BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC NÔNG LÂM TP HCM

KHOA CÔNG NGHỆ THÔNG TIN



TIỂU LUẬN TỐT NGHIỆP

**Tên đề tài: STOCKPREDICT - DỰ ĐOÁN CHỨNG KHOÁN**

Ngành: Công Nghệ thông tin

Niên khoá: 2022 – 2026

Lớp: DH22DTA

Sinh viên thực hiện:

1. Trần Quốc Hưng - 22130094

Cán bộ hướng dẫn: Ths. Nguyễn Đức Công Song

TP. HỒ CHÍ MINH , tháng , năm

DANH SÁCH CHỮ VIẾT TẮT

ANSI **A**merican **N**ational **S**tandards **I**nstitute

Học viện chuẩn hoá quốc gia Mỹ

ATM **A**utomatic **T**eller **M**achine

Máy rút tiền tự động

CA **C**ertification **A**uthority

Chứng thực

ISO **I**nternational **S**tandardizations **O**rganization

Tổ chức chuẩn hoá quốc tế

MAC **M**essage **A**uthentication **C**ode

Mã chứng thực thông điệp

UUCP **U**nix-to-**U**nix **C**opy **P**rotocol

Giao thức truyền nhận thông điệp trên Unix

DANH MỤC HÌNH ẢNH

[Hình 2.1 Mạch chỉnh lưu và bộ lọc 3](#_Toc98229800)

# DANH MỤC BẢNG

Bảng 1.1: Tên bảng 19

TÓM TẮT

**MỤC LỤC**

[DANH SÁCH CHỮ VIẾT TẮT 1](#_Toc164966931)

[DANH MỤC HÌNH ẢNH 2](#_Toc164966932)

[DANH MỤC BẢNG 3](#_Toc164966933)

[TÓM TẮT 4](#_Toc164966934)

[MỞ ĐẦU 6](#_Toc164966935)

[1. LÝ DO CHỌN ĐỀ TÀI 6](#_Toc164966936)

[2. MỤC TIÊU VÀ PHẠM VI NGHIÊN CỨU 6](#_Toc164966937)

[3. Ý NGHĨA KHOA HỌC VÀ THỰC TIỄN CỦA ĐỀ TÀI 6](#_Toc164966938)

[CHƯƠNG 1. TỔNG QUAN ĐỀ TÀI 7](#_Toc164966939)

[1. 1. Phân tích, đánh giá các công trình nghiên cứu đã có liên quan đến đề tài nghiên cứu 7](#_Toc164966940)

[1. 2. Nêu những vấn đề còn tồn tại 7](#_Toc164966941)

[1. 3. Chỉ ra những vấn đề mà tiểu luận cần tập trung nghiên cứu, giải quyết 7](#_Toc164966942)

[CHƯƠNG 2. PHƯƠNG PHÁP VÀ NỘI DUNG NGHIÊN CỨU 8](#_Toc164966943)

[2. 1. Trình bày cơ sở lí thuyết, lí luận, giả thiết khoa học 8](#_Toc164966944)

[2. 2. Trình bày mô hình lý thuyết của giải pháp đã đề xuất sử dụng trong tiểu luận 8](#_Toc164966945)

[CHƯƠNG 3. GIẢI PHÁP CHO BÀI TOÁN/VẤN ĐỀ/MÔ HÌNH 9](#_Toc164966946)

[3.1. Phát biểu mô hình/bài toán trong đề tài: cụ thể, rõ ràng, có thể phát biểu bằng ngôn ngữ tự nhiên, hay mô hình toán học, … 9](#_Toc164966947)

[3.2. Giải pháp cụ thể để giải quyết mô hình/bài toán: 9](#_Toc164966948)

[3.3 HIỆN THỰC GIẢI PHÁP 10](#_Toc164966949)

[CHƯƠNG 5. KẾT QUẢ, KẾT LUẬN VÀ KIẾN NGHỊ 12](#_Toc164966950)

[5.1. KẾT QUẢ 12](#_Toc164966951)

[5. 2. KẾT LUẬN 12](#_Toc164966952)

[5. 3. KIẾN NGHỊ 12](#_Toc164966953)

[PHỤ LỤC (nếu có) 14](#_Toc164966954)

MỞ ĐẦU

1. **Lý do chọn đề tài**

Trong bối cảnh kinh tế toàn cầu ngày càng phát triển, thị trường chứng khoán đóng vai trò quan trọng trong việc huy động vốn và tạo điều kiện cho các nhà đầu tư tìm kiếm lợi nhuận. Tuy nhiên, do sự biến động phức tạp của thị trường, việc dự đoán giá cổ phiếu luôn là một thách thức lớn đối với các nhà đầu tư, chuyên gia tài chính cũng như các tổ chức kinh tế.

Truyền thống, việc phân tích và dự báo giá cổ phiếu dựa trên các phương pháp phân tích cơ bản (fundamental analysis) và phân tích kỹ thuật (technical analysis). Tuy nhiên, với sự phát triển mạnh mẽ của trí tuệ nhân tạo (AI), đặc biệt là các mô hình học máy (Machine Learning) và học sâu (Deep Learning), việc áp dụng AI vào dự đoán giá cổ phiếu đã và đang mang lại những kết quả đáng khích lệ. Các mô hình AI có khả năng phân tích khối lượng dữ liệu khổng lồ từ nhiều nguồn khác nhau như dữ liệu thị trường, tin tức, mạng xã hội, tình hình kinh tế vĩ mô và vi mô… giúp nhận diện xu hướng và đưa ra dự đoán chính xác hơn.

Lựa chọn đề tài "Ứng dụng AI để dự đoán giá cổ phiếu" có ý nghĩa thực tiễn cao, góp phần giải quyết một trong những bài toán quan trọng trong lĩnh vực tài chính.

1. **Mục tiêu**

- Tìm hiểu tổng quan về thị trường chứng khoán và các phương pháp dự đoán giá cổ phiếu

- Tìm hiểu các mô hình Machine Learning (ML) và Deep Learning (DL) phổ biến trong dự đoán chuỗi thời gian ( Hồi quy tuyến tính, Cây quyết định, Mạng nơ-ron nhân tạo… ).

# - Thu thập dữ liệu lịch sử giá cổ phiếu từ các nguồn đáng tin cậy như Yahoo Finance, Google Finance, hoặc các sàn giao dịch chứng khoán. Xử lý dữ liệu, loại bỏ nhiễu, điền giá trị thiếu và chuẩn hóa dữ liệu để phù hợp với mô hình AI.Khám phá dữ liệu (EDA - Exploratory Data Analysis) nhằm tìm ra các đặc trưng quan trọng ảnh hưởng đến giá cổ phiếu.

- Xây dựng, huấn luyện và đánh giá mô hình AI:

* Xây dựng mô hình AI dựa trên dữ liệu đã tiền xử lý.
* Huấn luyện mô hình với tập dữ liệu huấn luyện, tối ưu hóa tham số và cải thiện độ chính xác.
* Đánh giá mô hình dựa trên các chỉ số hiệu suất như RMSE (Root Mean Squared Error), MAE (Mean Absolute Error) và R² (R-squared).
* So sánh kết quả giữa các mô hình để tìm ra mô hình có hiệu suất tốt nhất.

1. **NỘI DUNG VÀ PHẠM VI NGHIÊN CỨU**
   1. Các nội dung cần nghiên cứu, thử nghiệm
      * Ứng dụng AI trong dự đoán giá cổ phiếu, tập trung vào các mô hình ML và DL như Random Forest, LSTM.
      * Thu thập và xử lý dữ liệu từ các sàn giao dịch chứng khoán.
      * Xây dựng, huấn luyện và đánh giá mô hình dự đoán theo các tiêu chí RMSE, MAE.
      * Kiểm định thực tế và đánh giá khả năng ứng dụng.
   2. **Các bước thực hiện:**
      * **Nghiên cứu lý thuyết**: Tìm hiểu mô hình AI và phương pháp dự đoán.
      * **Thu thập dữ liệu**: Lấy dữ liệu giá cổ phiếu, tiền xử lý và chuẩn hóa.
      * **Xây dựng mô hình**: Huấn luyện, tối ưu hóa và kiểm tra độ chính xác.
      * **Đánh giá mô hình**: So sánh hiệu suất giữa các mô hình.
      * **Kiểm định thực tế**: Ứng dụng mô hình vào dự đoán và đề xuất cải tiến.
2. **CƠ SỞ KHOA HỌC VÀ THỰC TIỄN**
   1. **Cơ sở lý thuyết**
      * **Lý thuyết chuỗi thời gian**: Dự đoán giá cổ phiếu dựa trên dữ liệu lịch sử.
      * **Trí tuệ nhân tạo (AI)**: Ứng dụng Machine Learning (ML) và Deep Learning (DL) trong phân tích tài chính.
      * **Các mô hình dự đoán**: Hồi quy tuyến tính, Random Forest, XGBoost, LSTM.

**Ngôn ngữ lập trình**: Python.

**Thư viện ML/DL**: Scikit-learn, TensorFlow, Keras.

**Xử lý dữ liệu**: Pandas, NumPy.

**Trực quan hóa dữ liệu**: Matplotlib, Seaborn.

**3. Giải thuật áp dụng**

**Machine Learning**: Hồi quy tuyến tính, Random Forest, XGBoost.

**Deep Learning**: Mạng nơ-ron nhân tạo (ANN), mạng nơ-ron hồi quy (LSTM, GRU).

**Tiền xử lý dữ liệu**: Chuẩn hóa, loại bỏ nhiễu, xử lý giá trị thiếu.

Các công nghệ và giải thuật này giúp xây dựng mô hình AI hiệu quả để dự đoán giá cổ phiếu, hỗ trợ ra quyết định trong đầu tư tài chính.

1. **SẢN PHẨM CỦA ĐỀ TÀI**

1. Sản phẩm ứng dụng

Sau khi hoàn tất đề tài, sản phẩm sẽ là một hệ thống dự đoán giá cổ phiếu dựa trên AI với các chức năng chính:

Thu thập và xử lý dữ liệu: Tải dữ liệu giá cổ phiếu từ các nguồn (Yahoo Finance, Google Finance), làm sạch và chuẩn hóa dữ liệu.

Huấn luyện mô hình AI: Xây dựng và huấn luyện các mô hình Machine Learning và Deep Learning để dự đoán giá cổ phiếu.

Dự đoán giá cổ phiếu: Dự báo giá cổ phiếu trong ngắn hạn dựa trên dữ liệu đầu vào.

Đánh giá hiệu suất mô hình: So sánh độ chính xác giữa các mô hình qua các chỉ số RMSE, MAE.

Giao diện trực quan (nếu có): Hiển thị kết quả dự đoán dưới dạng đồ thị và số liệu phân tích.

2. Sản phẩm nghiên cứu

Báo cáo phân tích: Tổng hợp kết quả nghiên cứu về AI trong dự đoán giá cổ phiếu.

So sánh mô hình: Đánh giá ưu, nhược điểm của các mô hình AI áp dụng.

Đề xuất cải tiến: Hướng phát triển nhằm nâng cao độ chính xác và tính ứng dụng thực tiễn.