

**HỌC VIỆN CHÍNH SÁCH VÀ PHÁT TRIỂN  
KHOA KINH TẾ SỐ**



**THI KẾT THÚC HỌC PHẦN  
MÔN HỌC: PHÂN TÍCH DỮ LIỆU LỚN VỚI R**

**ĐỀ TÀI: PHÂN TÍCH VÀ DỰ ĐOÁN GIÁ CỔ PHIẾU  
CÔNG TY CỔ PHẦN XÂY DỰNG COTECCONS - BẰNG  
PHƯƠNG PHÁP HỌC MÁY MACHINE LEARNING**

**Nguyễn Đắc Chiến – 71131101040**

**Hà Nội 2023**

# MỤC LỤC

<b>PHẦN MỞ ĐẦU .....</b>	<b>1</b>
1. <b>Lý do chọn đề tài .....</b>	<b>1</b>
2. <b>Mục tiêu nghiên cứu.....</b>	<b>2</b>
3. <b>Phạm vi nghiên cứu và cấu trúc bài thi.....</b>	<b>2</b>
<b>PHẦN NỘI DUNG .....</b>	<b>4</b>
<b>CHƯƠNG 1: GIỚI THIỆU KHÁI QUÁT VỀ NGÔN NGỮ R VÀ THỊ</b>	
<b>TRƯỜNG CHỨNG KHOÁN VIỆT NAM .....</b>	<b>4</b>
1.1. <b>Ngôn ngữ R và ứng dụng trong phân tích dữ liệu .....</b>	<b>4</b>
1.2. <b>Tổng quan về thị trường chứng khoán Việt Nam và yếu tố ảnh</b>	
<b>hưởng đến giá cả chứng khoán .....</b>	<b>6</b>
<b>CHƯƠNG 2: CƠ SỞ LÝ THUYẾT VỀ MACHINE LEARNING VÀ DỰ</b>	
<b>ĐOÁN GIÁ CỔ PHIẾU .....</b>	<b>10</b>
2.1. <b>Tổng quan về Machine Learning và ứng dụng trong dự đoán giá cổ</b>	
<b>phiếu 10</b>	
2.2. <b>2.2. Các mô hình trong Machine Learning thường được sử dụng.12</b>	
<b>CHƯƠNG 3: XÂY DỰNG MÔ HÌNH DỰ ĐOÁN VÀ ĐÁNH GIÁ HIỆU</b>	
<b>SUẤT MÔ HÌNH .....</b>	<b>15</b>
3.1. <b>Mô tả dữ liệu .....</b>	<b>15</b>
3.2. <b>Giới thiệu các thư viện .....</b>	<b>16</b>
3.3. <b>Xây dựng mô hình dự đoán giá cổ phiếu bằng phương pháp</b>	
<b>Machine Learning .....</b>	<b>17</b>
3.4. <b>Đánh giá mô hình dự đoán và đưa ra nhận xét.....</b>	<b>23</b>
3.5. <b>Các khuyến nghị cho nhà đầu tư về cổ phiếu Công ty Cổ phần Xây</b>	
<b>dựng Cotecons .....</b>	<b>24</b>
<b>PHẦN KẾT LUẬN.....</b>	<b>27</b>

## **PHẦN MỞ ĐẦU**

### **1. Lý do chọn đề tài**

Thị trường chứng khoán Việt Nam đang trở thành một trong những tâm điểm thu hút sự quan tâm của nhà đầu tư và các tổ chức tài chính trên khắp thế giới. Sự phát triển nhanh chóng và tiềm năng tăng trưởng lớn của thị trường này đã thu hút ánh mắt của những người quan tâm đến tài chính và đầu tư. Tuy nhiên, thị trường chứng khoán không phải lúc nào cũng trải qua những thời kỳ ổn định, và nó thường biến động mạnh dưới sự ảnh hưởng của nhiều yếu tố khác nhau.

Những yếu tố này không chỉ bao gồm tình hình kinh tế và chính trị của Việt Nam, mà còn bao gồm cả tin tức quốc tế và tâm lý thị trường toàn cầu. Sự biến động và không chắc chắn trong thị trường chứng khoán là điều không thể tránh khỏi, và nhà đầu tư đang phải đối mặt với những quyết định đầy rủi ro khi tham gia vào thị trường này.

Trong thị trường chứng khoán với rất nhiều các doanh nghiệp được niêm yết, ta có thể dễ dàng thấy được một doanh nghiệp rất nổi bật được quan tâm bởi nhiều nhà đầu tư đó là Công ty Cổ phần Xây dựng Coteccons. Công ty Cổ phần Xây dựng Coteccons là một trong những công ty xây dựng uy tín trên thị trường Việt Nam. Trải qua nhiều năm hoạt động trong lĩnh vực, hiện tại Coteccons đã ghi điểm trong khách hàng với nhiều dự án hoàn hảo, quy mô tầm cỡ và gây tiếng vang trên thị trường. Hiện tại, công ty có trụ sở chính tại 236/6 Điện Biên Phủ, Phường 17, Quận Bình Thạnh, TP HCM. Công ty này có mã chứng khoán CTD là công ty xây dựng được thành lập ngày 24/08/2004 từ quá trình cổ phần hóa một công ty thành viên thuộc Tổng công ty Fico), cho phép người đầu tư mua bán cổ phiếu của công ty trên thị trường chứng khoán. Điều này làm cho việc dự đoán giá cổ phiếu của Coteccons trở nên quan trọng vì nó có thể ảnh hưởng đến nhiều khía cạnh của thị trường chứng khoán và nền kinh tế.

Để đưa ra quyết định đầu tư hiệu quả, nhà đầu tư đòi hỏi một sự hiểu biết sâu rộng về thị trường và có khả năng dự đoán chính xác về giá cả chứng khoán. Trong bối cảnh công nghệ phát triển, phương pháp học máy đã trở thành một công cụ hấp

dẫn để dự đoán giá cả thị trường chứng khoán. Các mô hình học máy có khả năng xử lý một lượng lớn dữ liệu lịch sử và sử dụng nó để dự đoán xu hướng tương lai của thị trường.

Sự ứng dụng của học máy trong dự đoán giá cổ phiếu mang lại nhiều ưu điểm. Nó có thể giúp nhà đầu tư và các tổ chức tài chính đưa ra quyết định đầu tư thông

minh, dự báo các biến động tiềm ẩn và tối ưu hóa lợi nhuận. Tuy nhiên, dự đoán giá cổ phiếu vẫn là một nhiệm vụ phức tạp và không phải lúc nào cũng chính xác. Học máy chỉ có thể cung cấp thông tin dự đoán dựa trên dữ liệu lịch sử và không thể dự đoán hoàn toàn các biến động không xác định.

Trong bối cảnh công nghệ phát triển, phương pháp học máy machine learning đã phát triển mạnh mẽ và được áp dụng rộng rãi trong nhiều lĩnh vực, bao gồm cả dự đoán giá cổ phiếu. Việc tìm hiểu và áp dụng phương pháp học máy machine learning vào dự đoán giá cổ phiếu có ý nghĩa quan trọng bởi việc dự đoán giá cổ phiếu có thể giúp nhà đầu tư và các tổ chức tài chính đưa ra quyết định đầu tư thông minh và tối ưu hóa lợi nhuận.

Nhận thức rõ ràng tầm quan trọng của ứng dụng công nghệ nói chung, mô hình học máy machine learning nói riêng vào việc dự đoán giá cổ phiếu, em đã chọn đề tài “Dự đoán giá cổ phiếu Công ty Cổ phần Xây dựng Coteccons bằng phương pháp học máy machine learning” làm đề tài nghiên cứu.

## **2. Mục tiêu nghiên cứu**

- Xây dựng một mô hình dự đoán giá cổ phiếu Công ty Cổ phần Xây dựng Coteccons Phát sử dụng phương pháp học máy machine learning.
- Đánh giá hiệu suất của mô hình dự đoán so với các phương pháp truyền thống khác.
- Tìm hiểu yếu tố ảnh hưởng đến giá cổ phiếu và đưa ra các khuyến nghị cho nhà đầu tư khi đầu tư vào cổ phiếu của Công ty Cổ phần Xây dựng Coteccons.

## **3. Phạm vi nghiên cứu và cấu trúc bài tiểu luận**

- Dự đoán giá cả cổ phiếu Công ty Cổ phần Xây dựng Coteccons dựa trên dữ liệu lịch sử và các yếu tố tác động như chỉ số kinh tế, tin tức, tâm lý thị trường,....
- Sử dụng các phương pháp học máy machine learning như hồi quy tuyến tính, cây quyết định, mạng nơ-ron,.... để xây dựng mô hình dự đoán.

- Đánh giá hiệu suất của mô hình dự đoán bằng các chỉ số như sai số trung bình, độ chính xác,....

- Tìm hiểu và phân tích yếu tố ảnh hưởng đến giá cổ phiếu Công ty Cổ phần Xây dựng Cotecons chỉ số kinh tế, tin tức, tâm lý thị trường,.... và đưa ra các khuyến nghị đầu tư.

Cấu trúc bài thi được chia thành 3 chương:

Chương 1: Giới thiệu khái quát về ngôn ngữ R và thị trường chứng khoán Việt Nam

Chương 2: Cơ sở lý thuyết về machine learning và dự đoán giá cổ phiếu

Chương 3: Xây dựng mô hình dự đoán và đánh giá hiệu suất mô hình

## PHẦN NỘI DUNG

### CHƯƠNG 1: GIỚI THIỆU KHÁI QUÁT VỀ NGÔN NGỮ R VÀ THỊ TRƯỜNG CHỨNG KHOÁN VIỆT NAM

#### 1.1. Ngôn ngữ R và ứng dụng trong phân tích dữ liệu

##### *1.1.1. Giới thiệu khái quát về ngôn ngữ R*

Ngôn ngữ R là một ngôn ngữ lập trình và môi trường phát triển phần mềm sử dụng cho phân tích dữ liệu và thống kê. R được phát triển bởi Ross Ihaka và Robert Gentleman vào những năm 1990 và đã trở thành một trong những ngôn ngữ phổ biến nhất trong cộng đồng phân tích dữ liệu.

R cung cấp một loạt các chức năng và gói phần mềm mạnh mẽ để xử lý dữ liệu, thực hiện các phép tính thống kê, vẽ đồ thị và phân tích dữ liệu. R có một cú pháp dễ hiểu và mạnh mẽ, cho phép người dùng thực hiện các phép tính phức tạp và tùy chỉnh một cách linh hoạt.

Một trong những điểm mạnh của R chính là sự đa dạng trong chức năng và gói phần mềm mà nó cung cấp. R có thư viện rất phong phú để xử lý dữ liệu, thực hiện các phép tính thống kê, tạo đồ thị và phân tích dữ liệu. Với cú pháp dễ hiểu, R cho phép người dùng thực hiện phép tính phức tạp một cách tự nhiên.

R là môi trường dành cho việc phân tích dữ liệu và đồ thị thống kê với đa dạng các kỹ thuật phân tích thông kê từ các mô hình tuyến tính và phi tuyến tính, các kỹ thuật kiểm định thống kê, phân tích chuỗi thời gian và vô số các thuật toán nâng cao khác như “machine learning” hay “deep learning”. Điều này làm cho R trở thành một công cụ toàn diện cho mọi nhu cầu phân tích dữ liệu.

Ngoài ra, điểm đáng chú ý khác của R là hỗ trợ các công cụ đồ thị rất chất lượng và linh hoạt. Điểm đặc biệt là R, một ngôn ngữ lập trình mã nguồn mở, được tải, cài đặt và sử dụng hoàn toàn miễn phí cho tất cả người dùng với hơn 17634 gói chức năng (package) cũng như rất nhiều các hàm tính toán thống kê và đồ thị.

R là một công cụ mạnh mẽ và đa dạng, là lựa chọn hàng đầu cho bất kỳ ai muốn thực hiện phân tích dữ liệu và thống kê chuyên nghiệp.

### ***1.1.2. Ứng dụng của ngôn ngữ R trong phân tích dữ liệu***

Ngôn ngữ lập trình R là một trong những công cụ tiên tiến mới nhất trong thống kê và phân tích dữ liệu. R là một công cụ rất mạnh cho học máy, thống kê và phân tích dữ liệu. R có đóng góp to lớn trong ngành phân tích dữ liệu. Tất cả những việc liên quan đến tìm tòi, khai thác, thu thập và phân tích dữ liệu để đưa ra được những thông tin có giá trị cho một doanh nghiệp R đều có thể thực hiện được. Ngày nay, hàng triệu nhà phân tích, nhà nghiên cứu và thương hiệu lớn như Facebook, Google, Bing, Accdvisor, Wipro đang sử dụng R để giải quyết các vấn đề phức tạp trong kinh doanh. Dưới đây là một số ứng dụng của R trong phân tích dữ liệu:

- **Thống kê mô tả:** R cung cấp các công cụ mạnh mẽ để tính toán các thống kê mô tả như trung bình, phương sai, độ lệch chuẩn, phân phối, và các đặc trưng khác của dữ liệu.
- **Khai phá dữ liệu:** R hỗ trợ các công cụ khai phá dữ liệu để tìm hiểu và khám phá thông tin từ dữ liệu. Điều này bao gồm việc tạo biểu đồ, biểu đồ hộp, biểu đồ phân phối, biểu đồ tương quan và các công cụ khác để trực quan hóa dữ liệu.
- **Mô hình hóa thống kê:** R cung cấp các gói phần mềm mạnh mẽ để xây dựng và kiểm tra các mô hình thống kê như hồi quy tuyến tính, hồi quy logistic, phân tích phương sai (ANOVA), và các mô hình khác.
- **Phân tích dữ liệu đa biến:** R có các gói phần mềm tiên tiến để thực hiện phân tích dữ liệu đa biến như phân tích thành phần chính (PCA), phân tích nhân tố, phân tích cụm, và các phương pháp khác để nén và tìm kiếm thông tin từ dữ liệu đa biến.
- **Mô phỏng và mô hình hóa:** R cho phép việc tạo ra biểu đồ và biểu đồ tương tác để trực quan hóa dữ liệu. Nó cũng hỗ trợ việc tạo ra mô hình dự đoán, mô hình phân loại, và mô hình phân cụm để giúp hiểu và dự đoán dữ liệu.
- **Xử lý và làm sạch dữ liệu:** R có các gói phần mềm để xử lý và làm sạch dữ liệu như loại bỏ giá trị trống, loại bỏ nhiễu, chuyển đổi kiểu dữ liệu, và xử lý các vấn đề khác liên quan đến dữ liệu không đồng nhất.



- **Trực quan hóa dữ liệu:** R cung cấp các công cụ để tạo ra biểu đồ và biểu đồ chất lượng cao để trực quan hóa dữ liệu. Điều này giúp hiểu và trình bày thông tin từ dữ liệu một cách dễ dàng và hấp dẫn.

- **Lập trình và tùy chỉnh:** R là một ngôn ngữ lập trình linh hoạt, cho phép người dùng tùy chỉnh và mở rộng các chức năng của nó. Người dùng có thể viết các chương trình R riêng để thực hiện các tác vụ phức tạp và giải quyết các vấn đề riêng của họ.

R không chỉ là một công cụ phân tích dữ liệu mạnh mẽ, mà còn là một môi trường linh hoạt cho việc thực hiện nhiều nhiệm vụ khác nhau, từ khám phá dữ liệu đến xây dựng mô hình thống kê và trực quan hóa thông tin.

## **1.2. Tổng quan về thị trường chứng khoán Việt Nam và yếu tố ảnh hưởng đến giá cả chứng khoán**

### ***1.2.1. Tổng quan về thị trường chứng khoán Việt Nam***

Dù là một nền kinh tế còn non trẻ, thị trường chứng khoán Việt Nam được hình thành và phát triển sau 20 năm nhưng cũng đóng góp nhiều cho hoạt động kinh tế vĩ mô của đất nước.

Mọi biến động trên thị trường chứng khoán Việt Nam đều có thể tác động đến nền kinh tế nước ta, vì đây là kênh dẫn vốn quan trọng cho nền kinh tế, cũng là kênh đầu tư tiềm năng của công chúng. Thị trường chứng khoán có khả năng tạo thanh khoản cao, tập trung phân phối vốn và chuyển thời hạn vốn phù hợp với nhu cầu phát triển kinh tế. Nhờ vào thị trường chứng khoán, Chính phủ có thể huy động các nguồn lực tài chính mà không phải chịu áp lực về lạm phát, đặc biệt khi nguồn vốn đầu tư khu vực của Nhà nước còn nhiều hạn chế.

Theo nhiều chuyên gia tài chính - kinh tế hàng đầu, thị trường chứng khoán có thể phản ánh chính xác triển vọng của một nền kinh tế thay đổi như thế nào theo chu kỳ nửa năm. Cụ thể là giá chứng khoán tăng sẽ cho thấy nền kinh tế đang phát triển và ngược lại, giá chứng khoán giảm lại là dự báo không mấy tốt đẹp về triển vọng của một nền kinh tế trong tương lai gần.

Thị trường chứng khoán Việt Nam đã hỗ trợ rất tốt cho hoạt động cổ phần hóa doanh nghiệp Nhà nước. Bên cạnh đó, TTCK cũng phối hợp với hệ thống tín dụng của ngân hàng tạo ra 1 cơ cấu thị trường vốn Việt Nam cân đối hơn, hiệu quả hơn, hỗ trợ mạnh mẽ cho nền kinh tế.

Thị trường chứng khoán Việt Nam đã có sự phát triển đáng kể trong những năm gần đây. Tổng giá trị vốn hóa thị trường và khối lượng giao dịch đã tăng lên đáng kể. Các chỉ số thị trường như VN-Index và HNX-Index thường xuyên ghi nhận sự biến động. Từ đó thị trường này cũng thu hút được sự quan tâm của cả nhà đầu tư cá nhân và tổ chức tài chính. Các công ty chứng khoán, quỹ đầu tư và các cá nhân đầu tư tự do là các người tham gia chính.

Mặc dù thị trường chứng khoán Việt Nam đã phát triển đáng kể, nó còn phải đối mặt với các thách thức như biến động lớn, quản lý rủi ro, và nâng cao tính minh bạch. Tuy nhiên, triển vọng của thị trường vẫn rất lớn, và nó có thể trở thành điểm đến hấp dẫn cho nhà đầu tư trong và ngoài nước trong tương lai.

### ***1.2.2. Yếu tố ảnh hưởng đến giá cổ phiếu***

Giá cổ phiếu là mức giá của cổ phiếu tại một thời điểm nhất định. Tức là số tiền mà nhà đầu tư cần bỏ ra để mua một đơn vị cổ phiếu tại thời điểm hiện tại đang giao dịch trên thị trường. Giá cổ phiếu là dữ liệu quan trọng để đánh giá doanh nghiệp có đáng đầu tư không, hay tình hình kinh tế của các đơn vị phát hành. Dựa trên giá cổ phiếu, người chơi quyết định giao dịch mua bán để nắm bắt cơ hội sinh lời, tạo ra lợi nhuận. Giá cổ phiếu của một công ty hoặc tài sản cụ thể có thể bị ảnh hưởng bởi nhiều yếu tố khác nhau. Một số yếu tố quan trọng có thể ảnh hưởng đến giá cổ phiếu:

- **Kết quả kinh doanh của công ty:** Hiệu suất tài chính của công ty, bao gồm doanh thu, lợi nhuận, và tỷ suất sinh lời, có thể ảnh hưởng đáng kể đến giá cổ phiếu. Khi công ty có kết quả kinh doanh tốt, giá cổ phiếu thường tăng lên và ngược lại.
- **Tình hình tài chính:** Tình hình tài chính của công ty, bao gồm nợ nần, lưu chuyển tiền mặt, và tài sản, có thể ảnh hưởng đến sự tin tưởng của nhà đầu tư. Một

công ty có tình hình tài chính ổn định thường có khả năng thu hút các nhà đầu tư hơn.

- Thị trường và ngành công nghiệp: Tình hình trong ngành công nghiệp của công ty và điều kiện thị trường chung cũng có thể ảnh hưởng đến giá cổ phiếu. Ví dụ, một ngành công nghiệp đang phát triển có thể giúp tăng giá cổ phiếu của các công ty trong ngành đó.

- Tin tức và sự kiện: Các sự kiện và tin tức, bao gồm thông tin về công ty, sản phẩm, dự án, và vấn đề pháp lý có thể tác động đến giá cổ phiếu. Tin tức tích cực hoặc tiêu cực có thể thay đổi quan điểm của thị trường về công ty.

- Chính trị và kinh tế toàn cầu: Chính trị và tình hình kinh tế toàn cầu có thể tác động đến thị trường chứng khoán. Biến động trong tình hình kinh tế toàn cầu, chính trị, và thương mại có thể ảnh hưởng đến giá cổ phiếu.

- Tâm lý thị trường: Tâm lý của các nhà đầu tư và giao dịch viên cũng có thể gây ra sự biến động trong giá cổ phiếu. Thị trường có thể phản ứng mạnh mẽ theo cả hai hướng, tích cực hoặc tiêu cực, dựa trên tâm lý thị trường.

- Yếu tố cơ bản: Các yếu tố cơ bản như P/E ratio (tỷ lệ giá trị thị trường/giá trị tài sản ròng), P/B ratio (tỷ lệ giá trị thị trường/giá trị sổ sách), và EPS (lợi nhuận trên mỗi cổ phiếu) cũng có thể ảnh hưởng đến giá cổ phiếu.

- Chính sách tiền tệ và lãi suất: Chính sách tiền tệ của ngân hàng trung ương và lãi suất cơ bản có thể ảnh hưởng đến cả thị trường chứng khoán và giá cổ phiếu.

### ***1.2.3. Tổng quan về giá cổ phiếu Công ty Cổ phần Xây dựng Coteccons***

- Lịch sử hình thành

Coteccons (Công ty Cổ phần Xây dựng Coteccons) là một trong những tập đoàn xây dựng lớn và uy tín tại Việt Nam. Công ty này hoạt động trong lĩnh vực xây dựng và phát triển bất động sản và đã có nhiều dự án đáng chú ý trên cả nước. Dưới đây là một số thông tin giới thiệu về Coteccons:

Lịch sử: Coteccons được thành lập vào năm 2004 và đã trải qua một hành trình phát triển đáng kể từ đó. Trong suốt hơn một thập kỷ, Coteccons đã trở thành một trong những công ty xây dựng hàng đầu tại Việt Nam.

- Lĩnh vực hoạt động:

**Xây dựng:** Coteccons chuyên về xây dựng dự án công trình dân dụng và công nghiệp. Công ty đã thực hiện nhiều dự án quy mô lớn như tòa nhà chung cư, văn phòng, trung tâm thương mại, và cơ sở hạ tầng.

**Phát triển bất động sản:** Coteccons cũng tham gia vào lĩnh vực phát triển bất động sản và đã xây dựng nhiều dự án căn hộ và khu đô thị.

**Dự án nổi bật:** Coteccons đã thực hiện nhiều dự án nổi bật tại Việt Nam. Một số dự án đáng chú ý bao gồm tòa nhà Vincom Center Landmark 81, tòa nhà Bitexco Financial Tower, khu đô thị Vinhomes Central Park, và nhiều dự án khác.

**Uy tín và chất lượng:** Coteccons được biết đến với cam kết về chất lượng và sự tận tâm đối với từng dự án. Công ty này đã đạt được nhiều giải thưởng và chứng chỉ uy tín trong ngành xây dựng.

**Nhân lực:** Coteccons có đội ngũ nhân viên giàu kinh nghiệm và được đào tạo chuyên sâu trong các lĩnh vực liên quan đến xây dựng và quản lý dự án.

**Phát triển bền vững:** Công ty luôn hướng đến phát triển bền vững và thực hiện các dự án xanh, hữu ích cho cộng đồng.

Coteccons đã đóng góp đáng kể cho sự phát triển hạ tầng và bất động sản của Việt Nam và được coi là một trong những đối tác đáng tin cậy trong ngành xây dựng.

## CHƯƠNG 2: CƠ SỞ LÝ THUYẾT VỀ MACHINE LEARNING VÀ DỰ ĐOÁN GIÁ CỔ PHIẾU

### 2.1. Tổng quan về Machine Learning và ứng dụng trong dự đoán giá cổ phiếu

#### 2.1.1. Tổng quan về Machine Learning

Machine learning là một nhánh của trí tuệ nhân tạo (AI) và khoa học máy tính, tập trung vào việc sử dụng dữ liệu và thuật toán để bắt chước hành động của con người, dần dần cải thiện độ chính xác của nó.

Machine learning còn là một thành phần quan trọng của lĩnh vực khoa học dữ liệu đang phát triển. Thông qua việc sử dụng các phương pháp thống kê, các thuật toán được đào tạo để đưa ra các phân loại hoặc dự đoán và khám phá những thông tin chi tiết từ chính các dự án khai thác dữ liệu.

Thông qua các thông tin chi tiết có được để thúc đẩy việc đưa ra quyết định đối với các ứng dụng và doanh nghiệp, tác động mạnh đến các chỉ số tăng trưởng. Khi dữ liệu lớn tiếp tục nhu cầu mở rộng và phát triển đòi hỏi nhu cầu tuyển dụng các nhà khoa học dữ liệu sẽ tăng lên. Họ sẽ được yêu cầu giúp xác định các câu hỏi kinh doanh có liên quan nhất và dữ liệu để trả lời chúng.

Bài toán của machine learning thường được chia làm hai loại là dự đoán (prediction) và phân loại (classification). Các bài toán dự đoán thường là giá nhà, giá xe, v.v, còn các bài toán phân loại thường là nhận diện chữ viết tay, đồ vật, v.v.

#### 2.1.2. Ứng dụng của Machine Learning trong dự đoán giá cổ phiếu

Như bất kỳ ai trong chúng ta cũng có thể đoán, thị trường chứng khoán không ổn định và thường không thể đoán trước được. Trong nhiều thập kỷ, các nhà nghiên cứu đã sử dụng dữ liệu chuỗi thời gian để dự đoán giá trị tương lai - trong đó ứng dụng thách thức nhất và có khả năng sinh lợi nhất là dự đoán giá trị cổ phiếu của một công ty nhất định. Tuy nhiên, đúng như dự đoán, sự thay đổi của thị trường phụ thuộc vào nhiều thông số mà chỉ có thể định lượng được một số thông số, chẳng hạn như dữ liệu chứng khoán lịch sử, khối lượng giao dịch, giá cả hiện tại. Tất nhiên, các yếu tố cơ bản như giá trị nội tại, tài sản, hiệu suất hàng quý, các khoản đầu tư

gần đây và chiến lược của công ty đều ảnh hưởng đến niềm tin của nhà giao dịch đối với công ty và do đó ảnh hưởng đến giá cổ phiếu của công ty đó. Chỉ một số ít trong số đó có thể được kết hợp một cách hiệu quả vào mô hình toán học. Điều này khiến cho việc dự đoán giá cổ phiếu bằng machine learning trở nên khó khăn và không đáng tin cậy ở một mức độ nhất định. Hơn nữa, gần như không thể lường trước được một tin tức nào đó sẽ làm rung chuyển hoặc thúc đẩy thị trường chứng khoán trong những tuần tới – một đại dịch hay một cuộc chiến tranh.

Vì vậy, thay vì tập trung vào việc khớp các giá trị thực tế với độ chính xác cao, các nhà phân tích chỉ tập trung vào việc đưa ra các dự đoán ngắn hạn để có được ước tính xác suất về những gì thị trường “có thể” sẽ sớm trông như thế nào. Với đủ dữ liệu lịch sử và các tính năng hữu ích, các mô hình học máy machine learning có thể dự đoán những biến động ngắn hạn trên thị trường trong một ngày thị trường bình thường, không có biến động.

Dự đoán giá cổ phiếu bằng phương pháp học máy machine learning là quá trình dự đoán giá trị tương lai của một cổ phiếu được giao dịch trên sàn giao dịch chứng khoán để thu lợi nhuận. Với nhiều yếu tố liên quan đến việc dự đoán giá cổ phiếu, việc dự đoán giá cổ phiếu với độ chính xác cao là một thách thức và đây là lúc máy học đóng vai trò quan trọng.

Machine learning có thể sử dụng các mô hình hồi quy tuyến tính hoặc phi tuyến tính để dự đoán giá cổ phiếu dựa trên các biến đầu vào như giá cổ phiếu trong quá khứ, thể hiện tài chính của công ty, và các yếu tố thị trường khác. Mô hình hồi quy có thể cung cấp một ước tính liên quan đến giá cổ phiếu trong tương lai. Ngoài ra, có thể sử dụng các mô hình chuỗi thời gian như ARIMA (Autoregressive Integrated Moving Average) hoặc các biến thể của LSTM (Long Short-Term Memory) để phân tích và dự đoán sự biến động của giá cổ phiếu theo thời gian. Machine learning có thể sử dụng phân tích tâm lý thị trường để đánh giá tác động của tin tức và cảm xúc của nhà đầu tư đối với giá cổ phiếu. Các công cụ này có thể theo dõi các diễn biến trên mạng xã hội, trang web tin tức, và phương tiện truyền thông xã hội khác để đưa ra dự đoán. Không những thế, Machine learning còn có khả năng khai thác dữ liệu không cấu trúc như văn bản, âm thanh và hình ảnh từ các nguồn như báo cáo tài

chính, phát biểu của CEO, hoặc hình ảnh sản phẩm để đưa ra dự đoán giá cổ phiếu....

## **2.2. Các mô hình trong Machine Learning thường được sử dụng**

### ***2.2.1. Random Forest***

Random forest là một phương pháp thống kê mô hình hóa bằng máy (machine learning statistic) 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 (Decision tree). Random Forest cho thấy hiệu quả hơn so với thuật toán phân loại thường được sử dụng vì có khả năng tìm ra thuộc tính nào quan trọng hơn so với những thuộc tính khác. Trên thực tế, nó còn có thể chỉ ra rằng một số thuộc tính là không có tác dụng trong cây quyết định. Trong phạm vi bài báo này, nhóm nghiên cứu giới hạn phạm vi trong công tác khảo sát tính khoa học của phương pháp và định hướng việc ứng dụng phương pháp cho công tác phân loại ảnh viễn thám có kiểm định. Kết quả thử nghiệm cho thấy khả năng ứng dụng phương pháp Random forest vào trong công tác phân loại có kiểm định ảnh viễn thám là hoàn toàn khả thi.

### ***2.2.2. Time Series Models (Mô hình chuỗi thời gian)***

Dự báo chuỗi thời gian (Time Series Forecasting) là một kỹ thuật dự đoán các sự kiện thông qua một chuỗi thời gian. Nó dự đoán các sự kiện trong tương lai bằng cách phân tích các xu hướng trong quá khứ, với giả định rằng các xu hướng trong tương lai sẽ tương tự như vậy.

Đối với dự đoán giá cổ phiếu theo thời gian, các mô hình chuỗi thời gian như ARIMA (Autoregressive Integrated Moving Average) và GARCH (Generalized Autoregressive Conditional Heteroskedasticity) thường được sử dụng để mô phỏng và dự đoán sự biến động giá cổ phiếu.

Mô hình ARIMA (Autoregressive Integrated Moving Average) là một trong những công cụ phổ biến được sử dụng trong phân tích chuỗi thời gian để dự đoán giá cổ phiếu. ARIMA có thể sử dụng để dự đoán giá cổ phiếu trong tương lai dựa trên dữ liệu lịch sử. Nó có thể phát hiện và xác định xu hướng (trend) và yếu tố mùa vụ (seasonality) trong dữ liệu giá cổ phiếu. Điều này giúp nhà đầu tư và nhà phân tích thị trường hiểu rõ hơn về cách giá cổ phiếu thay đổi theo thời gian. Nhà đầu tư

có thể sử dụng dự đoán từ ARIMA để hỗ trợ quyết định giao dịch, chẳng hạn như quyết định mua hoặc bán cổ phiếu dựa trên sự biến động được dự đoán.

Mô hình Dự đoán Google Search Trends cũng là một mô hình được ứng dụng rộng rãi trong việc dự đoán giá cổ phiếu và quản lý rủi ro trong lĩnh vực tài chính. Mô hình Dự đoán Google Search Trends có thể giúp dự đoán mức độ biến động của giá cổ phiếu và điều này quan trọng cho việc đánh giá rủi ro và xác định kích thước của vị thế giao dịch. Mô hình Dự đoán Google Search Trends cho phép xác định các ngưỡng rủi ro dựa trên biến động dự đoán. Nhà đầu tư và người quản lý tài sản có thể sử dụng thông tin này để xác định ngưỡng mức rủi ro chấp nhận được và đưa ra quyết định giao dịch tương ứng. Mô hình Dự đoán Google Search Trends có thể được sử dụng để đánh giá hiệu suất của các chiến lược đầu tư dựa trên giá cổ phiếu. Nó giúp xác định liệu một chiến lược cụ thể có độ ổn định và hiệu quả không, bằng cách so sánh biến động thực tế với biến động dự đoán.

Hai mô hình này đã trở nên rất phổ biến trong thời gian hiện nay. Chúng có những ưu và nhược điểm khác nhau. Mô hình Dự đoán Google Search Trends tập trung vào việc mô hình hóa biến động (volatility) của dữ liệu, trong khi ARIMA tập trung vào việc mô hình hóa giá trung bình của chuỗi thời gian. Mô hình Dự đoán Google Search Trends thường được sử dụng cho dữ liệu có tính biến động không đều, trong khi ARIMA thích hợp cho dữ liệu có tính ổn định hơn. Sự lựa chọn giữa hai mô hình này phụ thuộc vào mục tiêu cụ thể của phân tích và đặc điểm của dữ liệu được nghiên cứu.

### ***2.2.3. Linear Regression (Hồi quy tuyến tính)***

Hồi quy tuyến tính là một kỹ thuật phân tích dữ liệu dự đoán giá trị của dữ liệu không xác định bằng cách sử dụng một giá trị dữ liệu liên quan và đã biết khác. Nó mô hình toán học biến không xác định hoặc phụ thuộc và biến đã biết hoặc độc lập như một phương trình tuyến tính.

Các mô hình hồi quy tuyến tính tương đối đơn giản và cung cấp một công thức toán học dễ giải thích để đưa ra các dự đoán. Hồi quy tuyến tính là một kỹ thuật thống kê được sử dụng từ lâu và áp dụng dễ dàng cho phần mềm và tính toán. Các



doanh nghiệp sử dụng nó để chuyển đổi dữ liệu thô một cách đáng tin cậy và có thể dự đoán được thành nghiệp vụ thông minh và thông tin chuyên sâu hữu ích. Các nhà

khoa học trong nhiều lĩnh vực, bao gồm sinh học và các ngành khoa học hành vi, môi trường, và xã hội, sử dụng hồi quy tuyến tính để tiến hành phân tích dữ liệu sơ bộ và dự đoán các xu hướng tương lai. Nhiều phương pháp khoa học dữ liệu, chẳng hạn như máy học và trí tuệ nhân tạo, sử dụng hồi quy tuyến tính để giải quyết các bài toán phức tạp.

## CHƯƠNG 3: XÂY DỰNG MÔ HÌNH DỰ ĐOÁN VÀ ĐÁNH GIÁ HIỆU SUẤT MÔ HÌNH

### 3.1. Mô tả dữ liệu

Dữ liệu mô tả chỉ số VNIndex (chỉ số cuối ngày) từ 10/23/2018 đến 10/23/2023, dữ liệu bao gồm giai đoạn ảnh hưởng của đại dịch Covid-19.

Trong tập dữ liệu thu thập được từ AmiBroker của chỉ số VN-index bao gồm 6 cột dữ liệu:

	A	B	C	D	E	F	G
1	Ticker	Date/Time	Open	High	Low	Close	Volume
2	CTD	10/23/2018 0:00	104.32	104.32	100.59	101.35	139680
3	CTD	10/24/2018 0:00	100.32	102.04	99.97	99.97	54090
4	CTD	10/25/2018 0:00	96.52	99.9	96.52	99.28	42890
5	CTD	10/26/2018 0:00	100.66	101.35	96.52	97.9	93570
6	CTD	10/29/2018 0:00	97.28	97.9	95.9	95.9	73120
7	CTD	10/30/2018 0:00	95.84	98.45	95.49	96.52	82480
8	CTD	10/31/2018 0:00	98.94	100.39	97.9	100.39	99370
9	CTD	11/1/2018 0:00	100.39	101.49	100.32	101.35	65150
10	CTD	11/2/2018 0:00	100.73	105.07	100.73	104.45	92460
11	CTD	11/5/2018 0:00	104.45	104.66	103.42	104.25	55960
12	CTD	11/6/2018 0:00	104.25	105.14	103.49	104.8	59490
13	CTD	11/7/2018 0:00	104.45	105	103.56	104.8	76160
14	CTD	11/8/2018 0:00	104.94	104.94	104.11	104.8	32920
15	CTD	11/9/2018 0:00	103.49	104.73	103.21	103.35	40270
16	CTD	11/12/2018 0:00	102.04	102.04	100.66	101.83	20690
17	CTD	11/13/2018 0:00	98.25	101.28	98.25	99.97	59500
18	CTD	11/14/2018 0:00	99.97	102.18	99.97	100.66	47960
19	CTD	11/15/2018 0:00	100.66	101.97	100.32	100.32	82760
20	CTD	11/16/2018 0:00	101.35	101.97	100.25	101.97	67630
21	CTD	11/19/2018 0:00	102.73	104.11	102.04	103.42	65950
22	CTD	11/20/2018 0:00	103.42	104.11	103.07	104.11	150980
23	CTD	11/21/2018 0:00	104.04	105.83	103.42	105.83	254590
24	CTD	11/22/2018 0:00	106.73	106.73	105.35	105.83	18320
25	CTD	11/23/2018 0:00	105.56	105.56	104.45	104.59	22320
26	CTD	11/26/2018 0:00	101.35	104.59	101.35	103.76	55340
27	CTD	11/27/2018 0:00	106.18	107.62	106.18	106.87	190240
28	CTD	11/28/2018 0:00	106.87	108.93	106.87	108.45	80420
29	CTD	11/29/2018 0:00	109.21	109.21	107.56	108.25	53610

Trong đó:

- *Date/Time*: Thể hiện ngày thị trường chứng khoán mở cửa hoạt động để giao dịch
- *Open*: Giá mở cửa
- *High*: Giá trần

- *Low*: Giá sàn
- *Close*: Giá đóng cửa
- *Volume*: Tổng khối lượng cổ phiếu đã thực hiện giao dịch trong ngày

### 3.2. Giới thiệu các thư viện

#### *Cài đặt và nạp các thư viện cần thiết*

**forecast**: Gói này cung cấp các công cụ và hàm để thực hiện dự đoán và dự báo dữ liệu, đặc biệt là trong lĩnh vực dự đoán chuỗi thời gian. Nó bao gồm các hàm cho việc xây dựng mô hình và thực hiện dự đoán chuỗi thời gian.

**readxl**: Gói này giúp bạn đọc dữ liệu từ các tệp Excel. Nó cho phép bạn nạp dữ liệu từ các tệp .xlsx vào R để xử lý và phân tích.

**ggplot2**: Gói này được sử dụng để tạo biểu đồ đẹp và linh hoạt bằng cách sử dụng cú pháp dựa trên lớp (grammar of graphics). Nó cung cấp một cách mạnh mẽ để tạo biểu đồ thống kê và trực quan hóa dữ liệu.

**TTR (Technical Trading Rules)**: Gói này cung cấp các hàm để phân tích dữ liệu tài chính và tạo các chỉ số kỹ thuật cho phân tích kỹ thuật. Nó thường được sử dụng trong lĩnh vực phân tích thị trường tài chính.

**gridExtra**: Gói này giúp bạn tạo ra các bố cục (layouts) phức tạp cho biểu đồ và đồ thị trong R, cho phép bạn kết hợp nhiều biểu đồ lại với nhau trong một hình ảnh.

**plotly**: Gói này cho phép bạn tạo ra các biểu đồ tương tác và đồ thị trực quan bằng cách sử dụng thư viện dựa trên web Plotly. Nó cho phép bạn tạo biểu đồ tương tác cho ứng dụng web hoặc bài viết.

**xts (Extensible Time Series)**: Gói này cung cấp một cấu trúc dữ liệu và các hàm cho việc làm việc với chuỗi thời gian. Nó giúp bạn quản lý và phân tích dữ liệu theo thời gian một cách hiệu quả.\

**tibble:** Gói này là một phiên bản hiện đại của DataFrame trong R. Nó cung cấp cách hiển thị và xử lý dữ liệu tabular một cách linh hoạt hơn so với DataFrame cổ điển.

**quantmod:** Gói này được sử dụng trong lĩnh vực tài chính và phân tích dữ liệu thị trường tài chính. Nó cung cấp các công cụ để tải và quản lý dữ liệu tài chính và tạo các biểu đồ và đồ thị liên quan đến tài chính.

**tseries:** Gói này cung cấp các hàm và công cụ cho việc phân tích chuỗi thời gian và dự đoán. Nó bao gồm các phân tích thống kê và mô hình hóa cho chuỗi thời gian.

**fGarch (Financial GARCH Models):** Gói này làm việc với mô hình GARCH (Generalized Autoregressive Conditional Heteroskedasticity) trong lĩnh vực tài chính. Nó cho phép bạn mô hình hóa và dự đoán biến động trong tài chính và giúp bạn quản lý rủi ro tài chính.

**library(prophet):** "Prophet" thường được sử dụng trong các ứng dụng dự đoán thời gian, như dự đoán doanh số bán hàng, lưu lượng truy cập trang web, dự đoán giá cổ phiếu, và nhiều ứng dụng khác liên quan đến thời gian.

### **3.3. Xây dựng mô hình phân tích và dự đoán giá cổ phiếu bằng phương pháp MachineLearning.**

#### **3.3.1. Đường MA của cổ phiếu CTD.**

Trực quan hóa dữ liệu tạo biểu đồ giá và dòng tiền của cổ phiếu CTD



Hình 1: Biểu đồ giá và dòng tiền của cổ phiếu CTD



Hình 2: Ba đường MA20, MA50, MA100.

Các đường trung bình MA20, MA50 và MA100 là ba loại đường trung bình thường được sử dụng trong phân tích kỹ thuật và dự báo chuỗi thời gian. Dưới đây là một sơ lược về mỗi loại :

**MA20 (Moving Average 20):** Đây là đường trung bình động của 20 ngày gần nhất của dữ liệu. MA20 được sử dụng để giảm đi sự biến động ngắn hạn và làm nổi bật xu hướng dài hạn trong dữ liệu. Nó thường được sử dụng để xác định xu hướng ngắn hạn của một tài sản hoặc thị trường tài chính. Nếu MA20 cắt đường giá và nằm trên đường giá thì trong ngắn hạn cổ phiếu sẽ giảm và ngược lại.

**MA50 (Moving Average 50):** Đây là đường trung bình động của 50 ngày gần nhất của dữ liệu. MA50 là một chỉ số phổ biến để theo dõi xu hướng giá trung hạn của một tài sản hoặc thị trường. Nó giúp làm mờ các biến động ngắn hạn và cung cấp một cái nhìn tổng quan về hành vi giá của tài sản. Nếu MA50 cắt đường giá và nằm trên đường giá thì trong trung hạn giá cổ phiếu sẽ có xu hướng giảm và ngược lại.

**MA100 (Moving Average 100):** Đây là đường trung bình động của 100 ngày gần nhất của dữ liệu. MA100 là một chỉ số dùng để theo dõi xu hướng dài hạn của một tài sản hoặc thị trường tài chính. Nó thường được sử dụng để xác định xu hướng chung và đánh giá sự biến động dài hạn. Nếu MA100 cắt đường giá, nằm trên đường giá và đâm xuống thì trong dài hạn cổ phiếu có xu hướng down trend.

### ***3.3.2. Phân tích và dự báo lãi theo thời gian của công ty CTD.***

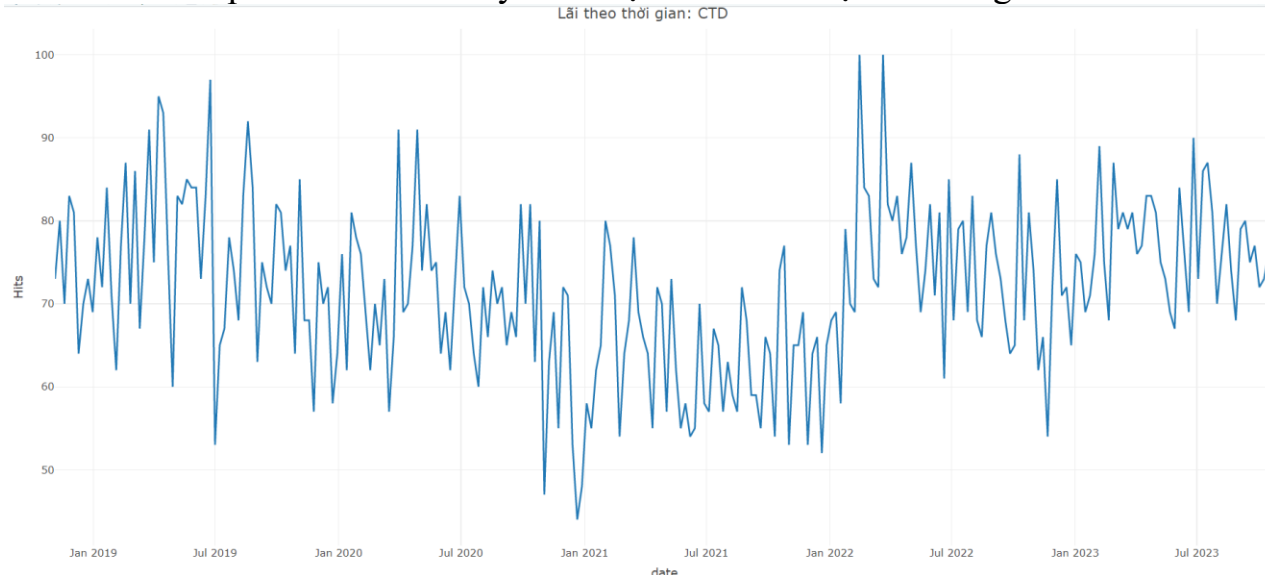
Biểu đồ lãi suất trên biểu đồ lựa chọn lợi suất hàng ngày của cổ phiếu CTD. Lãi suất được tính bằng cách lấy sự thay đổi tỷ lệ tự nhiên của giá đóng cửa hàng ngày theo thời gian, cho thấy biến động hàng ngày của giá cổ phiếu. Dưới đây là các điểm chính trong phân tích này:

**Biểu đồ Lãi Suất:** Biểu đồ hiển thị sự biến động hàng ngày của lãi suất của cổ phiếu CTD. Các giá trị dương cho thấy lãi suất tăng, trong khi các giá trị âm cho thấy lãi suất giảm. Biểu đồ này giúp xác định mức biến động trong giá cổ phiếu và có thể là một dấu hiệu về sự biến động trong thị trường.

**Điểm Chốt Lãi Suất (Horizontal Line):** Đường ngang ở mức 0 trên biểu đồ là điểm chốt lãi suất. Nếu biểu đồ lãi suất vượt qua đường này từ dưới lên, điều này có thể chỉ ra sự gia tăng của lãi suất và ngược lại, nếu nó vượt qua từ trên xuống, điều này có thể chỉ ra sự giảm lãi suất.

Phân tích biểu đồ lãi suất theo thời gian có thể giúp nhà đầu tư và nhà phân tích thị trường hiểu rõ hơn về biến động của giá cổ phiếu CTD theo thời gian và có thể dự

đoán sự biến động trong tương lai. Tuy nhiên, việc hiểu biểu đồ này đòi hỏi kiến thức về phân tích kỹ thuật và thị trường tài chính.



Hình 3 : Biểu đồ lãi của CTD qua từng năm

Lợi ích từ dự báo tình hình thị trường: Biểu đồ này cho phép công ty CTD theo dõi và đánh giá sự quan tâm của người dùng tới sản phẩm hoặc dịch vụ của họ thông qua số lượt tìm kiếm trên Google. Khi số lượt tìm kiếm tăng lên, có thể hiểu rằng có một sự quan tâm lớn hơn từ phía khách hàng về sản phẩm hoặc thông tin liên quan đến công ty. Điều này có thể giúp công ty dự báo tình hình thị trường và thay đổi chiến lược kinh doanh một cách hiệu quả hơn.

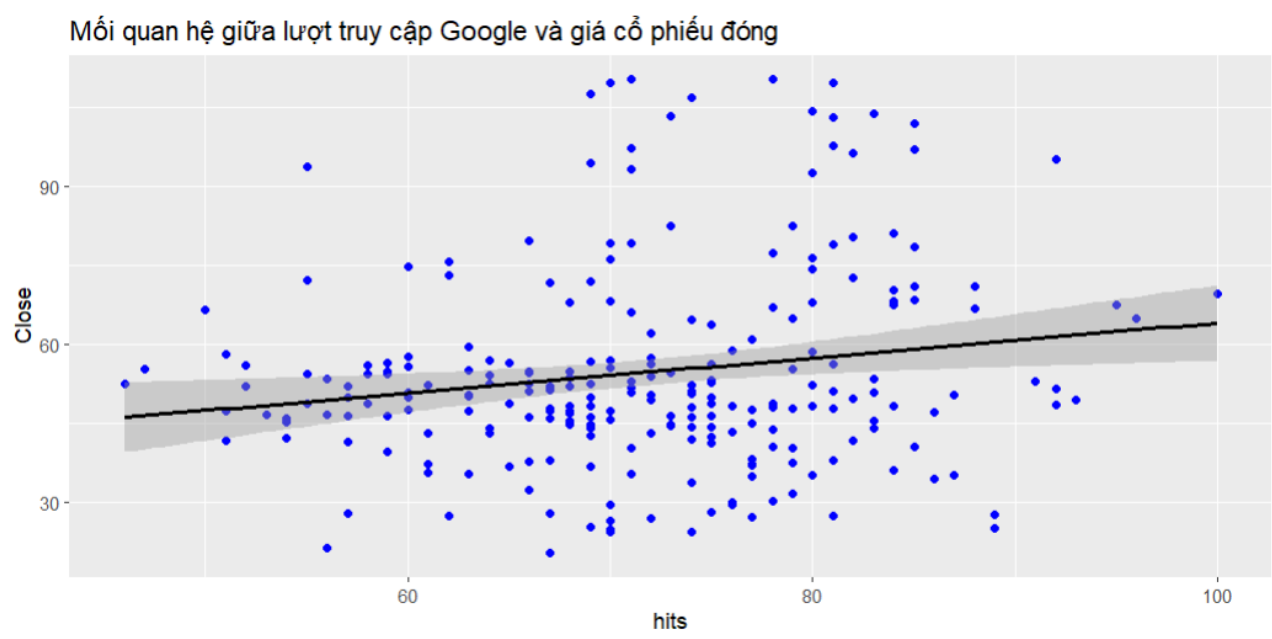
Đánh giá tác động của sự kiện và chiến dịch tiếp thị: Biểu đồ này cũng có thể giúp công ty đánh giá tác động của các sự kiện hoặc chiến dịch tiếp thị trực tuyến đối với sự quan tâm của người dùng. Khi có một chiến dịch quảng cáo trực tuyến hoặc sự kiện quan trọng, công ty có thể theo dõi sự thay đổi trong số lượt tìm kiếm Google để đo lường hiệu suất của chiến dịch và tác động của nó đến sự quan tâm của khách hàng.

Hỗ trợ ra quyết định kế hoạch sản phẩm và tiếp thị: Mối quan hệ giữa số lượt tìm kiếm và giá cổ phiếu có thể giúp công ty ra quyết định về kế hoạch sản phẩm và tiếp thị. Khi có một tăng trưởng đáng kể trong số lượt tìm kiếm, công ty có thể xem xét việc mở rộng sản phẩm hoặc tăng nguồn cung cấp để đáp ứng nhu cầu tăng của thị trường.

Tạo sự tin tưởng từ nhà đầu tư: Sự gia tăng sự quan tâm từ cộng đồng đầu tư có thể tạo ra sự tin tưởng từ phía nhà đầu tư. Khi có một tương quan tích cực giữa sự quan tâm từ Google và giá cổ phiếu tăng, nhà đầu tư có thể có sự tin tưởng vào tiềm năng tăng trưởng của công ty, và điều này có thể thúc đẩy giá cổ phiếu tăng cao hơn.



Phản ánh tình hình thị trường và đối thủ: Biểu đồ cũng có thể giúp công ty theo dõi sự quan tâm đối với các đối thủ cạnh tranh. Nếu có sự gia tăng đáng kể trong số lượt tìm kiếm đối với đối thủ cùng lĩnh vực, công ty có thể cần xem xét lại chiến lược của họ để duy trì hoặc nâng cao tầm nhìn của mình trong thị trường.



Hình 4: Biểu đồ mối quan hệ giữa lượt truy cập google và giá cổ phiếu

### 3.4. Đánh giá mô hình dự đoán và đưa ra nhận xét

#### 1. Mô hình Dự đoán Google Search Trends:

- Ưu điểm:

Biểu đồ Google Search Trends có thể cung cấp thông tin quan trọng về sự quan tâm của khách hàng đối với sản phẩm hoặc dịch vụ của công ty (CTD).

Có thể theo dõi sự thay đổi trong số lượt tìm kiếm theo thời gian, giúp công ty đánh giá tác động của các sự kiện hoặc chiến dịch tiếp thị.

Đánh giá mối quan hệ giữa lượt tìm kiếm Google và giá cổ phiếu giúp công ty hiểu rõ hơn về tương quan giữa hai yếu tố này.

- Nhược điểm:

Biểu đồ Google Search Trends chỉ cho thấy sự quan tâm của khách hàng dựa trên số lượt tìm kiếm trên Google, không phản ánh được các yếu tố khác có thể ảnh

hưởng đến giá cổ phiếu như hoạt động tài chính của công ty, thông tin về sản phẩm mới, và sự biến động tổng thể của thị trường.

## 2. Mô hình Dự đoán Giá Cổ Phiếu CTD:

- Ưu điểm:

Biểu đồ về giá cổ phiếu CTD và đường trung bình động (Moving Average) cung cấp cái nhìn trực quan về xu hướng giá cổ phiếu và mức hỗ trợ/kháng cản chú ý.

Giúp nhà đầu tư và người quản lý dễ dàng theo dõi sự biến động của giá cổ phiếu theo thời gian.

- Nhược điểm:

Biểu đồ này không tích hợp sự quan tâm từ Google Search Trends, do đó không cho phép đánh giá tương quan trực tiếp giữa sự quan tâm của khách hàng và giá cổ phiếu.

Chỉ dự đoán giá cổ phiếu dựa trên dữ liệu quá khứ, không phản ánh được các yếu tố bên ngoài có thể ảnh hưởng đến giá cổ phiếu trong tương lai.

## 3. Nhận xét và Kết hợp thông tin từ cả hai mô hình:

Mô hình Dự đoán Google Search Trends có thể cung cấp thông tin về tình hình thị trường và tương quan giữa sự quan tâm của khách hàng và giá cổ phiếu.

Mô hình Dự đoán Giá Cổ Phiếu CTD cung cấp thông tin về xu hướng giá cổ phiếu và mức hỗ trợ/kháng.

Nhận xét cuối cùng là sự kết hợp thông tin từ cả hai mô hình có thể giúp công ty CTD có cái nhìn tổng quan và chính xác hơn về tình hình thị trường và sự quan tâm của khách hàng. Tuy nhiên, để đưa ra quyết định kế hoạch sản phẩm và tiếp thị hiệu quả, công ty cần xem xét thêm các yếu tố khác như hoạt động tài chính và chiến lược kinh doanh tổng thể.

### **3.5. Các khuyến nghị cho nhà đầu tư về cổ phiếu Công ty Cổ phần Xây dựng Cotecons**

#### **3.5.1. Phân tích chung ngành xây dựng**

Khi ngành xây dựng của Việt Nam phát triển nhanh chóng, CTD có triển vọng tích cực. Dự kiến nhu cầu về xây dựng các công trình cơ sở hạ tầng, nhà ở, và dự án

thương mại sẽ tiếp tục tăng trong giai đoạn 2022-2031. Điều này có thể là một lợi thế cho CTD với vai trò là một trong những công ty hàng đầu trong lĩnh vực xây dựng tại Việt Nam.

### ***3.5.2. Phân tích triển vọng cổ phiếu***

Về triển vọng cổ phiếu CTD, dự kiến doanh thu và lợi nhuận sau thuế (LNST) của CTD trong năm 2023 có thể tăng lên. Tuy nhiên, điều này cũng phụ thuộc vào thị trường xây dựng tổng thể và khả năng quản lý tài chính của công ty. Cần theo dõi sự biến động của giá cổ phiếu CTD để đưa ra quyết định đầu tư.

Hiện tại, CTD có mức giá cổ phiếu có thể cao hơn so với một số chỉ số như P/E. Điều này có thể làm cho việc đầu tư ngắn hạn trở nên khó khăn hơn. Tuy nhiên, CTD có thể được xem xét cho mục đích đầu tư dài hạn do vị thế và thị phần dẫn đầu trong ngành xây dựng.

### ***3.5.3. Một số giải pháp***

#### *Dưới góc nhìn của nhà đầu tư*

Nhà đầu tư cần quan sát sự biến động của giá cổ phiếu CTD và nắm bắt cơ hội mua vào trong trường hợp giá cổ phiếu giảm. Tuy giảm giá có thể tạo cơ hội mua vào, nhưng cần xem xét kỹ thuật phân tích cổ phiếu và tình hình tài chính của công ty trước khi quyết định mua.

Để tận dụng tiềm năng tăng giá cổ phiếu trong tương lai, nhà đầu tư có thể xem xét mua cổ phiếu CTD với giá thấp hơn và bán khi giá cao hơn. Tuy nhiên, cần đánh giá rủi ro và thực hiện phân tích kỹ thuật và tài chính trước khi đưa ra quyết định mua hoặc bán.

### *Dưới góc nhìn của nhà quản trị*

Sự biến động của giá cổ phiếu CTD có thể ảnh hưởng đến tài chính của công ty. Nhà quản trị cần theo dõi tình hình và đưa ra biện pháp để quản lý tài chính trong trường hợp giá cổ phiếu giảm. Điều này có thể bao gồm cắt giảm chi phí hoặc tối ưu hóa hoạt động kinh doanh.

Để cải thiện hiệu suất công ty, có thể cần xem xét tái cơ cấu tổ chức, tối ưu hóa quy trình sản xuất, hoặc tìm kiếm cơ hội hợp tác mới. CTD cũng cần quản lý rủi ro và sử dụng công cụ tài chính để bảo vệ khỏi biến động giá cổ phiếu.

Duy trì sự phát triển bền vững và quản lý rủi ro là quan trọng để duy trì và tăng giá trị cho cổ đông của CTD.

Những khuyến nghị và phân tích này nhằm giúp nhà đầu tư và nhà quản trị hiểu rõ hơn về triển vọng và tiềm năng của Công ty Cổ phần Xây dựng Coteccons (CTD) và đưa ra quyết định đầu tư và quản lý hiệu quả. Tuy nhiên, cần lưu ý rằng đầu tư luôn có rủi ro và nên được thực hiện sau khi xem xét cẩn trọng tình hình và thông tin cụ thể của công ty và thị trường.

## PHẦN KẾT LUẬN

Trong đề tài này, chúng tôi đã tiến hành nghiên cứu và phân tích về việc sử dụng các mô hình học máy, đặc biệt là mô hình Dự đoán Google Search Trends, để dự đoán giá cổ phiếu của Công ty Cổ phần Xây dựng Coteccons (CTD). Chúng tôi đã thu thập dữ liệu lịch sử về giá cổ phiếu, các chỉ số tài chính liên quan và các yếu tố thị trường để xây dựng các mô hình dự đoán.

Từ việc sử dụng các mô hình trên, chúng tôi đã đưa ra các biểu đồ để hỗ trợ việc quan sát sự biến động giá của CTD trong quá khứ và dự đoán giá cổ phiếu của công ty trong tương lai. Những dự đoán này có thể hỗ trợ nhà đầu tư trong việc ra quyết định đầu tư, tối ưu hoá lợi nhuận và giảm thiểu rủi ro. Đồng thời, nó cũng hỗ trợ việc lập kế hoạch đầu tư dài hạn và xác định các mục tiêu tài chính cá nhân và tổ chức.

Các kết quả của đề tài này cho thấy tiềm năng lớn của Machine Learning trong việc dự đoán giá cổ phiếu của Công ty Cổ phần Xây dựng Coteccons (CTD). Tuy nhiên, chúng tôi muốn nhấn mạnh rằng việc dự đoán giá cổ phiếu là một vấn đề phức tạp và phụ thuộc vào nhiều yếu tố khách quan và không chắc chắn.

Khả năng ứng dụng của các mô hình Machine Learning sẽ phụ thuộc vào việc cải thiện dữ liệu đầu vào, quá trình tiền xử lý dữ liệu và sự tối ưu hóa của mô hình. Ngoài ra, việc theo dõi và cập nhật mô hình liên tục theo thời gian cũng rất quan trọng để đảm bảo tính hiệu quả của nó.

Sử dụng Machine Learning để dự đoán giá cổ phiếu là một lĩnh vực nghiên cứu hứa hẹn và có thể cung cấp giá trị cho các nhà đầu tư và nhà quản lý tài sản trong việc ra quyết định đầu tư. Tuy nhiên, việc áp dụng cần được thực hiện cẩn thận và liên tục theo dõi và cải thiện để đảm bảo tính chính xác và tin cậy

