AI Course

**Action Plan &**

**Capstone Project   
Final Report**

For students (instructor review required)

ⓒ2023 SAMSUNG. All rights reserved.

Samsung Electronics Corporate Citizenship Office holds the copyright of this document.

This document is a literary property protected by copyright law so reprint and reproduction without permission are prohibited.

To use this document other than the curriculum of Samsung Innovation Campus, you must receive written consent from copyright holder.

|  |
| --- |
| ỨNG DỤNG THUẬT TOÁN HỌC MÁY DỰ ĐOÁN XU HƯỚNG GIÁ CỔ PHIẾU TRÊN THỊ TRƯỜNG CHỨNG KHOÁN VIỆT NAM |

22/08/2024

GROUP 9

Nguyễn Tuấn Đạt

Vũ Minh Đức

Đặng Văn Mạnh

**MỤC LỤC**

[ỨNG DỤNG THUẬT TOÁN HỌC MÁY DỰ ĐOÁN XU HƯỚNG GIÁ CỔ PHIẾU TRÊN THỊ TRƯỜNG CHỨNG KHOÁN VIỆT NAM 2](#_Toc175281644)

[DANH MỤC CÁC TỪ VIẾT TẮT 6](#_Toc175281645)

[1. Introduction 7](#_Toc175281646)

[1.1. Thông tin chung 7](#_Toc175281647)

[1.1.1. Khảo sát bài toán 7](#_Toc175281648)

[1.1.2. Phát biểu bài toán phát hiện ung thư phổi 8](#_Toc175281649)

[1.1.3. Giới thiệu về LightGBM và RNN trong dự đoán giá cổ phiếu 9](#_Toc175281650)

[1.2. Mục tiêu và động lực 9](#_Toc175281651)

[1.2.1. Động lực 9](#_Toc175281652)

[1.2.2. Mục tiêu 9](#_Toc175281653)

[1.3. Thành viên và vai trò phân công 10](#_Toc175281654)

[1. Nguyễn Tuấn Đạt – Tìm hiểu lí thuyết về mô hình, các kĩ thuật 10](#_Toc175281655)

[2. Đặng Văn Mạnh – Nghiên cứu mô hình và các chỉ số đánh giá 10](#_Toc175281656)

[3. Vũ Minh Đức - Phát triển giao diện 10](#_Toc175281657)

[1.4. Lịch trình và các mốc quan trọng 10](#_Toc175281658)

[2. Project Execution 11](#_Toc175281659)

[2.1. Thu thập dữ liệu 11](#_Toc175281660)

[2.2. Phương pháp huấn luyện 12](#_Toc175281661)

[2.2.1. Xử lý và tăng cường dữ liệu 12](#_Toc175281662)

[2.2.2. Mô hình LightGBM 12](#_Toc175281663)

[2.2.3. Mô hình RNN 12](#_Toc175281664)

[2.2.4. Tiêu chí đánh giá 13](#_Toc175281665)

[2.3.Quy trình làm việc 14](#_Toc175281666)

[3. Results 15](#_Toc175281667)

[3.1. Tiền xử lý dữ liệu 15](#_Toc175281668)

[3.2. Phân tích dữ liệu và thăm dò 15](#_Toc175281669)

[3.3. Modeling 15](#_Toc175281670)

[3.4 Kết quả huấn luyện 16](#_Toc175281671)

[4. Projedct impact 17](#_Toc175281672)

[4.1. Thành tựu và lợi ích 17](#_Toc175281673)

[4.2. Hướng phát triển 18](#_Toc175281674)

[5. Team Member Review and Comment 18](#_Toc175281675)

**DANH SÁCH BẢNG**

[Bảng 3.4.1 Kết quả huấn luyện mô hình LightGBM và RNN 16](#_Toc175282409)

# DANH MỤC CÁC TỪ VIẾT TẮT

|  |  |  |
| --- | --- | --- |
| **STT** | **Ký hiệu** | **Nguyên nghĩa** |
| 1 | RNN | Recurrent Neural Network |
| 2 | LightGBM | Light Gradient Boosting Machine |
| 3 | MAE | **Mean Absolute Error** |
| 4 | RMSE | **Root Mean Squared Error** |

**DANH MỤC HÌNH ẢNH**

[Hình 3.1.1 Mã tiền xử lý dữ liệu mô hình LightGBM 15](#_Toc175282437)

[Hình 3.3.1 Huấn luyện mô hình LightGBM 16](#_Toc175282438)

[Hình 3.3.2 Huấn luyện mô hình RNN 16](#_Toc175282439)

# 1. Introduction

## 1.1. Thông tin chung

### 1.1.1. Khảo sát bài toán

Thị trường chứng khoán đóng vai trò quan trọng trong nền kinh tế toàn cầu, là nơi diễn ra các hoạt động mua bán cổ phiếu của các công ty niêm yết, từ đó tạo ra dòng tiền và giá trị cho các doanh nghiệp cũng như nhà đầu tư. Tại Việt Nam, thị trường chứng khoán đang phát triển mạnh mẽ, thu hút nhiều nhà đầu tư trong và ngoài nước. Tuy nhiên, với sự biến động liên tục và chịu tác động bởi nhiều yếu tố, việc dự đoán chính xác xu hướng giá cổ phiếu trở nên vô cùng quan trọng.

Trong bối cảnh này, việc sử dụng các phương pháp truyền thống như phân tích kỹ thuật và phân tích cơ bản đã dần bộc lộ những hạn chế trong khả năng dự đoán chính xác và ứng phó kịp thời với biến động thị trường. Điều này tạo ra nhu cầu cấp thiết cho việc áp dụng các công nghệ tiên tiến, đặc biệt là các thuật toán học máy, để cải thiện độ chính xác và hiệu quả trong dự đoán giá cổ phiếu.

Học máy đã trở thành một công cụ mạnh mẽ trong việc xử lý và phân tích dữ liệu lớn, đặc biệt là trong các lĩnh vực yêu cầu dự đoán và ra quyết định dựa trên dữ liệu. Trong lĩnh vực tài chính, các thuật toán học máy được áp dụng rộng rãi để dự đoán xu hướng giá cổ phiếu, từ đó hỗ trợ nhà đầu tư trong việc đưa ra các quyết định mua bán hiệu quả.

Mô hình LightGBM và RNN là hai trong số các công nghệ học máy hiện đại được sử dụng để dự đoán giá cổ phiếu. LightGBM là một thuật toán tăng cường cây quyết định (Gradient Boosting Decision Tree), nổi bật với khả năng xử lý dữ liệu lớn, tốc độ huấn luyện nhanh và hiệu suất cao. RNN (Recurrent Neural Network) là một loại mạng neuron hồi tiếp, phù hợp với các bài toán thời gian, giúp khai thác mối quan hệ tuần tự trong dữ liệu lịch sử giá cổ phiếu.

Việc áp dụng các mô hình học máy vào thị trường chứng khoán Việt Nam không phải là một nhiệm vụ đơn giản. Thị trường Việt Nam còn khá non trẻ so với các thị trường lớn như Mỹ hay châu Âu, đồng thời dữ liệu thị trường chưa được chuẩn hóa và đôi khi thiếu ổn định. Ngoài ra, thị trường chứng khoán Việt Nam cũng chịu ảnh hưởng mạnh mẽ từ các yếu tố chính sách và tâm lý, làm cho việc dự đoán trở nên phức tạp hơn.

Tuy nhiên, nếu có thể vượt qua các thách thức này, việc ứng dụng các mô hình như LightGBM và RNN sẽ mở ra nhiều cơ hội mới. Chúng không chỉ giúp tăng cường độ chính xác của các dự báo mà còn góp phần làm tăng tính minh bạch và hiệu quả của thị trường chứng khoán Việt Nam, từ đó thu hút thêm nhiều nhà đầu tư và đóng góp vào sự phát triển kinh tế bền vững.

### 1.1.2. Phát biểu bài toán phát hiện ung thư phổi

**Mục tiêu:** Phát triển một hệ thống dự đoán giá cổ phiếu dựa trên giá cổ phiếu của các ngày trước đó.

**Đầu vào:** Bộ dữ liệu lịch sử giá cổ phiếu, bao gồm giá mở cửa, giá đóng cửa, giá cao nhất, giá thấp nhất và khối lượng giao dịch trong một khoảng thời gian nhất định.

**Đầu ra:** Hệ thống sẽ cung cấp dự đoán về giá của cổ phiếu trong các ngày tương lai.

**Yêu cầu chính:**

- **Thu thập dữ liệu:**

* Sử dụng bộ dữ liệu được lấy từ thị trường chứng khoán Việt Nam. Đảm bảo tính chính xác của dữ liệu

- **Tiền xử lý dữ liệu:**

* Đảm bảo dữ liệu thu về đã được điều chỉnh các giá trị thiếu và phù hợp với mô hình học máy

- **Thiết kế mô hình:**

* Áp dụng thuật toán học máy như LightGBM và RNN để xây dựng mô hình dự đoán giá cổ phiếu

- **Đánh giá hiệu suất:** Đánh giá hiệu suất của mô hình dựa trên các chỉ số như MAE, RMSE. So sánh kết quả với các mô hình khác để đảm bảo độ tin cậy.

- **Ứng dụng thực tế:**

* Triển khai mô hình vào thực tế, hỗ trợ nhà đầu tư trong việc ra quyết định dựa trên các dự đoán từ mô hình.

**Kết quả mong đợi:** Mô hình sẽ cung cấp dự báo có độ chính xác cao, giúp tối ưu hóa chiến lược đầu tư và quản lý rủi ro hiệu quả.

### 1.1.3. Giới thiệu về LightGBM và RNN trong dự đoán giá cổ phiếu

LightGBM (Light Gradient Boosting Machine) là một trong những thuật toán học máy dựa trên kỹ thuật cây quyết định (Decision Tree). LightGBM được tối ưu hóa về tốc độ huấn luyện và hiệu quả trong việc xử lý các tập dữ liệu lớn, đặc biệt là các bài toán hồi quy và phân loại. Với khả năng xử lý nhanh và hiệu quả, LightGBM đã chứng tỏ là một mô hình dự đoán giá cổ phiếu mạnh mẽ.

Mạng nơ-ron hồi quy (Recurrent Neural Network - RNN) là một loại mạng nơ-ron đặc biệt, có khả năng học từ chuỗi thời gian (time series). RNN có khả năng ghi nhớ các trạng thái trước đó và sử dụng chúng để dự đoán trạng thái tiếp theo, điều này rất hữu ích trong việc dự đoán giá cổ phiếu vì nó giúp mô hình học được xu hướng từ dữ liệu lịch sử.

## 1.2. Mục tiêu và động lực

### 1.2.1. Động lực

Động lực chính của dự án này là sự ứng dụng mạnh mẽ của học máy trong tài chính, đặc biệt là trong việc dự đoán giá cổ phiếu. Thị trường chứng khoán Việt Nam đã chứng kiến sự tăng trưởng mạnh mẽ, cùng với đó là nhu cầu cần có các công cụ dự báo chính xác để giúp các nhà đầu tư đưa ra quyết định đúng đắn hơn. Việc áp dụng các mô hình học máy tiên tiến như LightGBM và RNN sẽ giúp tăng cường độ chính xác của các dự đoán và mang lại lợi ích lớn cho các nhà đầu tư.

### 1.2.2. Mục tiêu

Mục tiêu chính của nghiên cứu này là:

* Áp dụng và so sánh hai mô hình dự đoán LightGBM và RNN để dự đoán xu hướng giá cổ phiếu trên thị trường chứng khoán Việt Nam.
* Đánh giá hiệu quả của các mô hình dự đoán dựa trên các chỉ số như độ chính xác, sai số bình phương trung bình (MSE), và sai số tuyệt đối trung bình (MAE).
* Cung cấp các phân tích và khuyến nghị nhằm cải thiện hiệu quả của mô hình trong các nghiên cứu tương lai.

## 1.3. Thành viên và vai trò phân công

### 1. Nguyễn Tuấn Đạt – Tìm hiểu lí thuyết về mô hình, các kĩ thuật

* Tìm hiểu lý thuyết về mô hình LightGBM và RNN
* Tìm hiểu về các kỹ thuật và nguyên lý hoạt động của mô hình.

### 2. Đặng Văn Mạnh – Nghiên cứu mô hình và các chỉ số đánh giá

* Nghiên cứu các chỉ số đánh giá và triển khai huấn luyện mô hình trên tập dữ liệu
* Đảm bảo các mô hình được huấn luyện một cách chính xác và tối ưu.

### 3. Vũ Minh Đức – Thu thập và xử lý dữ liệu

* Thu thập dữ liệu lịch sử của các cổ phiếu và thực hiện tiền xử lý dữ liệu.
* Đảm bảo dữ liệu được chuẩn hóa và sẵn sàng cho các mô hình.

## 1.4. Lịch trình và các mốc quan trọng

**Giai đoạn 1: Lập kế hoạch cho đề tài và thu thập dữ liệu**

* Họp khởi động dự án: phân công vai trò và trách nhiệm cho các thành viên trong nhóm
* Xác định phạm vi và mục tiêu của dự án
* Xác định đề tài
* Xác định lý thuyết và thu thập dữ liệu

**Giai đoạn 2: Tiền xử lý dữ liệu và xây dựng mô hình**

* Kiểm tra dữ liệu: thu thập các bộ dữ liệu liên quan, kiểm tra và xác nhận dữ liệu sử dụng
* Tiền xử lý dữ liệu: thực hiện làm sạch dữ liệu, chuẩn hóa dữ liệu, phân tích và trực quan hóa dữ liệu sử dụng
* Xây dựng mô hình: thiết kế kiến trúc mô hình LightGBM, xây dựng mô hình RNN

**Giai đoạn 3: Huấn luyện, tinh chỉnh và kiểm thử mô hình**

* Huấn luyện mô hình: thực hiện huấn luyện mô hình ban đầu trên tập dữ liệu huấn luyện
* Đánh giá: thực hiện đánh giá hiệu xuất mô hình trên tập dữ liệu kiểm tra

# 2. Project Execution

## 2.1. Thu thập dữ liệu

**- Mô tả bộ dữ liệu:** Dự án này sử dụng bộ dữ liệu bao gồm 30 mã cổ phiếu chứa thông tin về giá cả (open, high, low, close) và khối lượng giao dịch chứng khoán từ ngày 01/01/2023 đến ngày 20/08/2024 được lấy từ package [vnquant](https://github.com/phamdinhkhanh/vnquant.git).

**- Thành phần bộ dữ liệu**

+ price\_new.csv: Giá cả và khối lượng giao dịch chứng khoán

+ price\_test.csv: Tập dữ liệu test dùng để đánh giá chất lượng mô hình

**- Các trường dữ liệu:**

* date: thời gian giao dịch
* open: giá mở cửa (giá thực hiện tại lần khớp lệnh đầu tiên trong ngày giao dịch chứng khoán)
* high: giá cao nhất trong ngày
* low: giá thấp nhất trong ngày
* close: giá đóng cửa (giá thị trường của các cổ phiếu vào thời điểm đóng cửa một phiên giao dịch) (target\_variable)
* volume: khối lượng cổ phiếu giao dịch trong ngày (target variable)
* symbol: mã cổ phiếu

## 2.2. Phương pháp huấn luyện

Trong dự án này, hai mô hình chính là **LightGBM** và **RNN** đã được sử dụng để dự đoán giá cổ phiếu.

### 2.2.1. Xử lý và tăng cường dữ liệu

- Dữ liệu được chia thành nhiều khoảng thời gian 14 ngày để tạo ra các đặc trưng cho mô hình. Các đặc trưng bao gồm giá trị trung bình, độ lệch chuẩn, giá trị tối đa và tối thiểu của các khoảng thời gian này. Đồng thời, lợi nhuận logarit của cổ phiếu cũng được tính toán để bổ sung thêm thông tin.

### 2.2.2. Mô hình LightGBM

Mô hình **LightGBM (Light Gradient Boosting Machine)** là một phương pháp học tăng cường dựa trên cây quyết định (decision trees). Điểm mạnh của LightGBM là khả năng xử lý nhanh và hiệu quả các tập dữ liệu lớn với nhiều đặc trưng. LightGBM hoạt động bằng cách xây dựng các cây quyết định nhỏ và dần dần cải thiện chúng qua các vòng lặp, với mục tiêu tối ưu hóa độ chính xác của dự đoán.

Trong dự án này, mô hình LightGBM được áp dụng vào dữ liệu chuỗi thời gian với các đặc trưng được tạo ra từ giá cổ phiếu trong quá khứ. Chúng tôi đã tinh chỉnh các tham số của mô hình, chẳng hạn như số lượng cây (n\_estimators), tốc độ học (learning\_rate) để đạt được hiệu quả tốt nhất.

### 2.2.3. Mô hình RNN

**RNN (Recurrent Neural Network)** là một loại mô hình mạng nơ-ron có khả năng xử lý dữ liệu tuần tự, đặc biệt là các chuỗi thời gian như dữ liệu giá cổ phiếu. Mô hình RNN có cấu trúc cho phép ghi nhớ các trạng thái trước đó và sử dụng chúng để dự đoán các trạng thái tiếp theo. Điều này đặc biệt hữu ích khi xử lý dữ liệu tài chính vì nó có tính phụ thuộc tuần tự rất cao.

Trong dự án này, mô hình RNN được thiết kế với nhiều lớp mạng nơ-ron tuần tự, sử dụng các kỹ thuật như dropout để giảm thiểu hiện tượng overfitting. RNN giúp mô hình nắm bắt các xu hướng dài hạn trong giá cổ phiếu và cải thiện khả năng dự đoán so với các mô hình truyền thống khác.

### 2.2.4. Tiêu chí đánh giá

Hai tiêu chí chính để đánh giá hiệu suất của mô hình là **Mean Absolute Error (MAE)** và **Root Mean Squared Error (RMSE)**.

#### 2.2.4.1. Chỉ số MAE (****Mean Absolute Error)****

MAE là chỉ số đo lường trung bình trị tuyệt đối sai số giữa giá trị dự báo và giá trị thực tế. Công thức tính MAE:

Trong đó:

* n: là số lượng mẫu thực nghiệm
* : là giá trị của biến mục tiêu cho mẫu thứ i
* : là giá trị dự đoán của biến mục tiêu cho mẫu thứ i

Chúng ta có thể thấy về bản chất thì MAE chính là norm chuẩn bậc 1. Khi MAE càng nhỏ thì khoảng cách giữa giá trị dự báo và giá trị thực tế càng nhỏ và mô hình càng tốt.

#### 2.2.4.2. RMSE (****Root Mean Squared Error)****

RMSE là một độ đo thước đo sai số giữa giá trị dự đoán và giá trị thực tế trong bài toán dự đoán. RMSE tính căn bậc hai trung bình bình phương của hiệu giữa giá trị dự đoán và giá trị thực tế các mẫu dữ liệu, RMSE có ưu điểm là nó giúp đánh giá được độ lớn của sai số so với giá trị dự đoán và giá trị thực tế. RMSE càng nhỏ thì mô hình dự đoán càng chính xác và ngược lại.

RMSE được tính theo công thức sau:

Trong đó:

- n: là số lượng mẫu thực nghiệm

* : là giá trị của biến mục tiêu cho mẫu thứ i
* : là giá trị dự đoán của biến mục tiêu cho mẫu thứ i

## 2.3.Quy trình làm việc

**- Khởi động dự án**

* Họp khởi động dự án
* Xác định phạm vi và mục đích nghiên cứu của dự án
* Xác định đề tài và bài toán thực hiện
* Tìm hiểu lý thuyết

**- Thu thập và tiền xử lý dữ liệu**

* Thu thập các bộ dữ liệu liên quan
* Tiền xử lý dữ liệu

**- Xây dựng mô hình**

* Thiết kế kiến trúc mô hình
* Xây dựng mô hình

**- Huấn luyện và tinh chỉnh mô hình**

* Thực hiện huấn luyện mô hình
* Điều chỉnh tham số và tinh chỉnh mô hình
* Lưu kết quả và thực hiện tải lại mô hình cho dự đoán

**- Đánh giá mô hình và dự đoán kết quả:**

* Xác định các biểu đồ đánh giá độ chính xác của mô hình
* Dự đoán kết quả trên các dữ liệu kiểm thử

**- Hoàn thiện và báo cáo:**

* Chuẩn bị báo cáo cuối cùng
* Đánh giá và điều chỉnh dự án lần cuối
* Nộp và báo cáo dự án

# 3. Results

## 3.1. **Tiền xử lý dữ liệu**

**Tiền xử lý dữ liệu:** Bước tiền xử lý dữ liệu tập trung vào việc làm sạch dữ liệu ban đầu, biến đổi chúng thành các đặc trưng có ý nghĩa cho mô hình học máy. Sau khi dữ liệu được tải về cần chỉnh lại cho đúng định dạng phù hợp với mô hình và dữ liệu bị thiếu được điền bằng giá trị trước đó:

A screenshot of a computer program

Description automatically generated

Hình 3.1.1 Mã tiền xử lý dữ liệu mô hình LightGBM

## 3.2. Phân tích dữ liệu và thăm dò

Dự án đã thực hiện cài đặt thử nghiệm cho kỹ thuật đề xuất, sử dụng tập dữ liệu price\_new.csv gồm 12181 dòng gồm dữ liệu về 30 mã cổ phiếu trong các ngày giao dịch từ 01/01/2023 đến 20/08/2024 với 7 trường dữ liệu đã được xử lý cẩn thận. Dữ liệu được chia thành các tập huấn luyện, kiểm tra và đánh giá với sự phân chia chi tiết như sau: tập huấn luyện gồm 9136 dòng, tập kiểm tra gồm 3045 dòng.

Việc phân chia này đảm bảo rằng mỗi tập dữ liệu đều có đủ số lượng mẫu để thực hiện các thử nghiệm một cách chính xác và hiệu quả.

## 3.3. Modeling

Mô hình LightGBM được chúng em thực thi với n-estimators=500.

A screenshot of a computer program

Description automatically generated

Hình 3.3.1 Huấn luyện mô hình LightGBM

Mô hình RNN được chúng em thực thi vói epoch = 100.

A screen shot of a computer

Description automatically generated

Hình 3.3.2 Huấn luyện mô hình RNN

Mọi quá trình được đào tạo bằng Pytorch 3, Colab với 100 compute units, T4 GPU, RAM hệ thống 12.7GB

## 3.4 Kết quả huấn luyện

|  |  |  |
| --- | --- | --- |
|  | LightGBM | RNN |
| MAE | 0.9096 | 0.7376 |
| RMSE | 3.0403 | 1.0376 |

Bảng 3.4.1 Kết quả huấn luyện mô hình LightGBM và RNN

Kết quả huấn luyện cho thấy mô hình RNN có hiệu suất vượt trội hơn so với LightGBM trong việc dự đoán xu hướng giá cổ phiếu. Cụ thể, RNN đạt MAE là 0.7376, thấp hơn đáng kể so với 0.9096 của LightGBM, cho thấy mô hình này dự đoán các giá trị gần đúng với giá trị thực tế chính xác hơn. Đặc biệt, RMSE của RNN là 1.0376, thấp hơn rất nhiều so với 3.0403 của LightGBM, điều này chỉ ra rằng RNN không chỉ có sai số nhỏ hơn mà còn ít gặp phải các sai số lớn trong dự đoán, cho thấy tính ổn định cao hơn trong việc xử lý dữ liệu. RNN thường được ưu tiên trong các bài toán liên quan đến chuỗi thời gian, nhờ vào khả năng nắm bắt các mối quan hệ tuần tự trong dữ liệu, điều này lý giải vì sao nó có kết quả tốt hơn trong bối cảnh dự đoán giá cổ phiếu. Mặc dù LightGBM là một mô hình mạnh mẽ và có thể hữu ích trong các tình huống khác, đặc biệt là khi xử lý dữ liệu không phụ thuộc hoàn toàn vào chuỗi thời gian hoặc khi yêu cầu tốc độ huấn luyện nhanh, nhưng trong kịch bản cụ thể này, RNN tỏ ra là lựa chọn ưu việt hơn để đạt được độ chính xác cao nhất. Kết quả này nhấn mạnh vai trò quan trọng của việc chọn mô hình phù hợp với đặc thù dữ liệu, giúp tối ưu hóa kết quả dự đoán trong các dự án liên quan đến phân tích tài chính và thị trường chứng khoán.

# 4. Projedct impact

## 4.1. Thành tựu và lợi ích

Do hạn chế về thời gian nên dự án vẫn còn thiếu sót nhưng đã đạt dược một số kết quả nhất định. Trước hết, việc ứng dụng các mô hình học máy như LightGBM và RNN vào việc dự đoán xu hướng giá cổ phiếu giúp nâng cao độ chính xác trong việc phân tích thị trường chứng khoán. LightGBM với khả năng xử lý dữ liệu lớn và tốc độ học nhanh, kết hợp với RNN, một mô hình mạnh mẽ trong việc xử lý dữ liệu chuỗi thời gian, cho phép dự đoán chính xác hơn dựa trên dữ liệu lịch sử. Điều này không chỉ hỗ trợ các nhà đầu tư đưa ra quyết định thông minh hơn mà còn giúp các tổ chức tài chính phát triển các chiến lược đầu tư hiệu quả.

Lợi ích thứ hai là việc cải thiện khả năng quản lý rủi ro. Với các mô hình dự đoán tốt hơn, nhà đầu tư có thể nhận diện sớm các xu hướng xấu, từ đó giảm thiểu rủi ro thua lỗ. Điều này cũng giúp tăng cường niềm tin của nhà đầu tư vào thị trường chứng khoán Việt Nam, đóng góp vào sự phát triển bền vững của nền kinh tế.

Cuối cùng, dự án này cung cấp nền tảng công nghệ và dữ liệu có giá trị cho các nghiên cứu và ứng dụng tiếp theo trong lĩnh vực tài chính và chứng khoán.

## 4.2. Hướng phát triển

Trong tương lai, dự án có thể được mở rộng theo một số hướng sau:

* **Mở rộng phạm vi ứng dụng:** Ngoài việc dự đoán xu hướng giá cổ phiếu, mô hình có thể được tùy chỉnh để áp dụng cho các loại tài sản khác như trái phiếu, hàng hóa, hoặc tiền điện tử. Điều này sẽ làm tăng khả năng ứng dụng thực tiễn của hệ thống trong việc quản lý danh mục đầu tư đa dạng.
* **Tích hợp thêm các yếu tố vĩ mô:** Hiện tại, mô hình chủ yếu dựa trên dữ liệu lịch sử của giá cổ phiếu. Tuy nhiên, việc tích hợp thêm các yếu tố vĩ mô như chính sách kinh tế, lãi suất, và các sự kiện toàn cầu có thể cải thiện đáng kể độ chính xác của dự báo.
* **Phát triển giao diện người dùng:** Xây dựng một giao diện trực quan và dễ sử dụng cho các nhà đầu tư, cho phép họ tương tác với mô hình và nhận được các dự báo một cách nhanh chóng. Việc tích hợp các công cụ này vào các ứng dụng di động hoặc nền tảng giao dịch trực tuyến sẽ làm tăng tính tiện dụng và hiệu quả.

# 5. Team Member Review and Comment

|  |
| --- |
| <ATTACH A TEAM PICTURE HERE> |

|  |  |
| --- | --- |
| NAME | REVIEW and COMMENT |
| Nguyễn Tuấn Đạt |  |
| Đặng Văn Mạnh |  |
| Vũ Minh Đức |  |

6. Instructor Review and Comment

|  |  |  |
| --- | --- | --- |
| CATEGORY | SCORE | REVIEW and COMMENT |
| IDEA | \_\_/10 |  |
| APPLICATION | \_\_/30 |  |
| RESULT | \_\_/30 |  |
| PROJECT MANAGEMENT | \_\_/10 |  |
| PRESENTATION & REPORT | \_\_/20 |  |
| TOTAL | \_\_/100 |  |