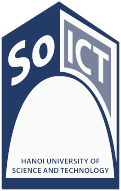
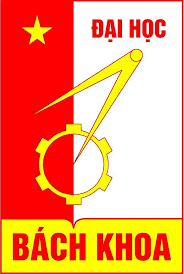
**TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**

**VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**



**Đồ án tốt nghiệp**

**Đề tài: Dự đoán xu hướng thị trường chứng khoán dựa trên các mô hình deep learning**

| Giáo viên hướng dẫn: | Trịnh Anh Phúc |
| --- | --- |
| Họ và tên: | Lê Trọng Đạt |
| MSSV: | 20183493 |
| Lớp: | CTTN-CNTT K63 |

**Hà Nội, 2022**

**LỜI CAM ĐOAN**

Tôi xin cam đoan rằng, luận văn **“Dự đoán xu hướng thị trường chứng khoán dựa trên các mô hình deep learning"** là công trình nghiên cứu dưới sự hướng dẫn của TS.Trịnh Anh Phúc, xuất phát từ nhu cầu thực tiễn và nguyện vọng của bản thân.

Ngoại trừ các kết quả nghiên cứu từ các công trình đã ghi rõ trong luận văn, các nội dung trình bày trong luận văn này là do chính tôi nghiên cứu do chính tôi thực hiện và kết quả của luận văn chưa từng công bố trước đây dưới bất kỳ hình thức nào.

Tôi xin chịu hoàn toàn trách nhiệm về lời cam đoan của mình. Nếu có gì sai trái, tôi xin chịu mọi hình thức kỷ luật theo quy định.

*Hà Nội, tháng … năm 2022*

Tác giả

**Lê Trọng Đạt**

Mục Lục

[**Chương 1: Giới thiệu**](#_heading=h.mzgmsl4hynqq) **1**

[**1.2 Các nghiên cứu liên quan**](#_heading=h.2vu2ogjsr1ie) **3**

[**1.5 Bố cục luận văn**](#_heading=h.wk44u2ud92po) **4**

[**Chương 3: Nội dung**](#_heading=h.tzpww9cuog3x) **4**

[Tiền xử lý dữ liệu](#_heading=h.2et92p0) 4

[LSTM](#_heading=h.tyjcwt) 6

[LSTM và CNN](#_heading=h.3dy6vkm) 7

[Transformer Encoder và Time Embedding](#_heading=h.1t3h5sf) 10

[Phần 4: Kết quả thực nghiệm](#_heading=h.4d34og8) **12**

[Lựa chọn tham số](#_heading=h.2s8eyo1) 12

[So sánh các thuật toán](#_heading=h.17dp8vu) 13

[Phần 5: Kết luận](#_heading=h.3rdcrjn) **15**

[Tài liệu tham khảo](#_heading=h.26in1rg) **16**

# Chương 1: Giới thiệu

**1.1 Đặt vấn đề**

Thị trường chứng khoán chính là thước đo sức mạnh nền kinh tế của nhiều nước trên thế giới và là công cụ phản ánh sự thay đổi của nền kinh tế trong tương lai. 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 nền kinh tế trong tương lai gần. Việc quan sát xu hướng phát triển của thị trường chứng khoán sẽ giúp các nhà kinh tế, ngân hàng, doanh nghiệp,... chuẩn bị những thứ cần thiết trước sự biến động của nền kinh tế.

Thị trường chứng khoán luôn thu hút các nhà đầu tư do khả năng sinh lời cao nhưng cũng đi kèm nhiều rủi ro. Lợi nhuận của các nhà đầu tư chủ yếu được tạo ra từ việc dự đoán trước sự biến động của thị trường từ đó có các hành động mua vào hay bán phù hợp. Việc lựa chọn mua hay bán cổ phiếu nào phụ thuộc vào kỳ vọng phát triển của một công ty trong tương lai. Do đó các nhà đầu tư khi tham gia thị trường chứng khoán cần phải dự đoán được xu hướng thị trường chứng khoán trong ngắn hạn hoặc dài hạn để giảm thiểu rủi ro và đưa ra các quyết định đúng đắn.

Mặt khác, trong thời kỳ dịch bệnh, sản xuất kinh tế bị đình trệ nên số lượng các nhà đầu tư f0 đầu tư vào chứng khoán tăng lên nhanh chóng. Đặc biệt là với sự nổi lên của tiền ảo, đây được xem là một kênh thu về lợi nhuận lớn nhưng kèm theo rủi ro cao. Với nhiều nhà đầu tư mới tham gia vào thị trường, kiến thức và kinh nghiệm về thị trường còn ít nên thường đưa ra các quyết định sai lầm. Khi đó, các công cụ phân tích và dự đoán xu hướng thị trường chứng khoán càng trở nên quan trọng để giúp các nhà đầu tư mới giảm thiểu rủi ro và sớm hòa nhập với thị trường.

Dự đoán về xu hướng của thị trường chứng khoán là một nhiệm vụ đầy thách thức do tính chất phi tuyến tính và đặc trưng bởi một hành vi đám đông của thị trường. Các yếu tố tạo nên sự biến động trong lĩnh vực này rất phức tạp và bắt nguồn từ nhiều nguyên nhân với các bản chất khác nhau. Từ các quyết định kinh tế, chính trị,·hiệu quả hoạt động cụ thể của công ty cho đến những nguyên nhân không rõ ràng như tin đồn, quyết định của các nhà đầu tư nổi tiếng, các quy luật của thị trường,... khiến cho hầu như không thể tính đến tất cả các yếu tố liên quan khi đưa ra quyết định giao dịch. Rất khó để phân tích các yếu tố này theo cách thủ công, vì vậy sẽ tốt hơn nếu có các công cụ hỗ trợ phân tích những dữ liệu này.

Ngày nay, cùng với sự phát triển mạnh mẽ của các mô hình học sâu (Deep Learning) trong nhiều lĩnh vực, ứng dụng của học sâu trong tài chính ngày càng được chú ý và phát triển. Việc áp dụng các mô hình học sâu để dự báo chứng khoán cũng rất phổ biến trên thế giới, không chỉ trong học thuật mà còn trong thực tiễn. Khác với các mô hình thống kê và kinh tế lượng truyền thống, các mô hình deep learning có thể mô tả các yếu tố ảnh hưởng phức tạp và phi tuyến. Các mô hình này đã trở thành công cụ hữu hiệu trong việc hỗ trợ các nhà đầu tư đưa ra các quyết định.

Luận văn này sẽ khám phá một ứng dụng của học sâu trong dự đoán xu hướng chứng khoán. Cụ thể, tôi sẽ sử dụng các mô hình học sâu để dự báo xu hướng thị trường chứng khoán trong ngắn hạn. Luận văn đã đạt được một số kết quả khả quan trong nghiên cứu kỹ thuật học sâu và so sánh được kết quả của việc áp dụng các mô hình học sâu vào các lý thuyết phân tích chứng khoán cổ điển. Tuy nhiên, luận văn không thể tránh khỏi những thiếu sót, rất mong nhận được những ý kiến đóng góp, nhận xét để tôi tiếp tục hoàn thiện và nâng cao kết quả nghiên cứu.

**1.2 Các nghiên cứu liên quan**

Nhiều bài báo sử dụng thành công các phương pháp học sâu trong việc phát triển các mô hình dự đoán xu hướng thị trường chứng khoán. Có thể chia các phương pháp này thành 2 hướng nghiên cứu là các mô hình chuỗi thời gian và học tăng cường. Ví dụ, Bao và cộng sự [11] đã sử dụng các phép biến đổi wavelet để loại bỏ nhiễu khỏi chuỗi giá chứng khoán trước khi đưa chúng vào một lớp autoencoder và lớp Long-short Term Memory (LSTM) để đưa ra dự đoán giá trong một ngày. M và cộng sự [12] đã so sánh Convolutional Neural Network (CNN) với Recurrent Neural Network (RNN) về dự đoán giá cổ phiếu của các công ty trong lĩnh vực công nghệ thông tin. Trong thử nghiệm của họ, mạng CNN cho kết quả tốt hơn so với RNN và LSTM. Sự khác biệt về hiệu suất được cho là do CNN không dựa trên dữ liệu lịch sử như trường hợp của các mô hình tuần tự. Ngoài ra, nhiều nhà nghiên cứu tập trung vào việc sử dụng các kỹ thuật Học tăng cường để giải quyết vấn đề giao dịch theo thuật toán. Ví dụ, Moody và Saell [14] đã giới thiệu một thuật toán học tăng cường để xác định các chính sách đầu tư sinh lời mà không cần xây dựng mô hình dự báo, và Dempster và Leemans [15] đã sử dụng học tăng cường để để giao dịch trên thị trường ngoại hối.

**1.3 Mục tiêu nghiên cứu**

Luận văn này tập trung nghiên cứu giải quyết bài toán dự đoán giá cổ phiếu bằng các mô hình deep learning áp dụng trên các lý thuyết phân tích cổ điển. Với bộ dữ liệu thu thập được từ các trang web về tài chính trên thế giới sau đó được tiền xử lý và áp dụng các phương pháp học máy và học sâu để đưa ra được kết quả dự đoán từ đó đề xuất mô hình tối ưu nhất cũng như lý thuyết phù hợp nhất với mô hình deep learning.

**1.4 Đối tượng nghiên cứu**

Đối tượng nghiên cứu của luận văn là các mô hình deep learning và ứng dụng của chúng vào các lý thuyết phân tích chứng khoán cổ điển.

**1.5 Bố cục luận văn**

Bố cục luận văn được chia thành 5 phần, bao gồm những nội dung như sau:

* Chương 1: Giới thiệu tổng quan về đề tài nghiên cứu, phát biểu bài toán, mục tiêu, đối tượng nghiên cứu
* Chương 2: Giới thiệu những kiến thức nền tảng. Cụ thể, chương này sẽ giới thiệu các phương pháp phân tích thị trường chứng khoán và các kiến thức nền tảng của học sâu (mạng nơ-ron nhân tạo, kỹ thuật lan truyền ngược), các mô hình học sâu sẽ được sử dụng.
* Chương 3: Mô tả các giải pháp và ứng dụng chúng vào dự đoán giá cổ phiếu. Sau đó là đánh giá hiệu quả của các mô hình đề xuất.
* Chương 4: Thực nghiệm, các bước tiền xử lý dữ liệu, huấn luyện mô hình, các phương pháp đánh giá mô hình và so sánh kết quả của từng mô hình, từ đó rút ra được các nhận xét về các mô hình đã cài đặt.
* Chương 5: Tóm tắt các ý chính của luận văn, đưa ra kết luận về ưu, nhược điểm của các mô hình từ đó đề xuất mô hình phù hợp.

# Chương 2: Cơ sở lý thuyết

## 2.1 Những vấn đề cơ bản của thị trường chứng khoán

### 2.2.1 Khái niệm chứng khoán

Chứng khoán là các giấy tờ có giá, có khả năng chuyển đổi, chuyển nhượng nhằm xác nhận quyền sở hữu, quan hệ vay nợ giữa người nắm giữ nó và chủ thể phát hành ra nó. Đối với mỗi loại chứng khoán thường có các tính chất sau:

* Tính thanh khoản của một chứng khoán là khả năng chuyển đổi giữa chứng khoán đó sang tiền mặt. Tính thanh khoản của chứng khoán thể hiện qua việc chứng khoán đó được mua bán, trao đổi trên thị trường.
* Tính sinh lời: Thu nhập của nhà đầu tư được sinh ra từ việc tăng giá chứng khoán trên thị trường, hay các khoản tiền lãi được trả hàng năm.
* Tính rủi ro: Đây là đặc trưng cơ bản của chứng khoán.Trong quá trình trao đổi, mua đi bán lại, giá của chứng khoán bị giảm hoặc mất hoàn toàn ta gọi là rủi ro.

### 2.2.2 Phân loại chứng khoán

* Cổ phiếu là giấy chứng nhận sự góp vốn kinh doanh vào các công ty cổ phần. Người nắm giữ cổ phiếu là cổ đông đồng thời là chủ sở hữu công ty phát hành. Giá trị ban đầu ghi trên cổ phiếu là mệnh giá của cổ phiếu. Mệnh giá là giá trị danh nghĩa. Số tiền nhận được từ khoản góp vốn gọi là cổ tức. Giá cổ phiếu giao động qua các phiên giao dịch trên thị trường chứng khoán và tách rời so với mệnh giá.
* Trái phiếu là giấy chứng nhận nợ dài hạn do các tổ chức phát hành cam kết với người vay sẽ chi trả lợi tức định kỳ và hoàn lại vốn khi đáo hạn. Mệnh giá ghi trên trái phiếu là số vốn gốc được trả cho trái chủ tại thời điểm đáo hạn.

Ngoài hai loại chứng khoán cơ bản trên, các doanh nghiệp đã phát hành ra nhiều loại chứng khoán khác bằng cách thêm hoặc bớt quyền lợi của nhà đầu tư để thu hút nguồn vốn và giảm thiểu rủi ro cho nhà đầu tư.

### 2.2.3 Khái niệm thị trường chứng khoán

Quá trình sản xuất ngày nay bao gồm các yếu tố: lao động, vốn, đất đai và năng lực kinh doanh. Trong đó vốn là yếu tố quan trọng hàng đầu vì nó là biểu hiện cho đối tượng lao động và tư liệu lao động. Trong khi những người có vốn thì không có cơ hội đưa vào sản xuất còn những người có cơ hội sản xuất lại thiếu hụt vốn, do đó cần một cơ chế để luân chuyển tiền nhàn rỗi trong xã hội vào quá trình sản xuất. Hình thức đơn giản nhất là quan hệ vay mượn trực tiếp giữa các tầng lớp dân cư hoặc quan hệ tín dụng giữa các doanh nghiệp. Song hình thức này quy mô không lớn và chủ yếu diễn ra trên cơ sở tín nhiệm giữa 2 chủ thể. Hình thức thứ 2 là qua các tổ chức tài chính trung gian như ngân hàng, tổ chức tín dụng,... Các tổ chức này ngày càng trở thành một phần không thể thiếu của nền kinh tế. Tuy nhiên xét về góc độ của doanh nghiệp hay nhà đầu tư thì hình thức thông qua các tổ chức tài chính trung gian lại không hiệu quả do chi phí luân chuyển vốn qua các trung gian tài chính lớn. Do đó, thị trường chứng khoán ra đời cho phép các doanh nghiệp phát hành các chứng từ có giá để huy động vốn trực tiếp từ các nhà đầu tư. Thị trường chứng khoán là nơi mua bán trao đổi các loại chứng khoán.

### 2.2.4 Vai trò của thị trường chứng khoán

Thứ nhất, thị trường chứng khoán với việc tạo ra các công cụ có tính thanh khoản cao, phân phối vốn, chuyển thời hạn của vốn phù hợp với yêu cầu phát triển kinh tế. Việc huy động vốn trên thị trường chứng khoán có thể làm tăng vốn tự có của các công ty và giúp họ tránh các khoản vay có chi phí cao cũng như sự kiểm soát chặt chẽ của các ngân hàng thương mại. Sự tồn tại của các thị trường chứng khoán cũng là yếu tố quyết định để thu hút vốn đầu tư nước ngoài. Đây là yếu tố đảm bảo cho sự phân bổ có hiệu quả các nguồn lực trong một quốc gia cũng như trong phạm vi quốc tế.

Thứ hai, thị trường chứng khoán góp phần thực hiện tái phân phối công bằng hơn, giải tỏa sự tập trung quyền lực kinh tế của các tập đoàn, huy động tầng lớp trung lưu trong xã hội, tăng cường sự giám sát của xã hội đối với quá trình phân phối. Từ đó, tạo ra sự cạnh tranh công bằng hơn qua đó thúc đẩy phát triển kinh tế, tạo ra hiệu quả, tiến tới xây dựng một xã hội công bằng và dân chủ.

Thứ ba, thị trường chứng khoán tạo điều kiện cho việc tách biệt giữa sở hữu và quản lý doanh nghiệp. Khi quy mô của doanh nghiệp tăng lên thì nhu cầu về quản lý chuyên trách cũng tăng theo. Thị trường chứng khoán tạo điều kiện cho tiết kiệm vốn và chất xám thúc đẩy quá trình cổ phần hoá doanh nghiệp nhà nước, giảm tiêu cực trong quản lý, tạo điều kiện kết hợp hài hoà lợi ích của chủ sở hữu, nhà quản lý và những người làm công.

Thứ tư, hiệu quả của quốc tế hoá thị trường chứng khoán. Việc mở cửa thị trường chứng khoán làm tăng tính cạnh tranh trên thị trường quốc tế, cho phép các công ty có thể huy động nguồn vốn rẻ hơn, tăng cường đầu tư từ bên ngoài, đồng thời tăng khả năng cạnh tranh quốc tế và mở rộng cơ hội kinh doanh cho các công ty trong nước.

Thứ năm, thị trường chứng khoán tạo điều kiện cho Chính phủ huy động các nguồn tài chính tiền tệ. Đồng thời, thị trường chứng khoán cũng cung cấp một dự báo tuyệt vời về các chu kỳ kinh doanh trong tương lai giúp Chính phủ và các công ty đánh giá kế hoạch đầu tư cũng như việc phân bổ các nguồn lực của họ. Thị trường chứng khoán cũng tạo điều kiện tái cấu trúc nền kinh tế.

Ngoài những tác động tích cực trên, thị trường chứng khoán cũng có những tác động tiêu cực nhất định như hiện tượng đầu cơ, hiện tượng xung đột quyền lực, bong bóng giá,…làm giảm thiệt hại cho quyền lợi của các cổ đông thiểu số, việc mua bán nội gián, thao túng thị trường làm nản lòng các nhà đầu tư tác động tiêu cực tới tiết kiệm và đầu tư. Nhiệm vụ của các nhà quản lý thị trường là giảm thiểu các tiêu cực của thị trường nhằm bảo vệ quyền lợi của các nhà đầu tư và đảm bảo tính hiệu quả của thị trường.

Như vậy, vai trò của thị trường chứng khoán được thể hiện ở nhiều khía cạnh khác nhau. Song để vai trò đó thật sự phát huy được hiệu quả thì phụ thuộc đáng kể vào các chủ thể tham gia vào thị trường và sự quản lý của nhà nước.

## 2.2 Các phương pháp phân tích thị trường

Ngay từ khi xuất hiện thị trường chứng khoán, nhiều nhà đầu tư đã tìm cách để phân tích xu hướng của thị trường một cách khoa học tạo ra các trường phái lý luận khác nhau. Mỗi trường phái đều có những ưu nhược điểm riêng.

### 2.2.1 Trường phái thị trường hiệu quả

Trường phái này cho rằng giá cả của các công cụ tài chính phản ánh thông tin của công ty sở hữu. Giá trị thực của công ty cũng được phản ánh qua giá cả.

*Thị trường tài chính hiệu quả* (efficient financial market) là thị trường tài chính trong đó giá hiện tại của tài sản tài chính phản ánh đầy đủ mọi thông tin có liên quan. Ba mức độ hiệu quả của thị trường:

* Hiệu quả yếu: giá cả hiện tại chỉ phản ánh thông tin trong quá khứ
* Hiệu quả trung bình: giá cả hiện tại phản ánh thông tin được công bố ở hiện tại
* Hiệu quả mạnh: giá cả hiện tại phản ánh đầy đủ thông tin kể cả thông tin quá khứ, thông tin được công bố, thông tin chỉ người trong công ty mới biết

Ví dụ: Cổ phiếu MU giao động mạnh với các thông tin về ronaldo. Giả sử vào sáng nay ronaldo bị chấn thương trong phòng tập. Với thị trường hiệu quả yếu, cổ phiếu MU không giảm mặc dù mọi người biết Ronaldo gãy chân rồi - giá cổ phiếu không phàn ánh được thông tin. Với thị trường hiệu quả trung bình, cổ phiếu MU giảm khi tin chấn thương được đăng trên các báo. Với hiệu quả mạnh, ngay khi Ronaldo bị chấn thương thì cổ phiếu giảm dù chưa ai công bố thông tin này.

Thị trường hiệu quả yếu như một canh bạc trong đó có kẻ ăn gian chưa bị phát hiện. Thị trường cần hiệu quả mạnh để tạo sự công bằng, minh bạch. Công bằng khi giá cả biểu hiện theo sự biến động của thị trường, minh bạch vì mọi thông tin không thể được che giấu.

Lý thuyết thị trường hiệu quả cho rằng những nhà tài chính giỏi là những người có nhiều thông tin nhất, đặc biệt là các thông tin mật. Họ cho rằng rủi ro của các công cụ tài chính đã được tính vào giá cả và mỗi cổ phiếu sẽ có biên độ giao động beta. Điều quan trọng là mua vào lúc beta thấp nhất và bán ra lúc beta cao nhất. Đây giống như hình thức đầu cơ lướt sóng.

### 2.2.2 Trường phái dựa trên giá trị

Trường phái giá trị được phát minh đầu tiên bởi Benjamin Graham và được Warren Buffett sử dụng để trở thành tỷ phú thuộc top giàu nhất thế giới. Trường phái giá trị cho rằng giá trị của một công cụ tài chính phản ánh giá trị thực của công ty nó thuộc về. Giá cả của các công cụ tài chính sẽ dần dần được thị trường điều chỉnh về giá trị thực của chúng. Điều này khá giống quy luật giá trị trong học thuyết mac-lenin.

Những nhà tài chính theo học thuyết này sẽ tính toán giá trị của công ty dựa vào báo cáo tài chính của các công ty đó và định giá chúng một cách logic. Họ cố gắng tìm những cổ phiếu được bán với giá thấp hơn giá trị thực và chờ trong 1 thời gian dài để chúng tự điều chỉnh.

Trường phái giá trị có vẻ đúng trong lý thuyết nhưng trong thực tế lại khó tránh được các rủi ro. Các cổ phiếu giá rẻ hơn giá trị thường khó có chỗ đứng nên cần một lượng lớn đầu tư và rất cần năng lực của nhà quản trị. Nếu không công ty khó có thể tránh được các yếu tố rủi ro và phá sản trước khi được định giá lại.

### 2.2.3 trường phái dựa trên biểu đồ

Các nhà tài chính trường phái này sẽ cố gắng đưa ra giải thích cho hiện tại và dự đoán tương lai từ việc quan sát các số liệu, đồ thị trong quá khứ. Họ cho rằng giá cả và các thông số của cổ phiếu đã mang đầy đủ thông tin và quy luật biến đổi giá trong quá khứ và hiện tại phản ánh giá cả của cổ phiếu trong tương lai gần. Trường phái này thường đưa ra các quy luật của giá cả như: cổ phiếu thường tăng vào tháng 1 giảm vào tháng 12.

## 2.3 Deep learning và phân tích thị trường

Cách mạng Công nghiệp 4.0 mang đến những thay đổi toàn diện trong mọi mặt của cuộc sống. “Thế giới vận hành bằng công nghệ” trở thành xu thế tất yếu, là mục tiêu tranh đua cho mọi nền kinh tế, mọi doanh nghiệp cùng hướng đến. Trong cuộc cách mạng Công nghiệp lần thứ IV, Trí tuệ nhân tạo (A.I) và deep learning là những thành phần quan trọng, nổi bật, đã và đang được nghiên cứu, ứng dụng rộng rãi trong nhiều lĩnh vực khác nhau trong đó có ngành ngân hàng nói chung và lĩnh vực công nghệ tài chính nói riêng, tạo ra những bứt phá mạnh mẽ về tính hiệu quả, kịp thời, chính xác trong công tác giám sát, dự báo và hỗ trợ ra quyết định.

2.3.1 Deep learning

Học sâu là một kỹ thuật tiên tiến của học máy dựa trên các thuật toán mạng nơ-ron nhân tạo. Là một nhánh đầy triển vọng của trí tuệ nhân tạo, học sâu đã thu hút được sự chú ý lớn từ cộng đồng công nghệ. So với các kỹ thuật học máy thông thường như máy vectơ hỗ trợ (SVM) và k-láng giềng gần nhất (kNN), các thuật toán học sâu sở hữu lợi thế của việc học tính năng không giám sát, khả năng tổng quát hóa mạnh mẽ và khả năng đào tạo mạnh mẽ với dữ liệu lớn. Hiện tại, học sâu đã được ứng dụng toàn diện trong các nhiệm vụ phân loại và dự đoán, thị giác máy tính, xử lý hình ảnh và âm thanh. Mặc dù DL được phát triển trong lĩnh vực khoa học máy tính, nhưng các ứng dụng của nó đã thâm nhập vào các lĩnh vực đa dạng như y học, khoa học thần kinh, vật lý và thiên văn học, tài chính và ngân hàng.

Học sâu được lấy cảm hứng từ mạng nơ-ron sinh học bao gồm nhiều nơ-ron liên kết với nhau. Mạng nơ-ron học sâu, hoặc mạng nơ-ron nhân tạo, cố gắng mô phỏng hệ cách hệ thần kinh hoạt động thông qua sự kết hợp của dữ liệu đầu vào, các trọng số và các đại lượng hiệu chỉnh. Các yếu tố này làm việc cùng nhau để nhận biết, phân loại và mô tả chính xác các đối tượng trong dữ liệu.

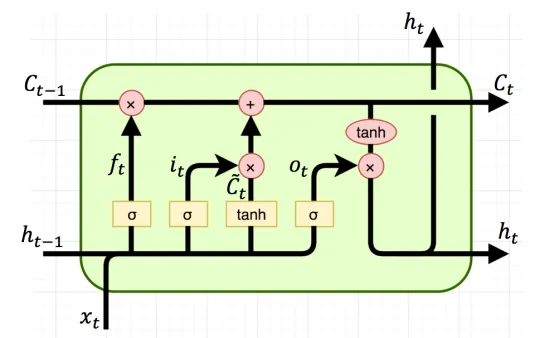
Mạng nơ ron bao gồm nhiều lớp được kết nối với nhau, mỗi lớp xây dựng dựa trên lớp trước đó để tinh chỉnh và tối ưu hóa dự đoán hoặc phân loại. Quá trình tính toán thông qua mạng này được gọi là quá trình chuyển tiếp. Lớp đầu vào là nơi mô hình học sâu nhập dữ liệu để xử lý và lớp đầu ra là nơi đưa ra dự đoán hoặc phân loại cuối cùng.

Trong phần tiếp theo tôi sẽ trình bày các mô hình học sâu được nghiên cứu trong luận văn: Học sâu có giám sát (LSTM, BiLSTM ) và mô hình Transformer.

2.3.1 Mạng nơ-ron ngắn dài LSTM (Long Short Term Memory)

Mạng nơ-ron ngắn dài LSTM là một cải tiến của mạng hồi quy, cho phép học được các thông tin ở quá khứ xa. Nó có khả năng xử lý vấn đề triệt tiêu gradient mà mạng hồi quy phải đối mặt. Giả sử khi xem video bạn nhớ cảnh trước đó hoặc khi đọc sách, bạn biết điều gì đã xảy ra ở chương trước. Tương tự như vậy, cách mạng hồi quy hoạt động, chúng ghi nhớ thông tin trước đó và sử dụng nó để xử lý đầu vào hiện tại. Thiếu sót của mạng hồi quy là không thể nhớ các thông tin lâu dài do vấn đề triệt tiêu gradient. LSTM được thiết kế để tránh được vấn đề phụ thuộc xa (long-term dependency). Việc nhớ thông tin trong suốt thời gian dài là đặc tính mặc định của chúng, chứ ta không cần phải huấn luyện nó để có thể nhớ được. Tức là ngay nội tại của nó đã có thể ghi nhớ được mà không cần bất kì can thiệp nào.

Mọi mạng hồi quy đều có dạng là một chuỗi các mô-đun lặp đi lặp lại của mạng nơ-ron. Với mạng RNN chuẩn, các mô-dun này có cấu trúc rất đơn giản, thường là một tầng tanh. LSTM cũng có kiến trúc dạng chuỗi như vậy, nhưng các mô-đun trong nó có cấu trúc phức tạp hơn. Thay vì chỉ có một tầng mạng nơ-ron, chúng có tới 4 tầng tương tác với nhau một cách rất đặc biệt.

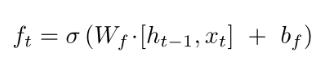


Hình: Mô hình một tế bào LSTM

Chìa khóa của LSTM là trạng thái tế bào (cell state) - chính đường chạy thông ngang từ đến . Trạng thái tế bào là một dạng như băng truyền chạy xuyên suốt tất cả các nút mạng và chỉ tương tác nhỏ để cập nhật thông tin. Nhờ có trạng thái tế bào mà các thông tin có thể truyền đi rất xa trong mạng. LSTM có khả năng bỏ đi hoặc thêm vào các thông tin cần thiết cho trạng thái tế bào, chúng được điều chỉnh cẩn thận bởi các nhóm được gọi là cổng (gate).

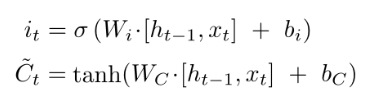
Cổng quên (forget gate - ):

Sử dụng để quyết định xem có bao nhiêu thông tin cũ của được giữ lại. Nó sẽ biến đổi tuyến tính đầu vào là và rồi đưa qua hàm kích hoạt sigmoid và đưa ra kết quả và số trong khoảng [0, 1] tương ứng với mỗi số trong trạng thái tế bào . Đầu ra là 1 tương ứng với giữ lại toàn bộ thông tin, đầu ra 0 tương ứng bỏ đi toàn bộ thông tin.

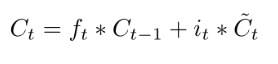


Cổng đầu vào (input gate - ):

Sử dụng để quyết định xem có bao nhiêu thông tin mới được cập nhật vào trạng thái tế bào. Đầu tiên là sử dụng một tầng sigmoid để quyết định giá trị nào ta sẽ cập nhập. Tiếp theo là một tầng tanh tạo ra một véc-tơ cho giá trị mới ​nhằm thêm vào cho trạng thái. Trong bước tiếp theo, ta sẽ kết hợp 2 giá trị đó lại để tạo ra một cập nhập cho trạng thái.

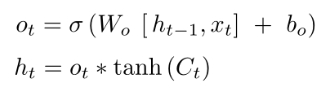


Tới đây ta cần kết hợp cổng quên và cổng đầu vào, để tạo ra trạng thái thực sự:



Cổng kết quả (output gate - ):

Giá trị đầu ra sẽ dựa vào trạng thái tế bào, nhưng sẽ được tiếp tục sàng lọc. Tương tự hai cổng trước, ta cũng dùng hàm sigmoid để ép giá trị kết hợp của và để quyết định điều này. Giá trị của trạng thái hiện tại cũng được ép trong khoảng -1 đến 1 bằng hàm tanh trước khi nhân với lại cổng ra:

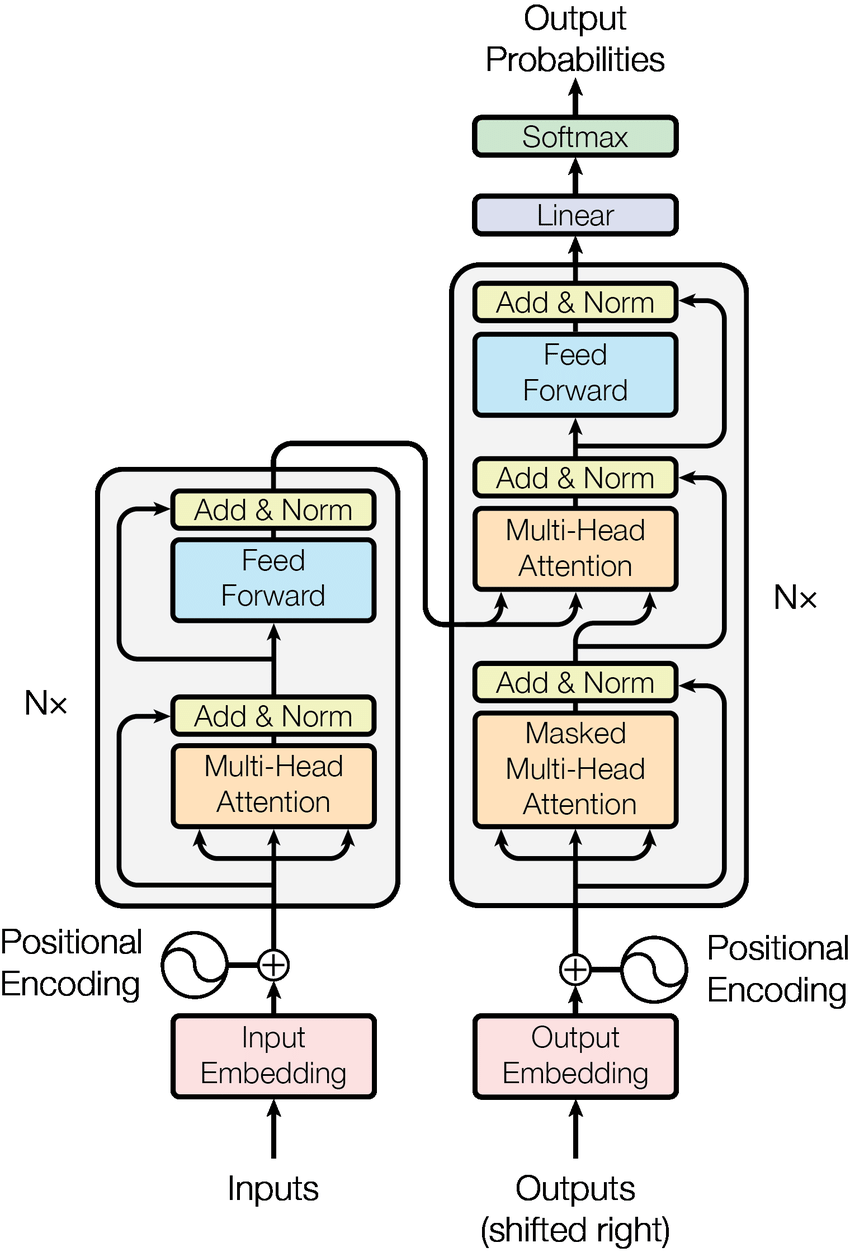


Mạng LSTM sau khi truyền tiến nhƣ trên là đã có thể dự đoán kết quả của một điểm dữ liệu nào đó. Ta có thể áp dụng kỹ thuật lan truyền ngƣợc để cập nhật trọng số cho phù hợp tƣơng tự nhƣ các phần trên, bằng cách tính đạo hàm từng phần ngƣợc lại từ đầu ra tới đầu vào.

2.3.2 Transformer Encoder

Transformer là một mô hình học sâu được thiết kế để phục vụ giải quyết nhiều bài toán trong xử lý ngôn ngữ và tiếng nói, ví dụ như bài toán dịch tự động, bài toán sinh ngôn ngữ, phân loại, nhận dạng thực thể, nhận dạng tiếng nói, chuyển văn bản thành tiếng nói. Tuy nhiên, khác với RNNs, Transformer không xử lý các phần tử trong một chuỗi một cách tuần tự mà sử dụng cơ chế self-attention. Nếu dữ liệu đầu vào là một câu ngôn ngữ tự nhiên, Transformer không cần phải xử lý phần đầu câu trước rồi mới tới phần cuối câu. Do tính năng này, Transformer có thể tận dụng khả năng tính toán song song của GPU và giảm thời gian xử lý đáng kể.

Giống như những mô hình dịch máy khác, kiến trúc tổng quan của mô hình transformer bao gồm 2 phần lớn là encoder và decoder. Encoder dùng để học vector biểu của câu với mong muốn rằng vector này mang thông tin hoàn hảo của câu đó. Decoder thực hiện chức năng chuyển vector biểu diễn kia thành ngôn ngữ đích.



Encoder của mô hình transformer có thể bao gồm nhiều encoder layer tượng tự nhau. Mỗi encoder layer của transformer lại bao gồm 2 thành phần chính là multi head attention và feedforward network, ngoài ra còn có cả skip connection và normalization layer. Encoder đầu tiên sẽ nhận ma trận biểu diễn của các từ đã được cộng với thông tin vị trí thông qua positional encoding. Sau đó, ma trận này sẽ được xử lý bởi Multi Head Attention. Multi Head Attention thật chất là self-attention, nhưng mà để mô hình có thể có chú ý nhiều pattern khác nhau, tác giả đơn giản là sử dụng nhiều self-attention.

Self-Attention là cơ chế giúp encoder nhìn vào các từ khác trong lúc mã hóa một từ cụ thể, vì vậy,Transformers có thể hiểu được sự liên quan giữa các từ trong một câu, kể cả khi chúng có khoảng cách xa. Các decoder cũng có kiến trúc giống như vậy nhưng giữa chúng có một lớp attention để nó có thể tập trung vào các phần liên quan của đầu vào.

Đầu tiên, với môi từ chúng ta cần tạo ra 3 vector: query, key, value vector bằng cách nhân ma trận biểu diễn các từ đầu vào với ma trận học tương ứng.

* query vector: vector dùng để chứa thông tin của từ được tìm kiếm, so sánh. Giống như là câu query của google search.
* key vector: vector dùng để biểu diễn thông tin các từ được so sánh với từ cần tìm kiếm ở trên. Ví dụ, như các trang web mà google sẽ so sánh với từ khóa mà bạn tìm kiếm.
* value vector: vector biểu diễn nội dung, ý nghĩa của các từ. Các bạn có thể tượng tượng, nó như là nội dung trang web được hiển thị cho người dùng sau khi tìm kiếm.

Để tính tương quan, chúng ta đơn giản chỉ cần tính tích vô hướng dựa các vector query và key. Sau đó dùng hàm softmax để chuẩn hóa chỉ số tương quan trong đoạn 0-1, và cuối cùng, tính trung bình cộng có trọng số giữa các vector values sử dụng chỉ số tương quan mới tính được

Cụ thể hơn, quá trình tính toán attention vector có thể được tóm tắt làm 3 bước như sau:

* Bước 1: Tính ma trận query, key, value bằng cách khởi tạo 3 ma trận trọng số query, key, vector. Sau đó nhân input với các ma trận trọng số này để tạo thành 3 ma trận tương ứng.
* Bước 2: Tính attention weights. Nhân 2 ma trận key, query vừa được tính ở trên với nhau để với ý nghĩa là so sánh giữa câu query và key để học mối tương quan. Sau đó thì chuẩn hóa về đoạn [0-1] bằng hàm softmax. 1 có nghĩa là câu query giống với key, 0 có nghĩa là không giống.
* Bước 3: Tính output. Nhân attention weights với ma trận value. Điều này có nghĩa là chúng ta biểu diễn một từ bằng trung bình có trọng số (attention weights) của ma trận value.

Chúng ta muốn mô hình có thể học nhiều kiểu mối quan hệ giữa các từ với nhau. Với mỗi self-attention, chúng ta học được một kiểu pattern, do đó để có thể mở rộng khả năng này, chúng ta đơn giản là thêm nhiều self-attention. Tức là chúng ta cần nhiều ma trận query, key, value mà thôi. Giờ đây ma trận trọng số key, query, value sẽ có thêm 1 chiều depth nữa.

# Chương 3: Các giải pháp đề xuất

## Tiền xử lý dữ liệu

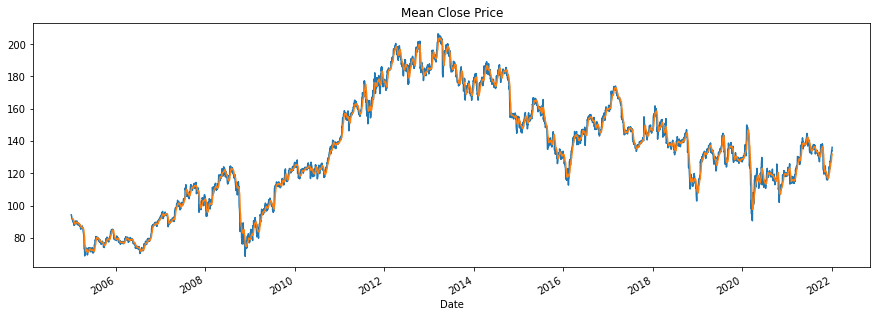
1. **Bộ dữ liệu**

Trong nghiên cứu này, tôi sử dụng giá đóng cửa của cổ phiếu IBM trong 17 năm (từ ngày 01-01-2005 đến ngày 01-01-2022) để huấn luyện các mô hình và bộ dữ liệu của bitcoin để kiểm thử. IBM có lịch sử lâu đời và là công ty công nghệ hàng đầu thế giới. Mọi biến động về công nghệ trên thế giới đều được phản ánh qua giá cổ phiếu của IBM. Cách dễ nhất để có được giá lịch sử của IBM là chỉ cần tải xuống bộ dữ liệu từ trang yahoo của IBM. Đối với mỗi ngày giao dịch, Yahoo cung cấp giá Mở, Cao, Thấp, Đóng và Khối lượng.

1. **Moving average**

Mục tiêu của đề tài là phát hiện các xu hướng của cổ phiếu tuy nhiên dữ liệu theo ngày thường có nhiều nhiễu do biến động giá. Để giảm nhiễu và làm biểu đồ mượt hơn, một kỹ thuật thường được sử dụng là đường trung bình động. Nhìn vào hướng của đường trung bình có thể thấy được cơ bản về cách giá đang di chuyển. Nếu nó nghiêng lên, giá đang trong tăng (hoặc gần đây), nếu nó nghiêng xuống và giá đang giảm, đi ngang và giá có khả năng biến động ổn định trong 1 phạm vi.

Đường [trung bình](https://www.investopedia.com/terms/m/movingaverage.asp) động (Moving Average) là một công cụ [phân tích kỹ thuật](https://www.investopedia.com/terms/t/technicalanalysis.asp) đơn giản giúp làm mịn dữ liệu giá bằng cách tạo ra [mức giá trung bình](https://www.investopedia.com/terms/a/averageprice.asp). Mức trung bình được lấy trong một khoảng thời gian cụ thể, như 10 ngày, 20 phút, 30 tuần hoặc bất kỳ khoảng thời gian nào mà nhà giao dịch chọn. Trong nghiên cứu này, tôi sử dụng cửa sổ trượt 10 ngày để tạo đường trung bình động.



Hình 1: Dữ liệu thô và dữ liệu sau khi moving average

1. **Chuyển đổi dữ liệu thành phần trăm thay đổi**

Cung cấp dữ liệu giá thô vào mô hình học sâu thường là một ý tưởng tồi do giá của cổ phiếu phụ thuộc vào đơn vị tiền tệ. Khi nhìn vào biểu đồ giá của IBM, bạn có thể thấy giá từ năm 2005 đến năm 2008 ($ 50- $ 100) ở một mức hoàn toàn khác so với giá của những năm 2020 và 2022 ($ 140- $ 220). Về bản chất, hai khoảng giá này không liên quan gì đến nhau. Để đưa các điểm giá trong quá khứ ngang bằng với điểm giá trong hiện tại và hữu ích hơn cho việc đào tạo mạng nơ-ron của, tôi chuyển đổi giá của một ngày thành phần trăm thay đổi so với ngày trước đó. Lợi ích của việc chuyển đổi này là đưa giá trong quá khứ và hiện tại về khoảng [-1, 1] để giá không còn phụ thuộc vào loại tiền tệ đồng thời thể hiện rõ hơn xu hướng tăng giảm của cổ phiếu.

Tiếp theo, sử dụng min-max scaler để chuyển data về khoảng [0, 1] để quá trình train hội tụ nhanh hơn và tránh vanishing và exploding gradient.

1. **Tạo nhãn**

Ý tưởng của mô hình là sử dụng sự biến động giá trong quá khứ để dự đoán tương lai. Do đó, tôi sử dụng cửa sổ trượt n ngày (n là tham số) để dự đoán xu hướng cho ngày thứ n + 1. Nghĩa là input của mô hình sẽ là giá đóng cửa cửa n ngày và output là giá cổ phiếu của ngày thứ n + 1. Ví dụ với cửa sổ trượt là 10 thì input thứ nhất là 10 ngày đầu tiên và nhãn là ngày thứ 11, tiếp đó cửa sổ trượt lấy input từ ngày thứ 2 đến ngày thứ 11 và nhãn là ngày thứ 12, cứ như vậy cho đến hết bộ dữ liệu.

1. **Chia bộ dữ liệu**

Trên bộ dữ liệu train (IBM), với mỗi thuật toán, tôi sử dụng phương pháp đánh giá chéo (cross validation) với tỷ lệ 80% train 20% validation để chọn ra bộ tham số tốt nhất. Sau đó mô hình với bộ tham số tốt nhất được đánh giá trên bộ dữ liệu kiểm thử (BTC).

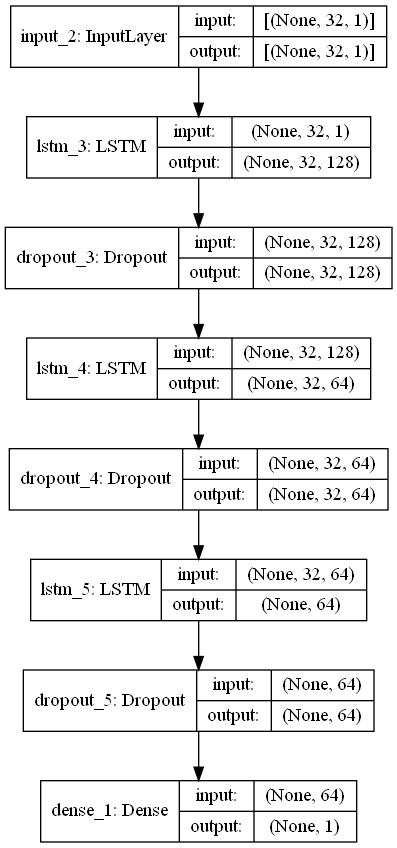
Ngoài ra bộ validation được sử dụng để kiểm tra overfitting và underfiting của mô hình.

## LSTM

LSTM là một cải tiến của mạng RNN truyền thống giúp giải quyết các vấn đề triệt tiêu gradient và bùng nổ gradient trong quá trình học. LSTM có thể lưu trữ thông tin trạng thái ở các tế bào trước bằng cell state. Cell state là một dạng giống như băng truyền. Nó chạy xuyên suốt tất cả các mắt xích (các nút mạng) và chỉ tương tác tuyến tính đôi chút. Vì vậy mà các thông tin có thể dễ dàng truyền đi thông suốt mà không sợ bị thay đổi. Mạng sẽ tự động học xem nên giữ những thông tin cũ nào và nên bỏ những thông tin nào đi. Điều này giúp LSTM trở nên phổ biến trong giải quyết các bài toán có dữ liệu tuần tự.

Mô hình LSTM được sử dụng trong nghiên cứu này được thể hiện như hình dưới.

Thay vì sử dụng một lớp LSTM, mô hình được thiết kế với 3 lớp LSTM nối tiếp nhau để tăng khả năng biểu diễn của mạng. Sau mỗi lớp LSTM sẽ có một lớp dropout với hệ số 0.25 để tránh overfitting và tăng tốc độ huấn luyện cho mô hình. Cuối cùng, tôi dùng một lớp tuyến tính để biến đổi đầu ra về 1 chiều.

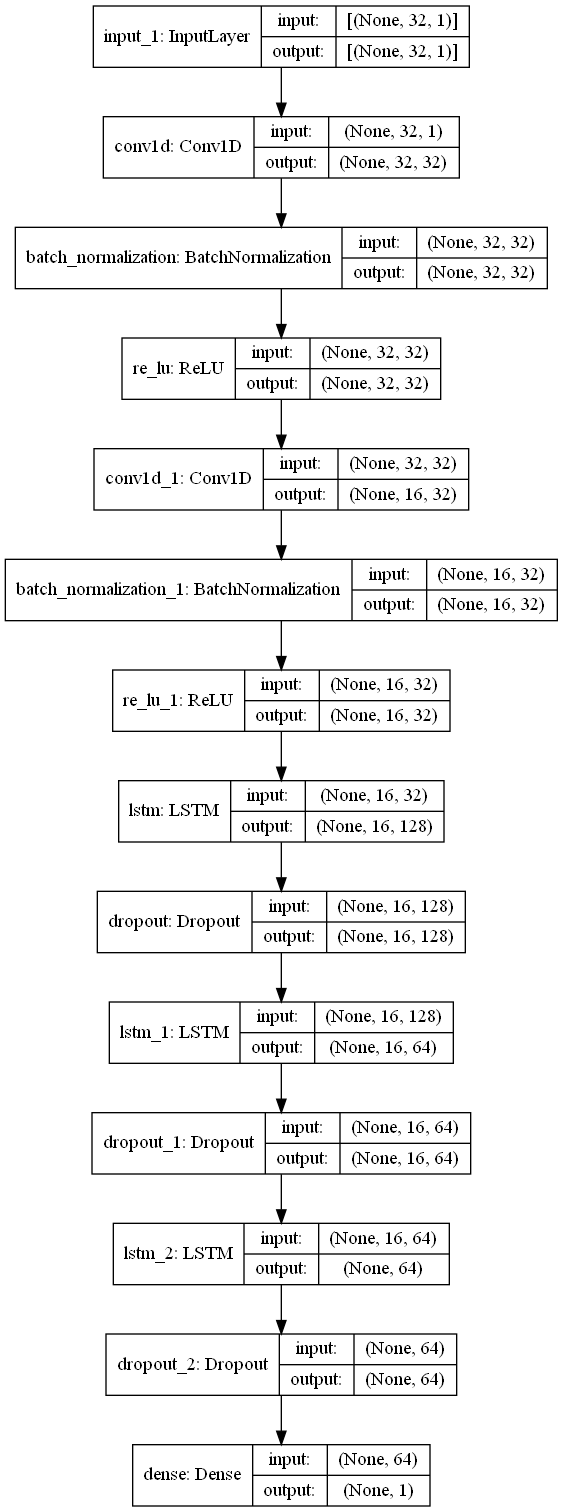


Hình 2: LSTM model

## LSTM và CNN

Thông thường, các mô hình CNN được sử dụng trong các bài toán với hình ảnh mà ở đó mỗi lớp CNN sẽ trích xuất các đặc trưng khác nhau từ hình ảnh. Tuy nhiên, những năm gần đây, người ta đã chỉ ra rằng các mô hình CNN cũng có hiệu quả tương tự khi áp dụng vào các bài toán có dữ liệu dạng chuỗi. Các lớp CNN hoạt động tốt trong việc trích xuất các đặc trưng về mặt không gian.

Cấu trúc model LSTM được giữ nguyên như trên. Đồng thời thêm 2 lớp CNN với hàm kích hoạt ReLU để trích xuất các đặc trưng về không gian trong dữ liệu. BatchNormalization được thêm vào sau lớp tích chập để tăng tốc độ huấn luyện mô hình và tránh triệt tiêu hay bùng nổ gradient.



Hình 3: LSTM và CNN model

## Transformer Encoder và Time Embedding

1. **Time2Vec**

Khi xử lý dữ liệu dạng time-series, thời gian là một đặc trưng cần thiết. Tuy nhiên trong kiến trúc transformer, các chuỗi được chuyển tiếp cùng một lúc qua mô hình, gây khó khăn trong việc trích xuất các đặc trưng về thời gian. Do đó Transformer có xu hướng sử dụng time embedding để cung cấp các thông tin về thời gian hay thứ tự cho mô hình.

Nếu không có thông tin về thứ tự hay thời gian, giá cổ phiếu của năm 2020 có thể ảnh hưởng đến giá cổ phiếu của năm 2008, điều này là không thực tế.

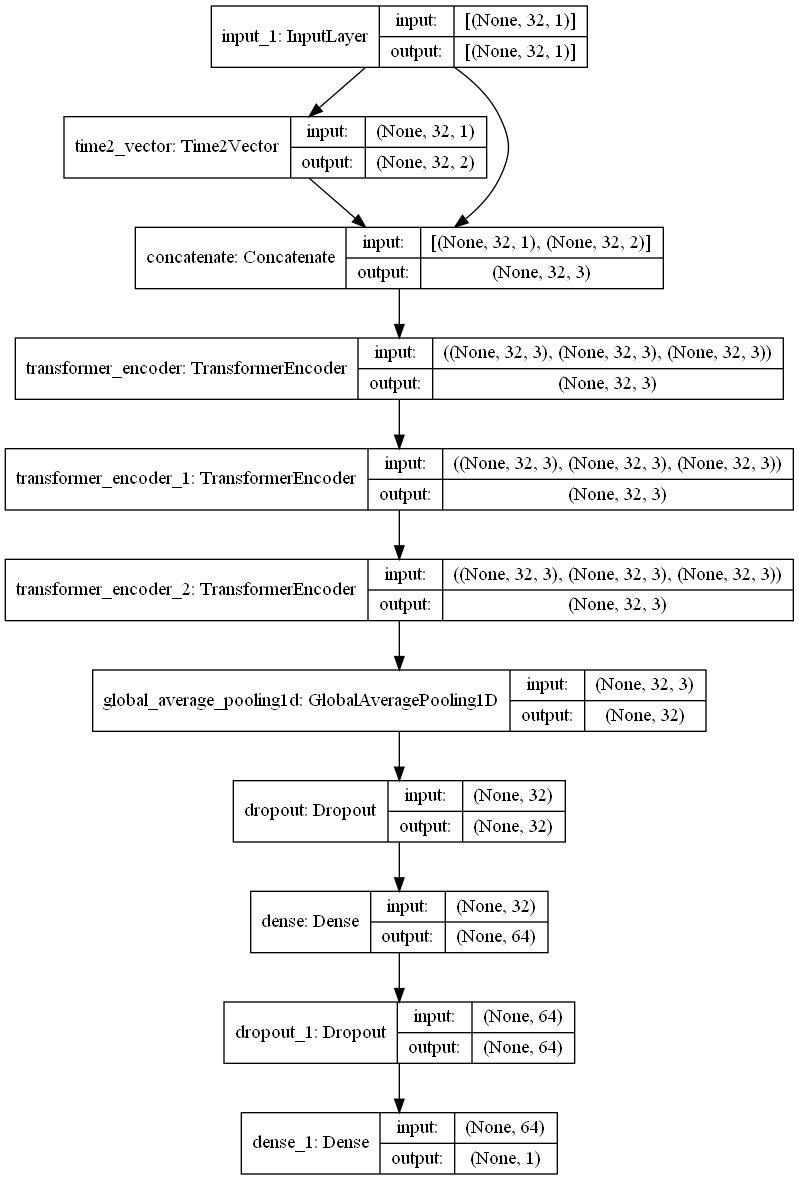
Để mã hóa các đặc trưng về thời gian, tôi sử dụng phương pháp được giới thiệu trong bài báo **Time2Vec: Learning a Vector Representation of Time [2].** Ý tưởng chính trong bài báo là chia đặc trưng về thời gian thành 2 loại có chu kỳ và không có chu kỳ. Không có chu kỳ là những đặc trưng có xác suất cao xảy ra trong khoảng thời gian. Các đặc trưng này được trích xuất qua một phép biến đổi tuyến tính ***t2v(τ)=ωᵢτ + φᵢ***. Có chu kỳ là những đặc trưng có tính lặp đi lặp lại theo thời gian như 4 mùa trong 1 năm. Loại thời gian này được biểu diễn bằng một hàm tuần hoàn ***t2v(τ)=F(ωᵢτ + φᵢ)*.** Trong kết quả cuối cùng, hàm sin cho kết quả thực nghiệm tốt nhất.

Như vậy sau lớp Time Embedding, đặc trưng về thời gian của mỗi timestep ban đầu sẽ được biểu diễn bằng 2 vector. Sau đó các đặc trưng về thời gian được thêm vào các timesteps.

1. **Transformer Encoder**

Transformer là một kiến trúc mạng neural sử dụng cơ chế self-attention, cho phép mô hình tập trung vào các phần có liên quan của chuỗi dữ liệu để cải thiện chất lượng dự đoán. Cơ chế self-attention bao gồm lớp Single-Head Attention và lớp Multi-Head Attention cho phép kết nối tất cả các bước trong chuỗi dữ liệu với nhau cùng một lúc, dẫn đến việc tạo ra các cách hiểu về sự phụ thuộc lâu dài. Cuối cùng, tất cả các quá trình này được song song hóa trong kiến trúc Transformer, cho phép tăng tốc quá trình học.

Kiến trúc model Transformer Encoder tương tự trong kiến trúc BERT [2]. Điểm khác biệt là model trong nghiên cứu này sử dụng Time2Vec để mã hóa thông tin thời gian và sử dụng 3 lớp Tranformer Encoder. Sau đó kết quả được đưa qua 2 lớp fully connected để giảm số chiều về 1.



Hình 4: Tranformer Encoder và Time Embedding model

# Phần 4: Kết quả thực nghiệm

Các mô hình được train trên colab free với 2 CPU 12GB RAM và GPU testla K80. Các mô hình được cài đặt bằng tensorflow và train 100 epochs với hàm loss Mean squared error và thuật toán tối ưu Adam. Mỗi mô hình sẽ được đánh giá 2 tham số là batch size và timesteps bằng phương pháp đánh giá chéo (cross validation) để chọn ra bộ tham số tốt nhất. Sau đó mỗi mô hình với bộ tham số tốt nhất được đánh giá trên bộ dữ liệu kiểm thử.

## Lựa chọn tham số

1. LSTM

Thực hiện grid search với 2 tham số timesteps và batch size với mô hình LSTM, kết quả thu được như sau:

| Batch size | Timesteps | Validate loss | Test loss | Train time |
| --- | --- | --- | --- | --- |
| **32** | **32** | **0.0027** | **0.0023** | **30m** |
| 32 | 64 | 0.0030 | 0.0032 | 47m |
| 32 | 128 | 0.0030 | 0.0042 | 60m |
| 64 | 32 | 0.0031 | 0.0037 | 48m |
| 128 | 32 | 0.0031 | 0.0050 | 60m |

Bảng 1: Kết quả grid search với LSTM model

Kết quả cho thấy với timesteps bằng 32 và batch size bằng 32 thì mô hình LSTM đạt kết quả tốt nhất.

1. LSTM và CNN

Thực hiện grid search với 2 tham số timesteps và batch size với mô hình LSTM, kết quả thu được như sau:

| Batch size | Timesteps | Validate loss | Test loss | Train time |
| --- | --- | --- | --- | --- |
| **32** | **32** | **0.0029** | **0.0022** | **40m** |
| 32 | 64 | 0.0031 | 0.0035 | 48m |
| 32 | 128 | 0.0035 | 0.0040 | 60m |
| 64 | 32 | 0.0030 | 0.0030 | 45m |
| 128 | 32 | 0.0031 | 0.0031 | 54m |

Bảng 2: Kết quả grid search với LSTM và CNN model

Kết quả cho thấy với timesteps bằng 32 và batch size bằng 32 thì mô hình LSTM kết hợp CNN đạt kết quả tốt nhất.

1. Transformer Encoder và Time Embedding

Thực hiện grid search với 2 tham số timesteps và batch size với mô hình LSTM, kết quả thu được như sau:

| Batch size | Timesteps | Validate loss | Test loss | Train time |
| --- | --- | --- | --- | --- |
| **32** | **32** | **0.0034** | **0.0048** | **30m** |
| 32 | 64 | 0.0039 | 0.0032 | 48m |
| 32 | 128 | 0.0042 | 0.0073 | 60m |
| 64 | 32 | 0.0030 | 0.0030 | 45m |
| 128 | 32 | 0.0031 | 0.0031 | 54m |

Bảng 3: Kết quả grid search với Transformer Encoder và Time Embedding model

Kết quả cho thấy với timesteps bằng 32 và batch size bằng 32 thì mô hình Transformer đạt kết quả tốt nhất.

## So sánh các thuật toán

Các mô hình với bộ tham số tốt nhất sẽ được đánh giá trên bộ dữ liệu test là giá cổ phiếu của bitcoin từ 01-01-2017 đến nay (ngày 21-01-2022). Kết quả trên tập kiểm thử của các model:

| Model | Test loss | Train time |
| --- | --- | --- |
| LSTM | 0.0023 | 30m |
| **LSTM + CNN** | **0.0022** | **30m** |
| Tranformer encoder | 0.0048 | 40m |

Bảng 4: So sánh kết quả các mô hình trên bộ dữ liệu test

A picture containing tool

Description automatically generated

Hình 5: kết quả dự đoán của LSTM model

Chart

Description automatically generated with medium confidence

Hình 6: kết quả dự đoán của LSTM và CNN model

A picture containing needle

Description automatically generated

Hình 7: kết quả dự đoán của Tranformer Encoder và Time Embedding model

Có thể thấy mô hình LSTM kết hợp CNN đạt kết quả tốt nhất sau đó là LSTM. Mô hình Transformer không hiệu quả cho bài toán này. Các đường dự đoán của LSTM, LSTM kết hợp CNN sát với kết quả thực tế. Với mô hình transformer, lý giải cho kết quả không tốt là việc các dự đoán trong quá khứ bị phụ thuộc vào tương lai do khả năng xử lý song song của mô hình.

# Phần 5: Kết luận

Trong nghiên cứu trên, tôi đã tiến hành giải quyết bài toán mang tính thực tế cao, đó là dự đoán xu hướng thị trường chứng khoán. Với cách tiếp cận xử lý dữ liệu mới là phân tích phần trăm thay đổi và thử nghiệm trên 3 mô hình LSTM, LSTM kết hợp CNN và Transformer, hai trong số 3 mô hình cho kết quả tốt. Tuy nhiên, mô hình Transformer chưa thực sự tốt và cần cải thiện thêm.

# Tài liệu tham khảo

[1] Stelios D. Bekiros (2010) Fuzzy Adaptive Decision Making for Boundedly Rational Traders in Speculative Stock Markets *European Journal of Operational Research* **202**(1):285-293.  
[2] Zhang, Y., Yang, X. (2016) Online Portfolio Selection Strategy based on Combining Experts Advice *Computational Economics* **50**(5).  
[3] Cavalcante, R.C., Brasileiro, R.C., Souza, V.F., Nobrega, J.P. and Oliveira, A. (2016) Computational Intelligence and Financial Markets: A Survey and Future Directions *Expert Systems with Applications* **55**:194-211  
[4] Wang, L., Hajric, V. (2020) The Cost of Bad Market Timing Decisions in 2020 was Annahilation *Bloomberg*.  
[5] Wang B, Huang H, Wang X. (2012) A novel text mining approach to financial time series forecasting *Neurocomputing* **83**(6): 136-145.  
[6] Guo Z, Wang H, Liu Q, Yang J. (2014) A Feature Fusion Based Forecasting Model for Financial Time Series *Plos One* **9**(6): 172-200.  
[7] Prasaddas S, Padhy S. (2012) Support Vector Machines for Prediction of Futures Prices in Indian Stock Market *International Journal of Computer Applications* **41**(3): 22-26.  
[8] Hinton GE, Salakhutdinov RR (2006) Reducing the Dimensionality of Data with Neural Networks *Science* **313**(5786): 504-507.  
[9] Bengio Y, Courville A, Vincent P. (2013) Representation Learning: A Review and New Perspectives *IEEE Transactions on Pattern Analysis Machine Intelligence* **35**(8): 1798-1828.  
[10] Arevalo, A., Nino, J., Hernandez, G. and Sandoval., J. (2016) High-Frequency Trading Strategy Based on Deep Neural Networks *ICIC*.  
[11] Bao, W.N., Yue, J. and Rao, Y. (2017) A Deep Learning Framework for Financial Time Series using Stacked Autoencoders and Long-Short Term Memory *Plos one* **12**.  
[12] M, H., Gopalakrishnan, E.A., Menon, V. and Kp, S. (2018) NSE Stock Market Prediction Using DeepLearning Models *Procedia Computer Science* **132**(10): 1351-1362.  
[13] Sutskever, I., Vinyals, O. and Le, Q. V., (2014) Sequence to Sequence Learning with Neural Networks *Advances in neural information processing systems*: 3104-3112.

[14] Moody, J.E. and Saffell, M. (2001) Learning to Trade via Direct Reinforcement *IEEE Transactions on Neural Networks* **12**(4): 875-889.  
[15] Dempster, M.A. and Leemans, V. (2006) An Automated FX Trading System using Adaptive Reinforcement Learning *Expert Systems Applications* **30**(5): 543-552