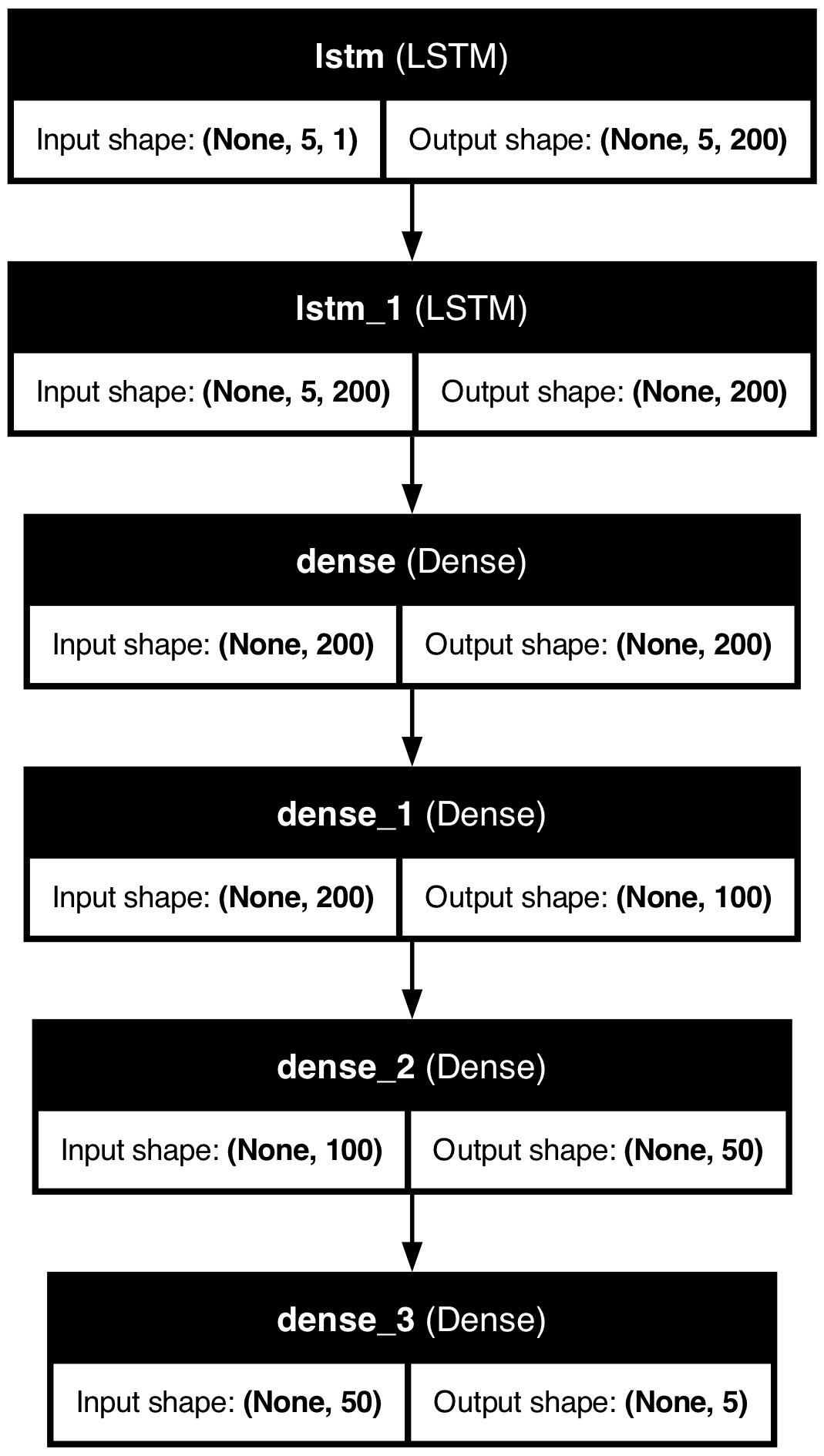
**MAE (Mean Absolute Error)**: Đo độ lệch tuyệt đối trung bình, cũng tương tự RMSE nhưng không khuếch đại lỗi lớn.



**SEM (Standard Error of the Mean)**: Sai số chuẩn của trung bình, đo mức độ biến động của dữ liệu.

*So sánh giữa hai mô hình:*

*Deep Learning Model:*



Có kết quả RMSE và MAE thấp hơn đáng kể so với mô hình ARIMA ở phần lớn các trường hợp (trừ FPT).

Đặc biệt, đối với các công ty như ACB, BIDV, BVH, và CTG, mô hình Deep Learning cho thấy kết quả RMSE, MAPE, và MAE tốt hơn, chứng tỏ mô hình này có độ chính xác cao hơn trong dự đoán.

Tuy nhiên, đối với công ty FPT, mô hình Deep Learning có giá trị RMSE (13.9527) và MAE (11.6538) rất cao, trong khi ARIMA có RMSE và MAE tốt hơn trong trường hợp này, chứng tỏ mô hình ARIMA phù hợp hơn cho FPT.

*ARIMA Model:*

Có RMSE và MAE cao hơn trong hầu hết các trường hợp (trừ FPT), đặc biệt đối với ACB và CTG, điều này cho thấy ARIMA kém chính xác hơn so với Deep Learning ở các công ty này.

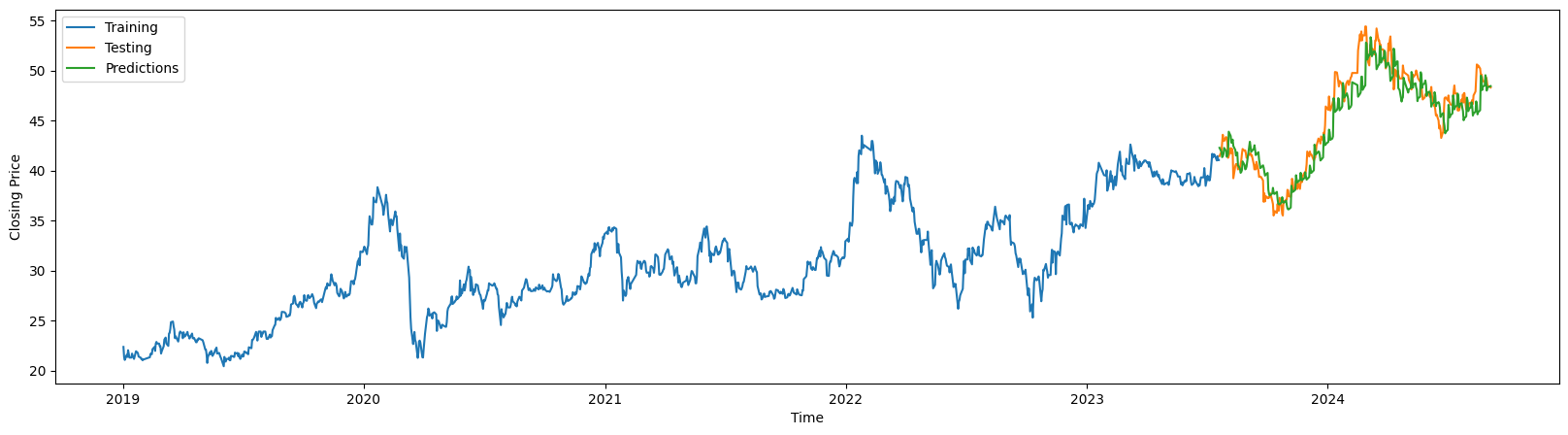
Tuy nhiên, với FPT, ARIMA có RMSE (21.0259) và MAE (15.8303), mặc dù không nhỏ, nhưng vẫn thấp hơn so với Deep Learning, điều này cho thấy mô hình này hiệu quả hơn cho trường hợp cụ thể này.

*Tổng kết:*

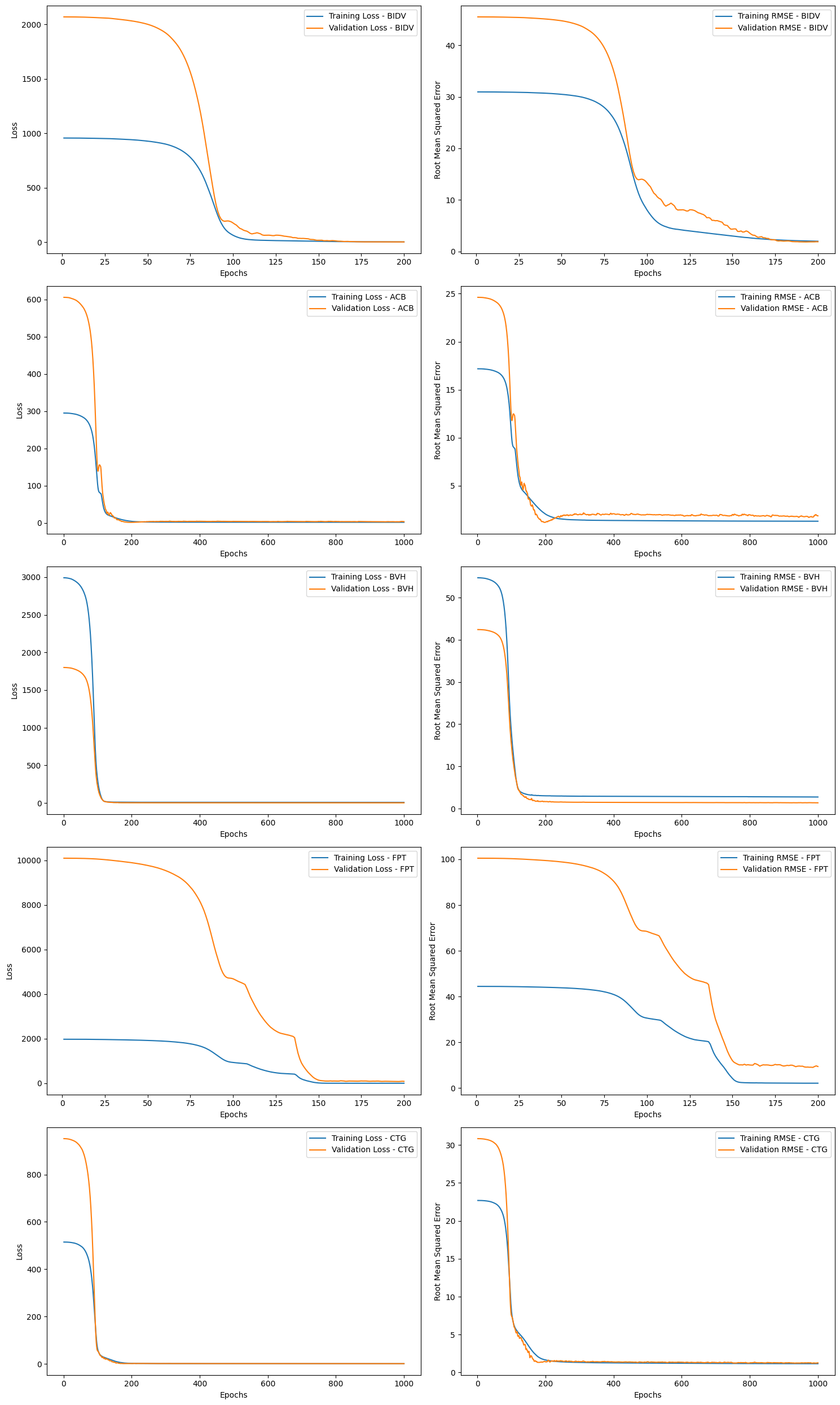
Deep Learning Model nhìn chung vượt trội hơn so với ARIMA về độ chính xác (RMSE, MAPE, MAE) trong các trường hợp ngoại trừ FPT.

ARIMA tỏ ra không hiệu quả trong việc dự đoán với độ chính xác cao, trừ trường hợp FPT, nơi mà ARIMA có kết quả tốt hơn so với Deep Learning.

****

*Hình* *1. Biểu đồ so sánh giá dự đoán trên tập kiểm tra so với giá gốc của ARIMA modelHình* *2. Biểu đồ so sánh giá dự đoán trên tập kiểm tra so với giá gốc của Deep Learning Model*

# 



*Hình 3: Phân tích kết quả huấn luyện và kiểm định mô hình*

**Phân tích Kết Quả Huấn Luyện và Kiểm Định Mô Hình**

Hình 1 trình bày kết quả huấn luyện và kiểm định mô hình dự đoán giá cổ phiếu trên năm mã cổ phiếu: BIDV (BID), ACB, BVH, FPT và CTG. Mỗi hàng tương ứng với một mã cổ phiếu, gồm hai biểu đồ: biểu đồ bên trái thể hiện sự thay đổi của Loss (lỗi huấn luyện và lỗi kiểm định) trong quá trình huấn luyện mô hình, còn biểu đồ bên phải thể hiện sự biến thiên của RMSE (Căn bậc hai của lỗi trung bình bình phương) cho tập huấn luyện và kiểm định theo số epoch.

***1. Mã cổ phiếu BIDV (BID):***

Kết quả cho thấy "Training Loss" và "Validation Loss" giảm mạnh sau khoảng 50 epoch và đạt đến giá trị hội tụ ổn định sau 150 epoch. Sự hội tụ này phản ánh khả năng mô hình nhanh chóng học được các quy luật tiềm ẩn trong dữ liệu. Tương tự, RMSE giảm rõ rệt và hội tụ tại cùng điểm, thể hiện sự tương đồng giữa khả năng dự đoán trên cả dữ liệu huấn luyện và kiểm định. Điều này cho thấy mô hình không bị overfitting và có thể khái quát tốt trên dữ liệu mới.

***2. Mã cổ phiếu ACB:***

"Training Loss" và "Validation Loss" giảm nhanh chóng trong 50 epoch đầu tiên và sau đó tiếp tục giảm nhẹ trước khi hội tụ về một giá trị rất thấp sau khoảng 200 epoch. Đường biểu diễn RMSE cũng cho thấy xu hướng tương tự, với mức giảm mạnh trong giai đoạn đầu và hội tụ sau khoảng 200 epoch. Kết quả này chỉ ra rằng mô hình có khả năng dự đoán chính xác cả trên tập huấn luyện lẫn tập kiểm định, đồng thời phản ánh hiệu suất dự đoán ổn định trên toàn bộ tập dữ liệu.

***3. Mã cổ phiếu BVH:***

Đối với BVH, cả "Training Loss" và "Validation Loss" đều cho thấy sự hội tụ nhanh chóng ngay từ những epoch đầu tiên và đạt trạng thái ổn định sau khoảng 200 epoch. Điều này phản ánh khả năng học nhanh chóng của mô hình đối với dữ liệu. Biểu đồ RMSE cũng xác nhận sự hội tụ tương đồng giữa tập huấn luyện và kiểm định, nhấn mạnh rằng mô hình có khả năng tái tạo các xu hướng quan trọng trong dữ liệu một cách nhất quán.

1. ***Mã cổ phiếu FPT:***

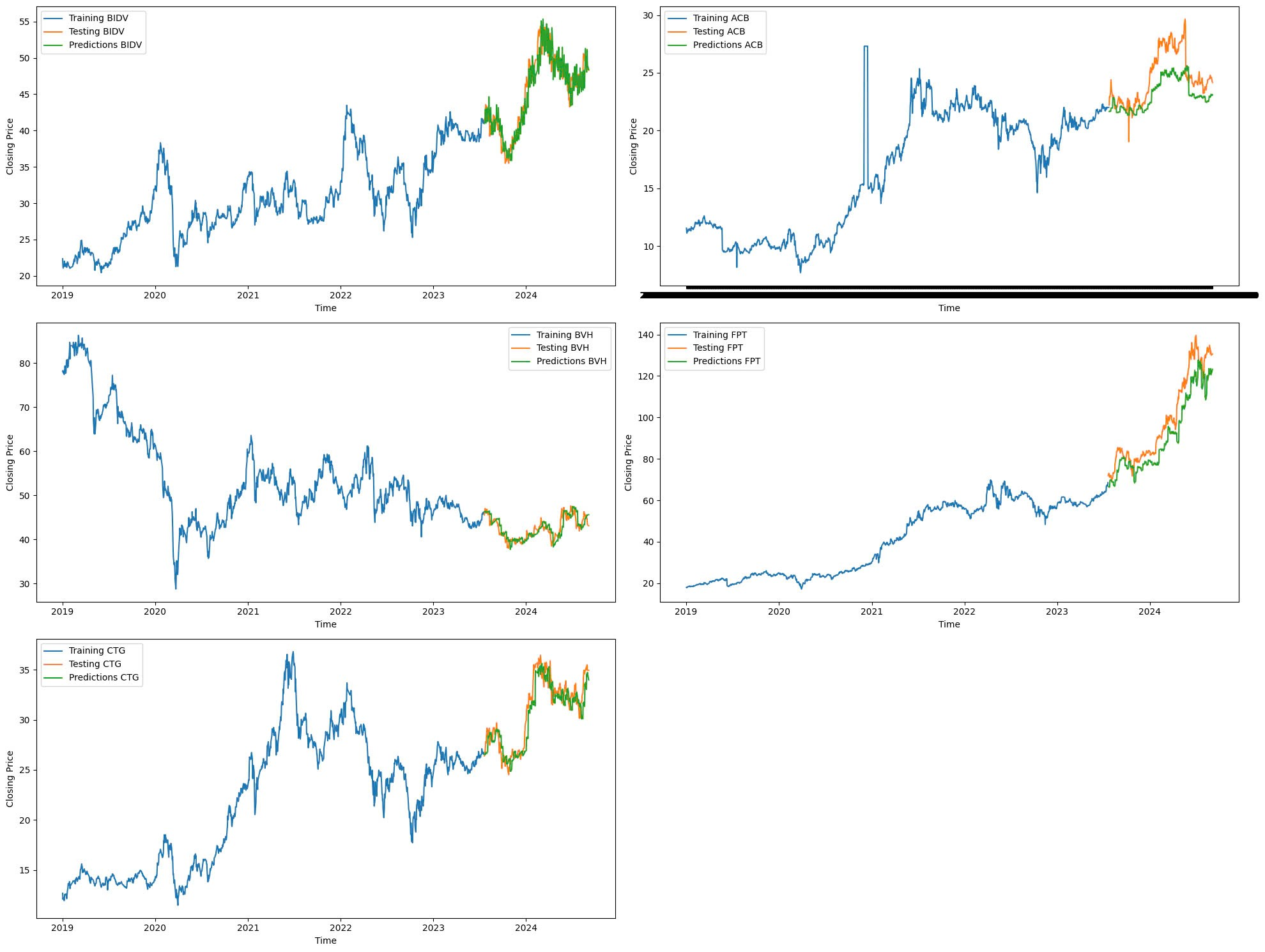
Biểu đồ của FPT cho thấy một mức độ dao động đáng kể trong "Validation Loss", đặc biệt là trong khoảng từ 50 đến 150 epoch, điều này có thể ám chỉ sự khó khăn của mô hình trong việc học quy luật từ dữ liệu kiểm định tại giai đoạn này. Tuy nhiên, sau khi vượt qua giai đoạn này, mô hình hội tụ tốt và cả "Training Loss" và "Validation Loss" đều giảm về mức thấp sau khoảng 200 epoch. Biểu đồ RMSE cũng cho thấy sự dao động tương tự, nhưng cuối cùng cả "Training RMSE" và "Validation RMSE" đều hội tụ về giá trị nhỏ. Điều này cho thấy mô hình đã cải thiện hiệu suất dần dần và có khả năng dự báo xu hướng giá cổ phiếu chính xác, mặc dù có sự bất ổn ngắn hạn trong quá trình huấn luyện.

***5. Mã cổ phiếu CTG:***

Đối với cổ phiếu CTG, mô hình cho thấy sự giảm mạnh của cả "Training Loss" và "Validation Loss" trong 50 epoch đầu tiên, với sự hội tụ rõ ràng sau khoảng 100 epoch. Điều này phản ánh khả năng học hiệu quả của mô hình. RMSE cũng cho thấy sự hội tụ tốt giữa tập huấn luyện và kiểm định, chứng minh rằng mô hình có thể đưa ra các dự đoán chính xác trên tập kiểm định với mức lỗi rất nhỏ, thể hiện khả năng dự đoán ổn định.

***Tổng kết:***

Kết quả huấn luyện và kiểm định cho thấy mô hình đã đạt được mức độ hội tụ ổn định trên tất cả các mã cổ phiếu, với cả Loss và RMSE giảm đáng kể và hội tụ về giá trị rất nhỏ sau quá trình huấn luyện. Đối với các mã cổ phiếu BID, ACB, BVH và CTG, mô hình nhanh chóng hội tụ sau khoảng 100 đến 200 epoch mà không gặp nhiều biến động lớn, phản ánh khả năng khái quát hóa tốt. Đối với cổ phiếu FPT, mặc dù có sự dao động trong giai đoạn giữa của quá trình huấn luyện, mô hình vẫn đạt được kết quả tốt sau khi số epoch tăng lên, cho thấy khả năng xử lý các đặc tính phức tạp của dữ liệu. Nhìn chung, kết quả biểu thị hiệu suất cao của mô hình trong việc dự đoán xu hướng giá cổ phiếu với độ chính xác cao và khả năng khái quát tốt trên tập dữ liệu kiểm định.



*Hình 4:Kết quả dự đoán giá cổ phiếu trên 5 mã cổ phiếu*

Hình 2 thể hiện kết quả dự đoán giá cổ phiếu của mô hình trên năm mã cổ phiếu: BIDV (BID), ACB, BVH, FPT và CTG. Mỗi biểu đồ cho thấy sự so sánh giữa giá thực tế trong giai đoạn huấn luyện và kiểm định cùng với giá trị dự đoán của mô hình trong giai đoạn kiểm định. Kết quả này giúp đánh giá sự chính xác của mô hình dự báo giá cổ phiếu cho từng mã và cung cấp cơ sở so sánh giữa các mã cổ phiếu.

Đối với mã cổ phiếu BID, mô hình đã dự đoán rất sát với giá trị thực tế trong suốt giai đoạn kiểm định. Điều này thể hiện qua việc giá trị dự đoán gần như trùng khớp với giá trị thực tế ở tất cả các điểm quan trọng, cho thấy mô hình có khả năng nắm bắt xu hướng ngắn hạn và dài hạn của thị trường. Tương tự, mô hình cũng hoạt động tốt với mã ACB, khi các dự đoán của mô hình bám sát xu hướng thực tế, dù có một số dao động nhỏ trong các giai đoạn biến động mạnh. Các giai đoạn tăng trưởng và suy thoái của ACB được dự đoán tương đối chính xác, chứng tỏ khả năng dự báo tốt của mô hình.

Đối với mã BVH, mặc dù có một số sai lệch nhỏ trong giai đoạn cuối, mô hình vẫn thể hiện khả năng dự đoán tốt, đặc biệt là trong các giai đoạn có xu hướng giảm mạnh. Khả năng bám sát xu hướng giá thực tế cho thấy mô hình hoạt động ổn định với mã cổ phiếu này. Trong khi đó, đối với mã FPT, mô hình đã nắm bắt rất tốt xu hướng tăng trưởng dài hạn, mặc dù có một số dao động giữa giá trị dự đoán và giá thực tế trong giai đoạn biến động ngắn hạn. Điều này chứng tỏ mô hình có khả năng dự đoán chính xác các xu hướng lớn, nhưng vẫn có thể gặp khó khăn trong việc dự báo những biến động ngắn hạn với cổ phiếu có xu hướng biến động mạnh như FPT.

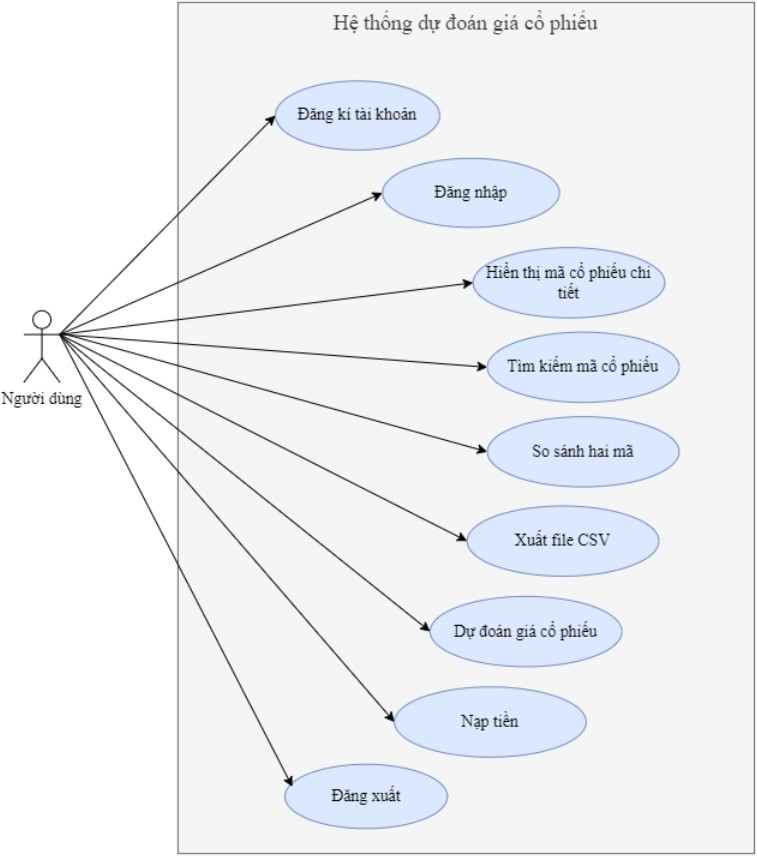
Cuối cùng, mã cổ phiếu CTG cũng cho kết quả dự đoán rất sát với giá trị thực tế, đặc biệt là trong các giai đoạn có xu hướng tăng và giảm. Mô hình đã không chỉ dự đoán tốt các xu hướng dài hạn mà còn nắm bắt được những thay đổi nhỏ trong giá trị cổ phiếu, cho thấy khả năng khái quát hóa tốt.

So sánh chung giữa các mã cổ phiếu cho thấy, mô hình dự báo hoạt động rất hiệu quả, đặc biệt là với các mã BID, ACB và CTG, khi sự sai lệch giữa giá trị dự đoán và giá trị thực tế là rất nhỏ. Đối với các mã có xu hướng biến động mạnh như FPT và BVH, mô hình vẫn giữ được khả năng dự đoán chính xác xu hướng tổng thể, dù có một số sai lệch nhỏ trong ngắn hạn. Nhìn chung, sự khác biệt giữa các mã cổ phiếu chủ yếu phụ thuộc vào mức độ biến động của từng mã. Các mã có sự biến động lớn hơn có xu hướng gặp khó khăn hơn trong việc dự báo ngắn hạn, nhưng mô hình vẫn duy trì hiệu suất tốt và nắm bắt các điểm chuyển pha quan trọng.

Kết quả này cho thấy mô hình có khả năng dự báo tốt trên cả năm mã cổ phiếu, với độ chính xác cao và khả năng khái quát hóa tốt. Điều này chứng minh mô hình có tiềm năng ứng dụng trong thực tiễn, hỗ trợ nhà đầu tư trong việc dự đoán xu hướng giá cổ phiếu và ra quyết định đầu tư.

# Chương 3. Phân tích thiết kế

## 3.1. Sơ đồ nghiệp vụ UseCase



*Hình* *5. Sơ đồ UseCasse*

## 3.2. Hệ thống quản lý thông tin Web-based

Trong bối cảnh phát triển ứng dụng web dự đoán giá cổ phiếu, việc lựa chọn một hệ thống quản lý thông tin web-based (Web-based Management Information System - MIS) là rất quan trọng để đảm bảo khả năng quản lý hiệu quả, truy cập linh hoạt, và vận hành mượt mà.

*3.2.1. Cập nhật và bảo trì dễ dàng*

Với hệ thống web-based, mọi cập nhật và nâng cấp cho phần mềm đều có thể được triển khai trực tiếp trên server mà không cần cài đặt riêng trên từng thiết bị người dùng.

Điều này giúp giảm thời gian bảo trì và đảm bảo rằng tất cả người dùng luôn có thể truy cập phiên bản mới nhất của hệ thống mà không bị gián đoạn.

*3.2.2. Truy cập từ xa*

Hệ thống cho phép người dùng làm việc từ bất kỳ đâu chỉ cần có kết nối internet và một trình duyệt web. Điều này giúp cải thiện tính linh hoạt trong công việc và hỗ trợ người dùng từ xa mà không cần yêu cầu cài đặt phần mềm đặc biệt.

*3.2.3. Triển khai cập nhật nhanh chóng*

Nhờ việc vận hành hoàn toàn qua môi trường web, các bản cập nhật có thể được triển khai ngay lập tức mà không cần phải phát hành các phiên bản client-side.

*3.2.4. Chi phí thấp*

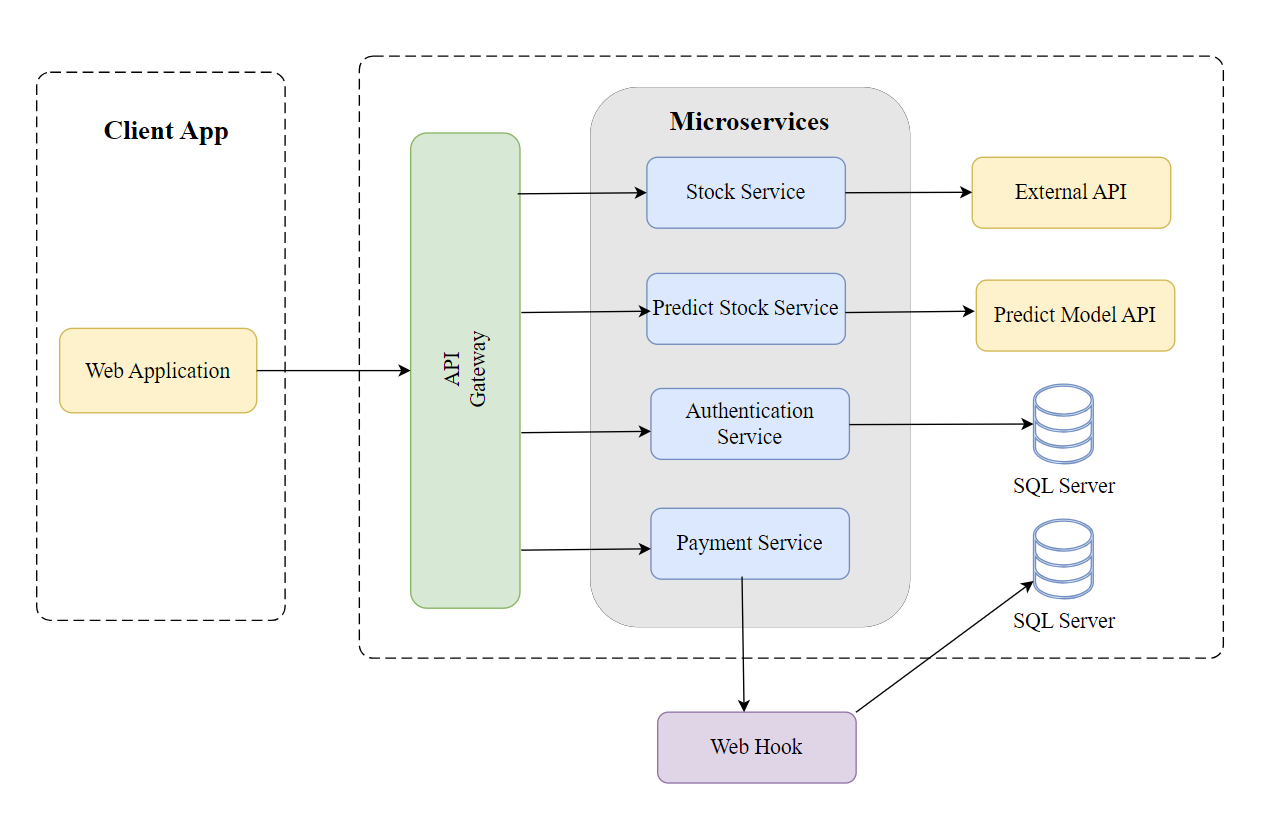
Hệ thống web-based giúp giảm chi phí liên quan đến cài đặt và bảo trì phần mềm trên máy khách (client-side), nhờ đó tiết kiệm chi phí hạ tầng.

*3.2.5. Tích hợp dễ dàng*

Hệ thống có khả năng tích hợp với các dịch vụ bên ngoài và hệ thống khác thông qua **API**. Điều này đặc biệt hữu ích cho việc đồng bộ hóa dữ liệu với các dịch vụ khác như thanh toán trực tuyến, quản lý tài khoản, và các dịch vụ phân tích tài chính.

## **3.3. Thiết kế kiến trúc phần mềm**

Kiến trúc của hệ thống dự đoán giá cổ phiếu được thiết kế dựa trên nguyên tắc Microservices, cho phép các dịch vụ hoạt động độc lập, dễ dàng bảo trì và mở rộng khi cần thiết. Sơ đồ dưới đây minh họa cấu trúc tổng thể của hệ thống, bao gồm các thành phần chính:



*Hình 6. Kiến trúc phần mềm*

*3.2.1. Client Application*

Ứng dụng web đóng vai trò là giao diện người dùng (frontend), cho phép người dùng tương tác với hệ thống.

Tất cả yêu cầu từ ứng dụng web sẽ được gửi qua **API Gateway** để điều hướng đến các dịch vụ microservices tương ứng.

*3.2.2. API Gateway*

**API Gateway** hoạt động như một điểm truy cập duy nhất, nhận các yêu cầu từ phía người dùng và phân phối chúng đến các microservices tương ứng.

Việc sử dụng API Gateway giúp hệ thống dễ dàng quản lý luồng dữ liệu, đồng thời đảm bảo bảo mật và hiệu suất.

*3.2.3. Microservices*

Các microservices trong hệ thống bao gồm:

**Stock Service:** Dịch vụ này chịu trách nhiệm lấy dữ liệu giá cổ phiếu từ các **API bên ngoài (External API)** và cung cấp thông tin về giá cổ phiếu hiện tại cho người dùng.

**Predict Stock Service:** Dịch vụ này thực hiện việc dự đoán giá cổ phiếu trong tương lai bằng cách gọi đến **Predict Model API**. Dữ liệu từ các mô hình học máy (machine learning) được xử lý để cung cấp các dự đoán chính xác cho người dùng.

**Authentication Service:** Xử lý việc xác thực và quản lý người dùng, đảm bảo rằng chỉ những người dùng hợp lệ mới có thể truy cập vào các tính năng của hệ thống. Dữ liệu người dùng được lưu trữ và quản lý trong **SQL Server**.

**Payment Service:** Quản lý các giao dịch thanh toán, bao gồm việc xử lý và lưu trữ thông tin giao dịch trong **SQL Server**. Dịch vụ này cũng tích hợp với **Web Hook** để phản hồi các sự kiện thanh toán ngay lập tức khi có sự thay đổi.

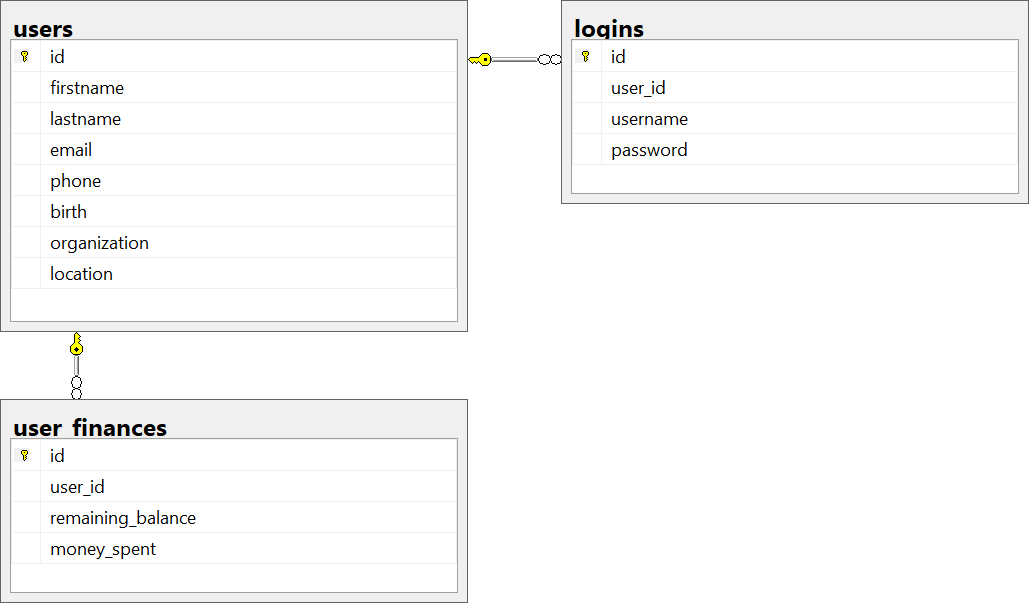
*3.2.4. Web Hook*

**Web Hook** được sử dụng để xử lý các sự kiện liên quan đến thanh toán hoặc thay đổi dữ liệu từ các dịch vụ bên ngoài một cách tự động và tức thì.

*3.2.5. Cơ sở Dữ liệu*

**SQL Server**: Tất cả dữ liệu liên quan đến người dùng và các giao dịch thanh toán đều được lưu trữ an toàn trong SQL Server, giúp hệ thống dễ dàng quản lý và truy xuất dữ liệu khi cần.

## 3.4. Thiết kế cơ sở dữ liệu



*Hình* *7. Database Diagram*

a, Bảng *Users*

Bảng Users lưu trữ thông tin về người dùng bao gồm: họ và tên, email, số điện thoại, ngày sinh, nới làm việc và địa chỉ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Column Name | Data Type | Allow Nulls | PK/FK | Unique |
| id | int |  | PK |  |
| firstname | nvarchar(100) |  |  | x |
| lastname | nvarchar(100) |  |  | x |
| email | varchar(100) |  |  |  |
| phone | varchar(50) |  |  | x |
| birth | Date |  |  | x |
| organazation | varchar(100) |  |  | x |
| location | varchar(100) |  |  | x |

b, Bảng *User Finances*

Bảng Finances lưu trữ thông tin giao dịch của người dùng.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Column Name | Data  Type | Allow Nulls | PK/FK | Unique |
| id | int |  | PK |  |
| user\_id | int |  | FK |  |
| remaining\_balance | int |  |  |  |
| money\_spent | int |  |  |  |

c, Bảng *Logins*

Bảng Login lưu trữ thông tin tài khoản của người dùng.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Column Name | Data Type | Allow  Nulls | PK/FK | Unique |
| id | int |  | PK |  |
| user\_id | int |  | FK |  |
| password | varchar(100) |  |  |  |
| username | varchar(100) |  |  |  |

# Chương 4. Xây dựng website

## Giới thiệu về công cụ sử dụng

*4.1.1.Cơ sở dữ liệu: Microsoft SQL Server*

SQL Server là một hệ quản trị cơ sở dữ liệu quan hệ (RDBMS) phát triển bởi Microsoft, được sử dụng để lưu trữ, quản lý và truy vấn dữ liệu.

*4.1.2. Framework Argular*

Angular là một framework mã nguồn mở do Google phát triển, được sử dụng để xây dựng các ứng dụng web đơn trang (single-page applications - SPA) hiện đại và mạnh mẽ.

*4.1.3. Ngôn ngữ Python*

Python là ngôn ngữ phổ biến trong phát triển AI nhờ cú pháp dễ đọc, hệ sinh thái phong phú với các thư viện như TensorFlow, PyTorch, Scikit-learn, và cộng đồng hỗ trợ rộng lớn. Điều này làm Python trở thành lựa chọn hàng đầu cho các dự án AI và học máy.

*4.1.4. Môi trường phát triển: Visual Studio*

Visual Studio là một môi trường phát triển tích hợp (IDE) mạnh mẽ của Microsoft, được thiết kế để hỗ trợ phát triển ứng dụng cho nhiều nền tảng, bao gồm Windows, web, và di động.

*4.1.5. Môi trường phát triển: Visual Studio Code*

Visual Studio Code (VS Code**)** là một trình soạn thảo mã nguồn nhẹ và mạnh mẽ do Microsoft phát triển, phù hợp cho nhiều loại dự án lập trình.

*4.1.6. Google Colab*

Google Colab là công cụ lý tưởng cho việc phát triển và thử nghiệm các mô hình học máy và phân tích dữ liệu mà không cần thiết lập môi trường máy tính cục bộ phức tạp.

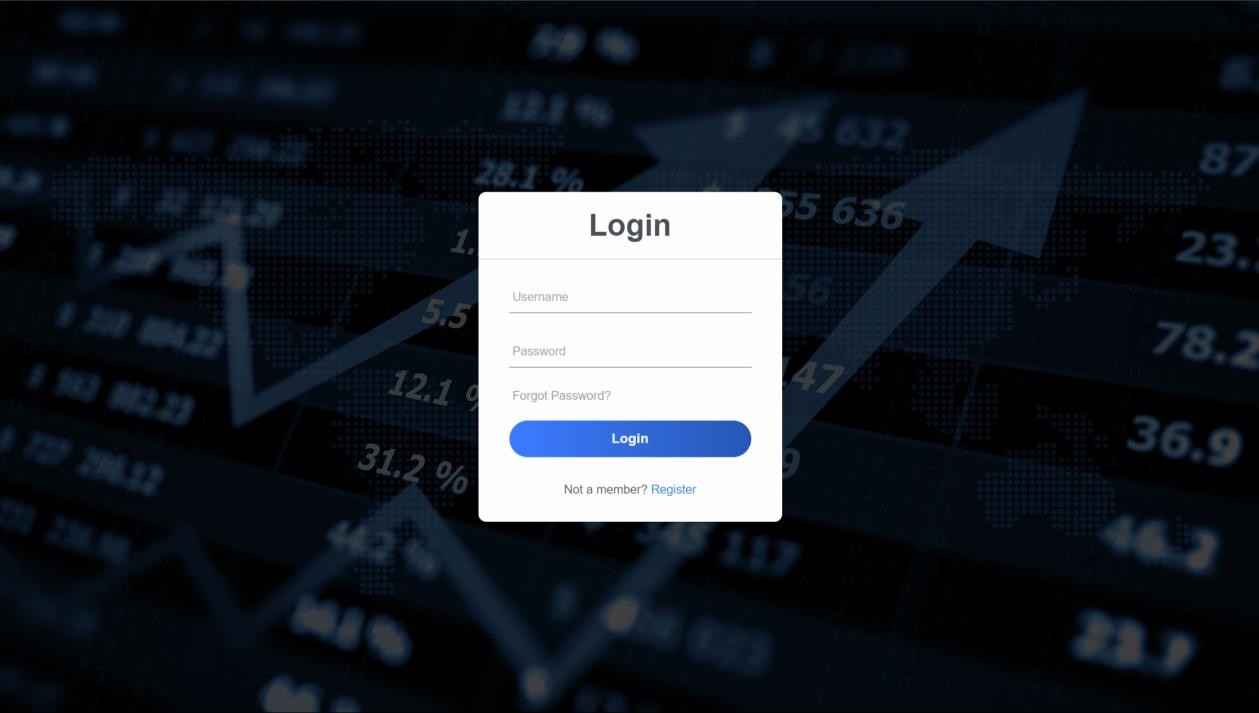
*4.1.7. VietQR API*

VietQR API là một dịch vụ giúp bạn dễ dàng tạo mã QR cho ứng dụng và website của mình. API hỗ trợ chuyển đổi các loại dữ liệu như URL, văn bản, số điện thoại, và thông tin thanh toán thành mã QR.

*4.1.8. Casso API*

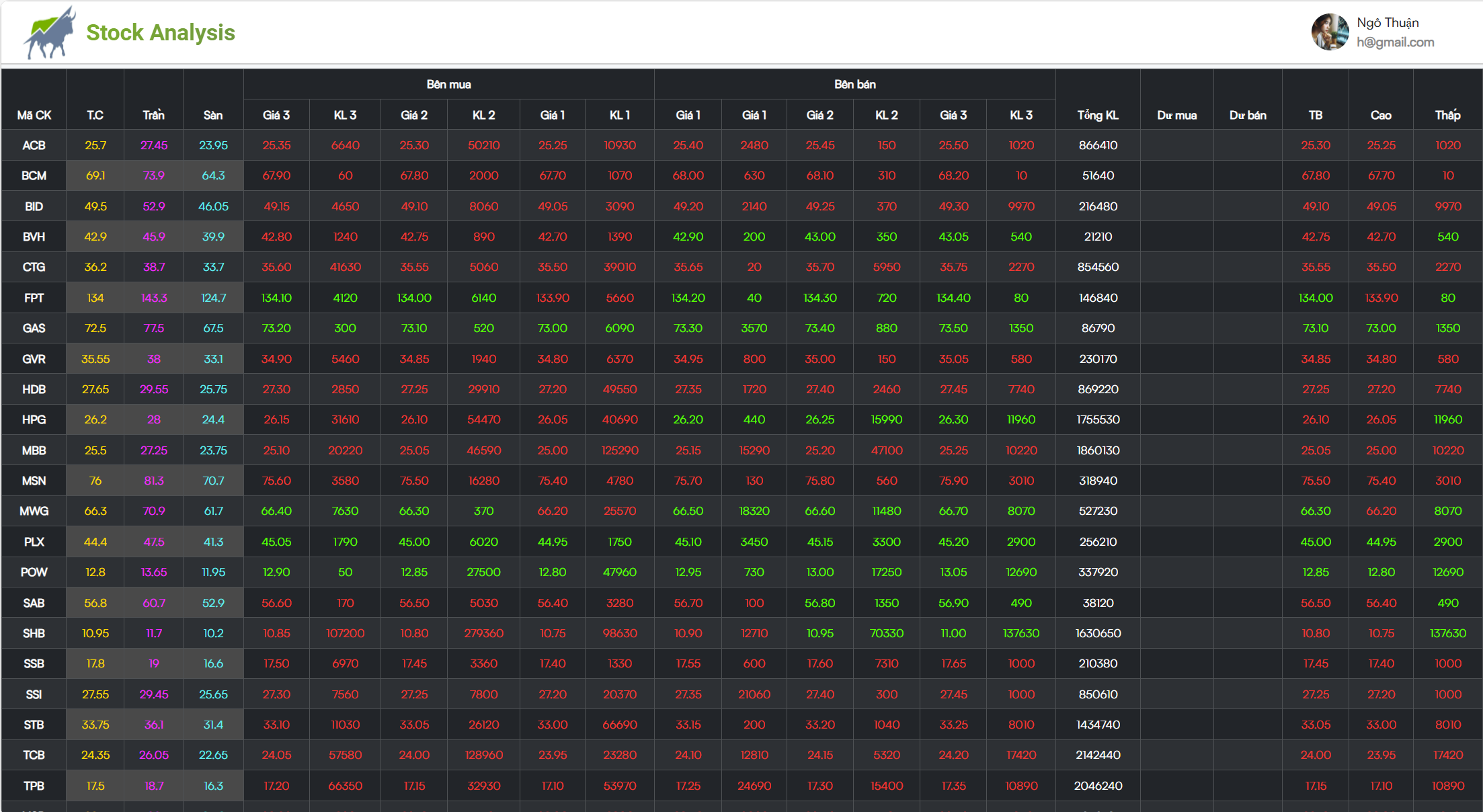
Casso.vn là nền tảng quản lý tài chính và thanh toán trực tuyến tại Việt Nam, hỗ trợ cá nhân và doanh nghiệp theo dõi chi tiêu, tạo báo cáo tài chính và thực hiện thanh toán dễ dàng.

## Xây dựng các chức năng chính

* + 1. *Chức năng đăng nhập*

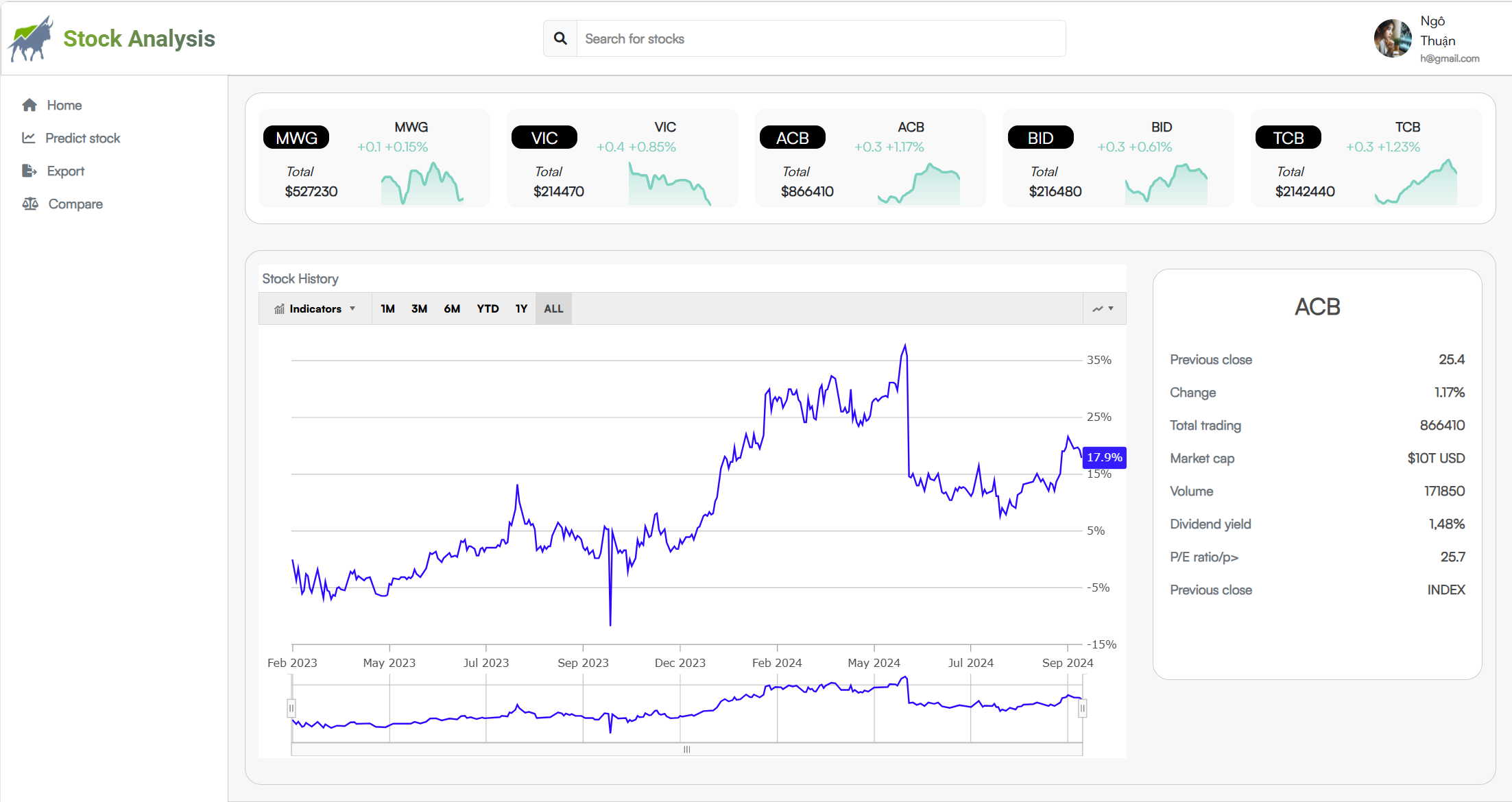
*Hình* *8. Màn hình đăng nhập*

* + 1. *Xem thông tin danh sách các mã cổ phiếu*



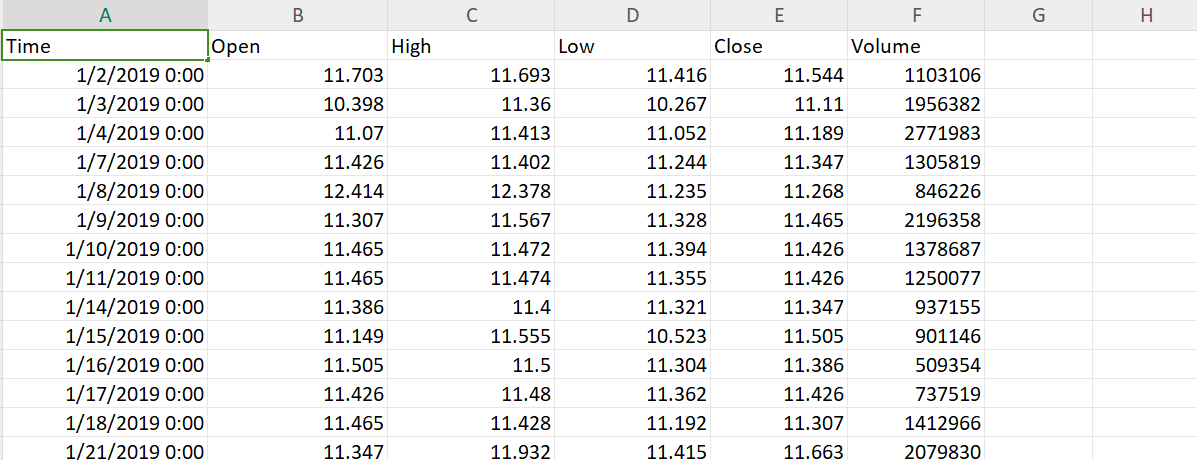
*Hình* *9. Màn hình hiển thị tổng quan các mã cổ phiếu*

* + 1. *Xem thông tin chi tiết từng mã cổ phiếu*



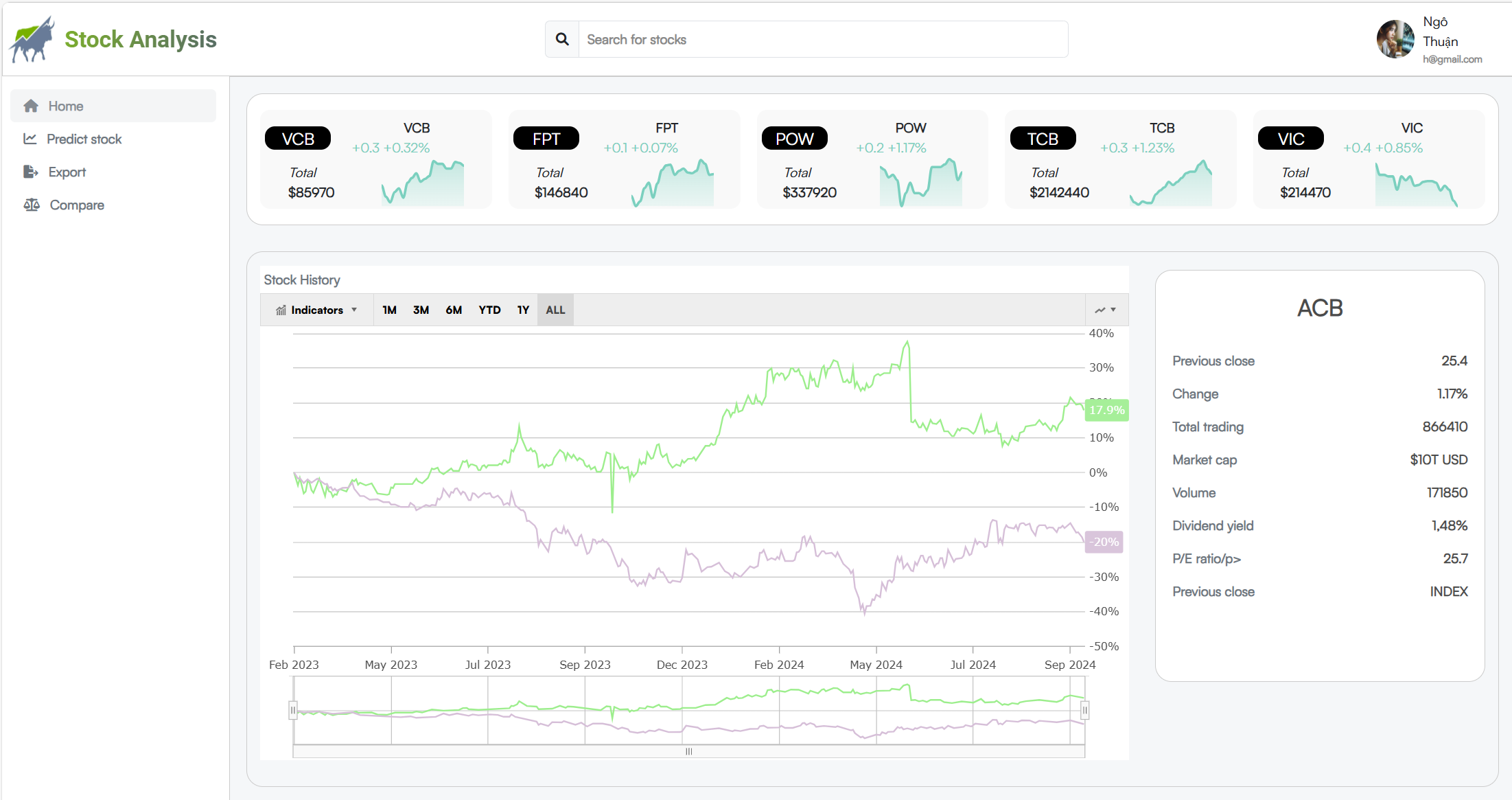
*Hình 10. Hiển thị thông tin chi tiết từng mã cổ phiếu*

* + 1. *Thực hiện chức năng xuất file excel*



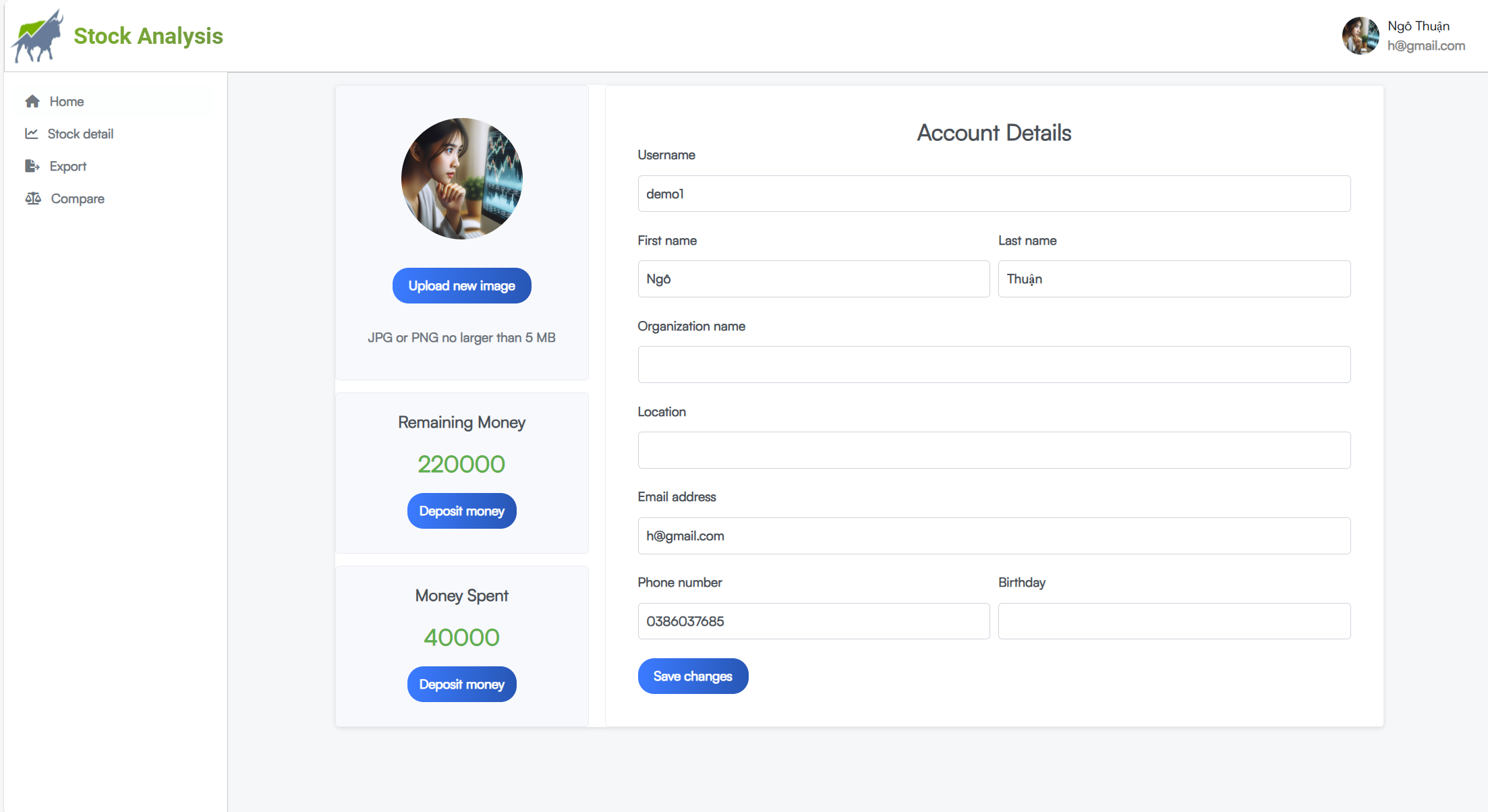
*Hình 11. File excel sau khi xuất*

* + 1. *So sánh hai mã cổ phiếu*



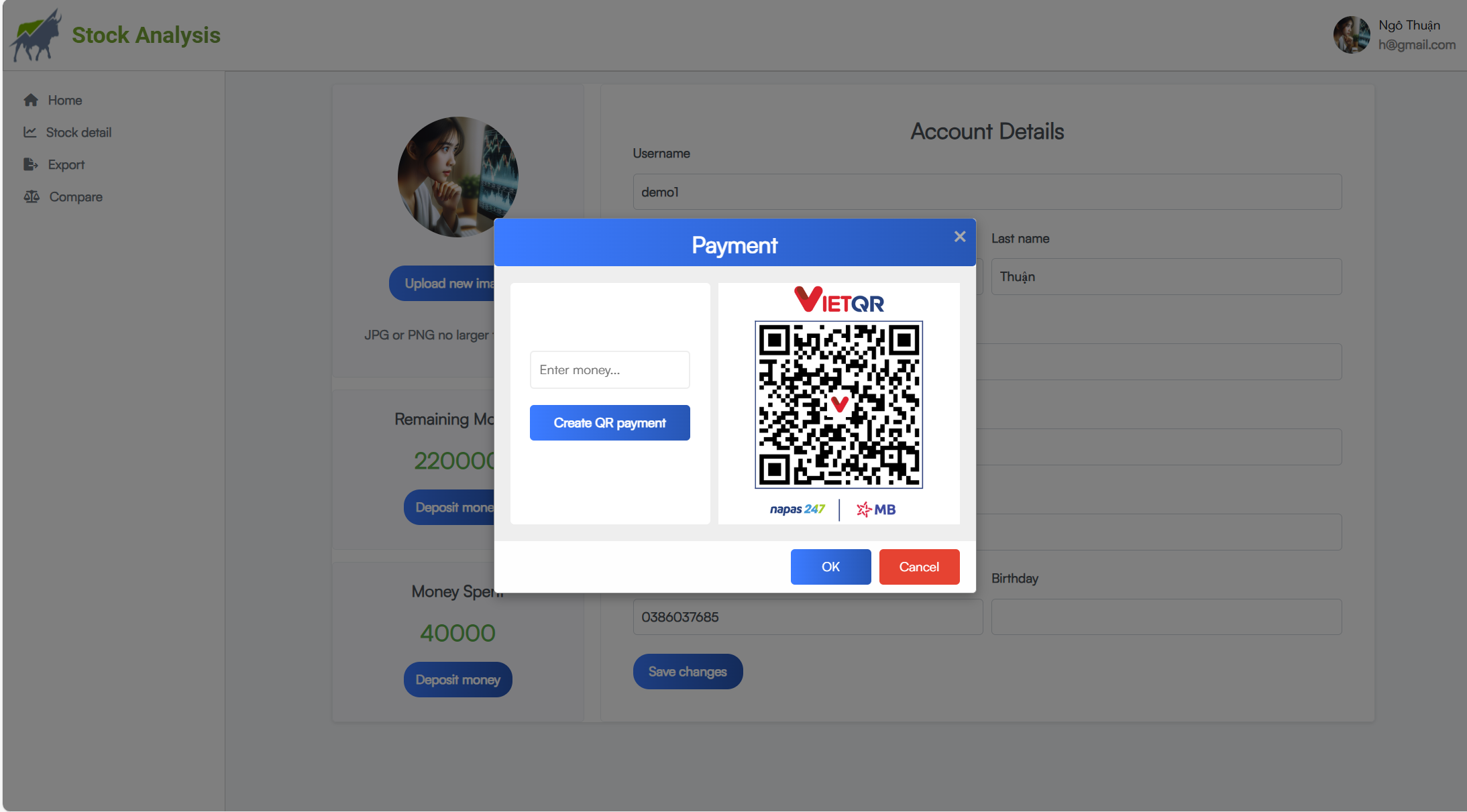
*Hình* *12. So sánh hai mã cổ phiếu*

* + 1. *Hiển thị thông tin cá nhân người dùng*



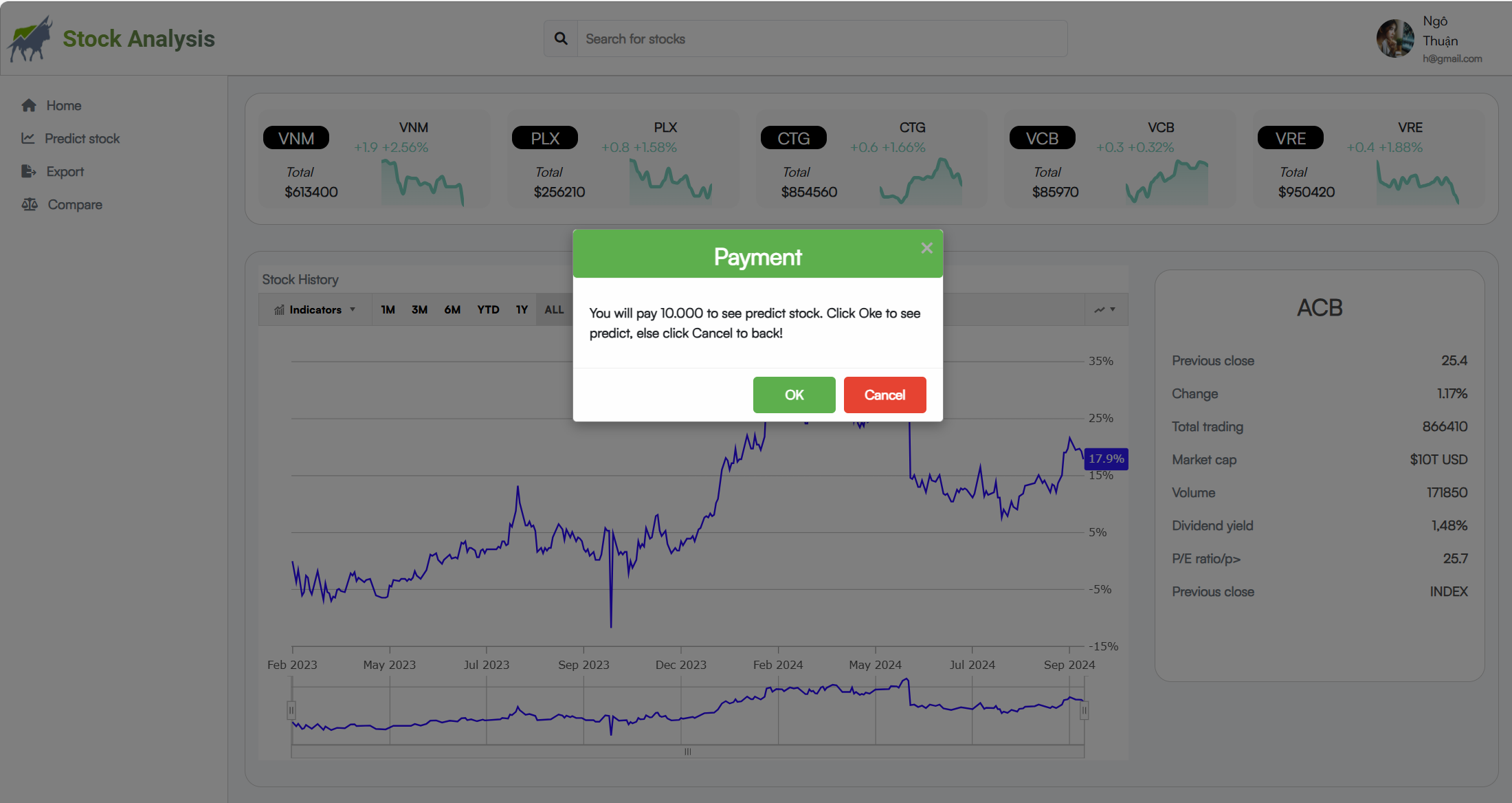
*Hình* *13. Màn hình hiển thị thông tin người dùng*

* + 1. *Nạp tiền vào tài khoản*

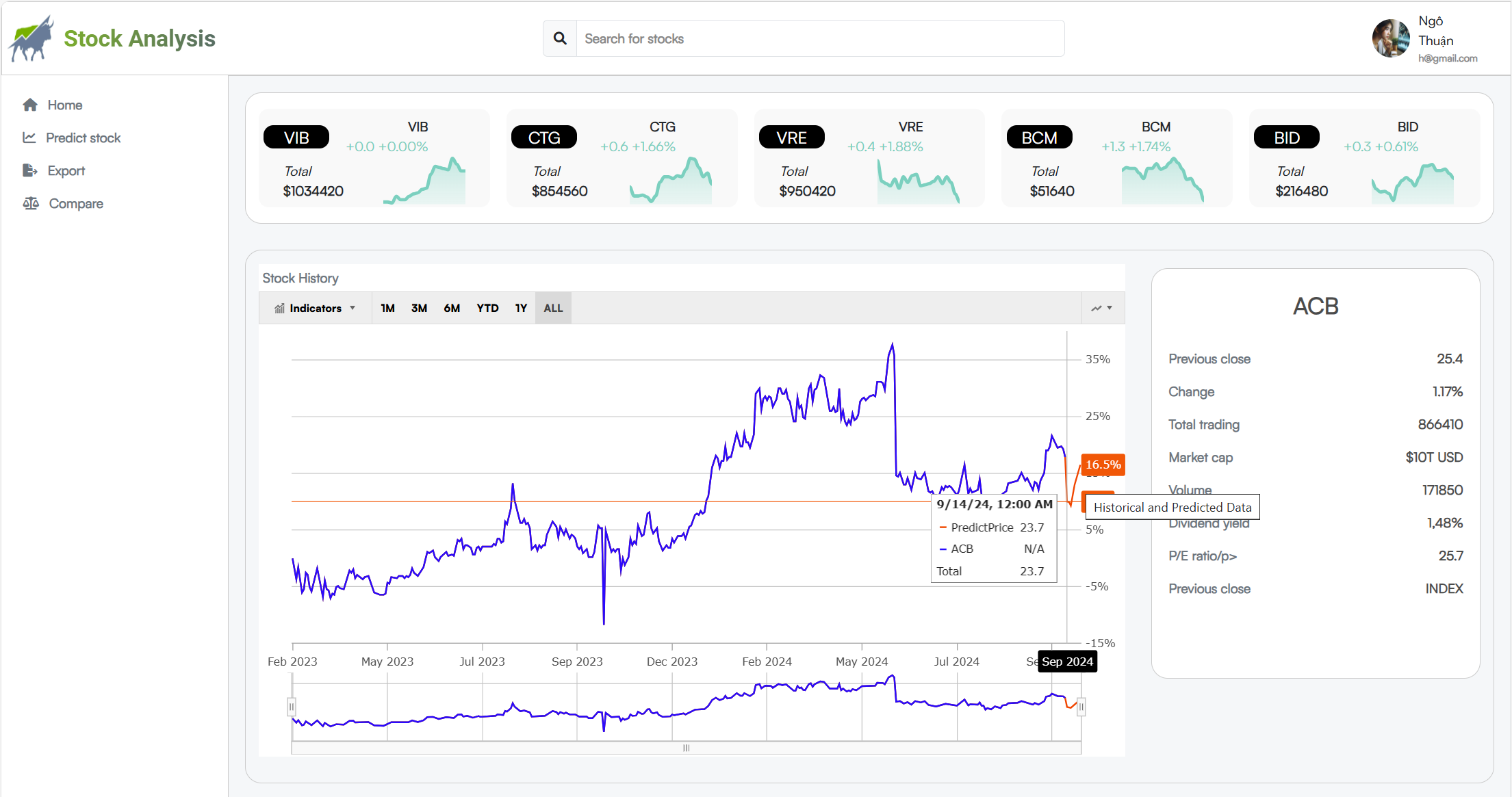


*Hình* *14. Nạp tiền vào tài khoản*

* + 1. *Dự đoán giá cổ phiếu*

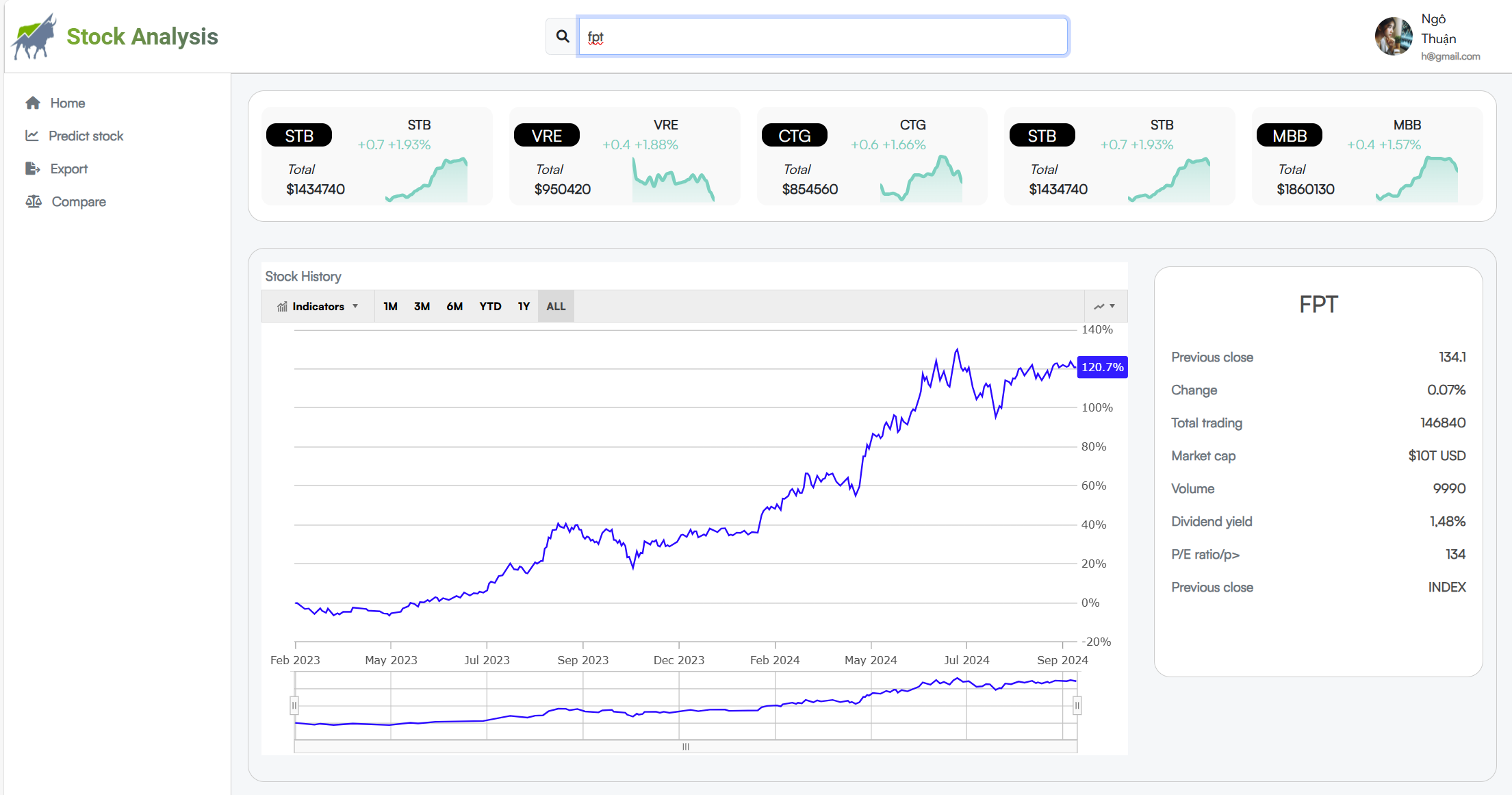


*Hìn**h 15. Màn hình thông báo khi bấm vào “Predict Stock”*



*Hình 16. Dự đoán giá cổ phiểu*

* + 1. *Tìm kiếm giá cổ phiếu*



*Hình 17: Tìm kiếm giá cổ phiếu*

# Chương 5: Kết quả đạt được

Hệ thống dự đoán giá cổ phiếu đã thành công trong việc kết hợp dữ liệu lịch sử và dự đoán bằng cách sử dụng thuật toán LSTM. Thuật toán này cho phép mô hình học từ các mẫu trong dữ liệu thời gian dài, giúp tối ưu hóa quá trình nhận diện xu hướng và dự đoán giá cổ phiếu tương lai một cách hiệu quả. Bằng cách tận dụng các đặc tính của LSTM, hệ thống có khả năng xử lý các chuỗi dữ liệu liên tục và phức tạp, giúp tăng độ chính xác trong dự đoán.

Hệ thống đã thực hiện phân tích chuỗi thời gian dựa trên các thông số quan trọng của thị trường, như giá mở cửa, giá đóng cửa, khối lượng giao dịch hàng ngày, và các chỉ số tài chính khác. Việc sử dụng LSTM giúp mô hình không chỉ nhận diện được mối quan hệ giữa các thông số này mà còn dự đoán được xu hướng tương lai, giúp người dùng nắm bắt thông tin một cách chính xác và có cơ sở trong việc ra quyết định đầu tư.

Thách thức liên quan đến dữ liệu nhiễu và biến động mạnh của thị trường đã được hệ thống giải quyết hiệu quả. Nhờ khả năng của LSTM trong việc ghi nhớ và quên có chọn lọc, mô hình có thể lọc ra các yếu tố nhiễu, tập trung vào các thông tin quan trọng để dự đoán chính xác. Điều này giúp hạn chế rủi ro do dữ liệu biến động mạnh gây ra, tăng cường sự ổn định của dự đoán.

Về mặt hiệu suất, hệ thống đã được tối ưu hóa nhằm đảm bảo khả năng xử lý nhanh chóng và đưa ra dự đoán kịp thời. Thuật toán LSTM được tinh chỉnh để xử lý lượng lớn dữ liệu một cách hiệu quả mà không làm giảm độ chính xác, từ đó đáp ứng yêu cầu khắt khe về thời gian và hiệu suất của người dùng.

Biểu đồ tài chính của hệ thống mang lại trải nghiệm trực quan, cho phép người dùng dễ dàng so sánh giữa dữ liệu lịch sử và dự đoán. Nhờ sự phân tích kỹ lưỡng của LSTM và khả năng trực quan hóa dữ liệu, người dùng có thể theo dõi sự thay đổi của giá cổ phiếu theo thời gian, hỗ trợ trong việc phân tích và đưa ra các quyết định đầu tư hợp lý.

# Chương 6: Kết luận và hướng phát triển

Dự án đã đạt được những bước tiến đáng kể trong việc xây dựng hệ thống dự đoán giá cổ phiếu dựa trên thuật toán LSTM. Hệ thống hiện tập trung vào việc dự đoán giá cổ phiếu từ dữ liệu lịch sử và các thông số thị trường, mang lại khả năng phân tích chính xác và hiệu quả. Nhờ vào LSTM, mô hình có thể xử lý chuỗi thời gian phức tạp, giảm thiểu rủi ro từ nhiễu và biến động của thị trường.

Hướng phát triển tương lai của dự án là mở rộng khả năng dự báo sang các yếu tố kinh tế xã hội, chẳng hạn như giá cà phê và các chỉ số thị trường khác. Điều này sẽ giúp hệ thống trở nên đa dạng hơn, không chỉ giới hạn trong việc dự đoán giá cổ phiếu mà còn mở rộng sang các lĩnh vực kinh tế rộng lớn hơn. Khả năng này sẽ tạo thêm giá trị cho hệ thống, giúp người dùng theo dõi nhiều lĩnh vực khác nhau.

Bên cạnh đó, việc áp dụng các mô hình tiên tiến như SVM và RNN sẽ giúp tăng cường khả năng phân tích và dự đoán trong tương lai. Khi kết hợp với LSTM, những thuật toán này sẽ tối ưu hóa hiệu suất của hệ thống. Điều này đáp ứng yêu cầu cao về tính kịp thời và độ chính xác của dự đoán, giúp hệ thống thích nghi với nhiều biến động của thị trường.

Trong thời gian tới, dự án sẽ tập trung vào hoàn thiện ứng dụng di động và xây dựng module web quản lý các công ty niêm yết chứng khoán. Những cải tiến này sẽ mang lại trải nghiệm tốt hơn cho người dùng, giúp họ dễ dàng theo dõi và phân tích dữ liệu doanh nghiệp. Hướng phát triển này không chỉ giúp quản lý thông tin một cách dễ dàng hơn mà còn cung cấp dự báo chính xác, hỗ trợ ra quyết định đầu tư thông minh.

# Chương 7: Tài liệu tham khảo

## 7.1. Argular Stock Chart

[*Argular Stock Chart*](https://www.infragistics.com/products/ignite-ui-angular/angular/components/charts/types/stock-chart)

## 7.2. Stock Analysis - Github

[*Stock Anslysis*](https://github.com/chungcacbon/PRN231_StockAnalysis)

## 7.3. Payment Bank

[*Tạo QR code VietQR Payment Bank API*](https://www.youtube.com/watch?v=J2wvv5bn6E8&t=899s)

## 7.4. Root Mean Square Error

[*Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, L., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Mo*](http://doi.org/10.48550/arXiv.1603.04467%20)

## 7.4. A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction.

[*Alpaydin, E. (2014). Introduction to Machine Learning. MIT Press. Chen, Y., & Hao, Y. (2017). A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. Expert Systems with Applications, 80, 340-355.*](http://doi.org/10.1016/j.eswa.2017.02.044%20)

## 7.4. Learning phrase representations using RNN encoder-decoder for statistical machine translation.

[*Cho, K., Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation.*](%20http:/doi.org/10.3115/v1/D14-1179%20)

## 7.5. Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction.

[*Chung, H., & Shin, K. S. (2020). Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction. Neural Computing and Applications, 32.*](%20http:/doi.org/10.1007/s00521-019-04236-3%20)

## 7.6. Brent crude oil price forecast utilizing deep neural network architectures.

[*Daneshvar, A., Ebrahimi, M., Salahi, F., & Rahmaty, M. (2022). Brent crude oil price forecast utilizing deep neural network architectures. Computational Intelligence and Neuroscience.*](%20http:/doi.org/1-13.10.1155/2022/6140796%20)

## 7.7. Taking the Human Out of the Loop: A Review of Bayesian Optimization.

[*de Freitas, N., Shahriari, B., Swersky, K., Wang, Z., & Adams, R. P. (2016). Taking the Human Out of the Loop: A Review of Bayesian Optimization. Proceedings of the IEEE, 104(1), 148–175.*](https://doi.org/10.1109/jproc.2015.2494218%20)

## 7.8. Adaptive subgradient methods for online learning and stochastic optimization.

[Duchi, J., Hazan, E., & Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. Journal of machine learning research, 12, 2121-2159. Guo, C., Liu, G., & Chen, C. H. (2020). Air pollution concentration forecast metho](%20https:/doi.org/10.1155/2020/8854649)

## 7.9. Long Short-Term Memory

[*Hochreiter, S., & Schmidhuber, J. (1997), Long Short-Term Memory. Neural Computation, 9, 1735–1780.*](http://doi.org/10.1162/neco.1997.9.8.1735%20)

## 7.10. Application of Seasonal SVR with Chaotic Immune Algorithm in Traffic Flow Forecasting

[*Hong, W. C. (2021). Application of Seasonal SVR with Chaotic Immune Algorithm in Traffic Flow Forecasting. Neural Computing and Applications, 21, 583–593.*](http://doi.org/10.1007/s00521-010-0456-7%20)

## 7.11. Hybrid Deep Learning Model for Stock Price Prediction

[*Hossain, M., Karim, R., Thulasiram, R., Bruce, N. D. B., & Wang, Y. (2018). Hybrid Deep Learning Model for Stock Price Prediction, 1837-1844.*](http://doi.org/10.1109/SSCI.2018.8628641%20)

## 7.12. Deep Learning with Long Short-Term Memory for Time Series Prediction.

[*Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z., & Zhang, H. (2019), Deep Learning with Long Short-Term Memory for Time Series Prediction. IEEE Communications Magazine, 1-6.*](%20https:/doi.org/10.1109/MCOM.2019.1800155%20)

## 7.13. Matplotlib: a 2D graphics environment

[*Hunter, J. D. (2007). Matplotlib: a 2D graphics environment. Computing in Science & Engineering, 9, 90–95.*](https://doi.org/10.1109/MCSE.2007.55%20)

## 7.14. Predicting Macro-Financial Instability - How Relevant is Sentiment? Evidence from Long Short-Term Memory Networks.

[*Kanzari, D., Nakhli, M. S., Gaies, B., & Sahut, J. M. (2023). Predicting Macro-Financial Instability - How Relevant is Sentiment? Evidence from Long Short-Term Memory Networks, 65*.](http://doi.org/10.1016/j.ribaf.2023.101912%20)

## 7.15. ADAM: A Method for Stochastic Optimization. International Conference on Learning Representations.

[*Kingma, D. P., & Ba, J. (2014). ADAM: A Method for Stochastic Optimization. International Conference on Learning Representations.*](http://doi.org/10.48550/arXiv.1412.6980%20)

## 7.16. Deep Learning. Nature

[*LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521, 436-44.*](http://doi.org/%2010.1038/nature14539%20)

## 7.17. Intelligent stock trading system based on improved technical analysis and Echo State Network.

[*Lin, X., Yang, Z., & Song, Y. (2011). Intelligent stock trading system based on improved technical analysis and Echo State Network. Expert systems with Applications, 38(9), 11347-11354.*](http://doi.org/%2010.1016/j.eswa.2011.03.001%20)

## 7.18. Application of Regularized GRU-LSTM Model in Stock Price Prediction.

[*Liu, Y., Wang, Z., & Zheng, B. (2019). Application of Regularized GRU-LSTM Model in Stock Price Prediction. 1886-1890.*](http://doi.org/10.1109/ICCC47050.2019.9064035%20)

## 7.19. A systematic review of fundamental and technical analysis of stock market predictions.

[*Nti, I. K., Adekoya, A. F., & Weyori, B. A. (2020). A systematic review of fundamental and technical analysis of stock market predictions. Artificial Intelligence Review, 53(4), 3007-3057.*](http://doi.org/10.1007/s10462-019-09754-z%20)