BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC THĂNG LONG

BÁO CÁO BÀI TẬP LỚN
MÔN: DỮ LIỆU LỚN

DỰ ĐOÁN GIÁ VÀNG

GIÁO VIÊN HƯỚNG DẪN:	Nguyễn Quang Huy	
LỚP:	K3N22425_IS330_03	
SINH VIÊN THỰC HIỆN:	Nguyễn Quang Bách	A47330
	Vũ Thành Đạt	A47996
	Tạ Văn Bằng	A48447

HÀ NỘI – 2025

LỜI CẢM ƠN

Lời đầu tiên, chúng em xin gửi lời cảm ơn chân thành nhất đến thầy Nguyễn Quang Huy. Trong quá trình học tập, nghiên cứu và xây dựng bài tập lớn **"Dự đoán giá vàng",** nhóm em đã nhận được rất nhiều sự quan tâm, góp ý, hướng dẫn tận tình của thầy. Thầy đã giúp nhóm tích lũy thêm nhiều kiến thức bổ ích để không chỉ hoàn thành được bài tập lớn, mà còn học được nhiều kinh nghiệm về lĩnh vực thị giác máy tính. Tất cả những kiến thức này đều rất hữu ích sau này cho công việc của chúng em.

Có lẽ kiến thức là vô hạn mà sự tiếp nhận kiến thức của bản thân mỗi người luôn tồn tại những hạn chế nhất định. Do đó, trong quá trình hoàn thành bài tập lớn, chắc chắn không tránh khỏi những thiếu sót. Nhóm em rất mong nhận được những góp ý đến từ thầy để bài tiểu luận của nhóm được hoàn thiện hơn.

Kính chúc thầy sức khỏe, hạnh phúc và thành công trên con đường sự nghiệp giảng dạy.

MỤC LỤC

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

Viết tắt	Tiếng Anh	Tiếng Việt
LSTM	Long Short-Term Memory	Bộ nhớ dài-ngắn hạn
MAE	Mean Absolute Error	Sai số tuyệt đối trung bình
MAPE	Mean Absolute Percentage Error	Sai số phần trăm tuyệt đối trung bĩnh
MSE	Mean Squared Error	Sai số toàn phương trung bình
RMSE	Root Mean Squared Error	Căn bậc hai của sai số bình phương trung bình
RNN	Recurrent Neural Network	Mạng nơ-ron hồi tiếp
MLlib	Machine Learning Library	Thư viện học máy
Spark	Apache Spark	Apache Spark

CHƯƠNG 1 - TỔNG QUAN VỀ ĐỀ TÀI

1.1. LÝ DO CHỌN ĐỀ TÀI

Trong bối cảnh kinh tế toàn cầu không ngừng biến động, giá vàng luôn là một trong những chỉ số tài chính được quan tâm hàng đầu. Dự đoán giá vàng không chỉ giúp nhà đầu tư đưa ra quyết định đúng đắn mà còn đóng vai trò quan trọng trong phân tích thị trường và quản lý rủi ro. Với sự phát triển mạnh mẽ của trí tuệ nhân tạo, đặc biệt là các mô hình học sâu như LSTM, việc áp dụng công nghệ để dự đoán giá vàng trở nên khả thi và hiệu quả hơn. Vì vậy, nhóm em chọn đề tài "Dự đoán giá vàng sử dụng mô hình LSTM" nhằm tìm hiểu cách áp dụng học sâu vào bài toán thực tế và nâng cao kiến thức về xử lý chuỗi thời gian.

1.2. MỤC ĐÍCH, ĐỐI TƯỢNG SỬ DỤNG

1.2.1. Mục đích

- Ứng dụng mô hình LSTM để dự đoán giá vàng trong tương lai dựa trên dữ liệu lịch sử.
- Đánh giá độ chính xác và hiệu quả của mô hình LSTM so với các phương pháp truyền thống.
- Củng cố kỹ năng tiền xử lý dữ liệu, xây dựng mô hình học máy và trực quan hóa kết quả.

1.2.2. Đối tượng sử dụng

- Các nhà đầu tư, phân tích tài chính cần công cụ hỗ trợ dự đoán giá vàng.
- Sinh viên, học viên đang học về khoa học dữ liệu, trí tuệ nhân tạo muốn tiếp cận các ứng dụng thực tế của LSTM.

1.3. PHƯƠNG PHÁP NGHIÊN CỨU

Để tiến hành xây dựng đề tài, nhóm sẽ thực hiện các nghiên cứu dưới đây:

1.3.1. Thu thập và xử lý dữ liệu

- **Nguồn dữ liệu**: Dữ liệu lịch sử giá vàng được thu thập từ Yahoo Finance dưới dạng tập .csv.
- **Các trường dữ liệu sử dụng**: Date, Open, High, Low, Close, Volume, v.v. (trong đó cột Close là cột mục tiêu dự đoán).
- **Xử lý thiếu dữ liệu**: Loại bỏ hoặc nội suy các giá trị bị thiếu để đảm bảo tính liên tục của chuỗi thời gian.
- **Chuẩn hóa dữ liệu**: Sử dụng MinMaxScaler để đưa dữ liệu về khoảng [0, 1], giúp mô hình học hiệu quả hơn.

1.3.2. Tạo tập dữ liệu cho mô hình LSTM

- **Tạo chuỗi đầu vào**: Xây dựng các chuỗi con từ dữ liệu lịch sử theo một cửa sổ thời gian cố định (ví dụ: 30 ngày).
- **Chia tập dữ liệu**:
  + Tập huấn luyện: 90% dữ liệu đầu.
  + Tập kiểm tra: 10% dữ liệu cuối.
- **Định dạng lại dữ liệu**: Chuyển đổi định dạng cho phù hợp với đầu vào 3 chiều của LSTM: (samples, timesteps, features).

1.3.3. Xây dựng mô hình và đánh giá mô hình LSTM

- **Thiết kế mô hình**:
  + Sử dụng các lớp LSTM với số lượng tế bào phù hợp (ví dụ: 50–100 units).
  + Thêm lớp Dropout để giảm hiện tượng overfitting.
  + Kết thúc bằng lớp Dense để xuất giá trị dự đoán.
- **Biên dịch mô hình**:

+ Hàm mất mát: mean_squared_error (MSE)

+ Bộ tối ưu hóa: Adam

- **Huấn luyện mô hình**:

  + Epochs: từ 50 đến 100 lần.

  + Batch size: 32.

  + Theo dõi loss để đánh giá quá trình học.

- **Dự đoán trên tập kiểm tra**: So sánh giá trị dự đoán với giá trị thực tế.

- **Đánh giá bằng các chỉ số**:

  + MSE (Mean Squared Error)

  + RMSE (Root Mean Squared Error)

  + MAE (Mean Absolute Error)

- **Trực quan hóa**:

  + Biểu đồ giá thực tế và giá dự đoán để so sánh.

  + Biểu đồ loss trong quá trình huấn luyện.

1.3.4. Triển khai và ứng dụng

- Mô hình sau huấn luyện có thể được lưu lại (dạng .h5) và triển khai vào hệ thống dự đoán thực tế hoặc tích hợp vào giao diện người dùng.

CHƯƠNG 2 - CƠ SỞ LÝ THUYẾT

2.1. GIỚI THIỆU

2.1.1. Giới thiệu về dữ liệu lớn (Big data)

Trong thời đại công nghệ số phát triển mạnh mẽ, dữ liệu lớn (Big Data) ngày càng đóng vai trò quan trọng trong việc hỗ trợ ra quyết định, phân tích và dự báo trong nhiều lĩnh vực như kinh tế, tài chính, y tế, giáo dục, truyền thông,... Dữ liệu lớn không chỉ đơn thuần là dữ liệu có khối lượng lớn, mà còn mang nhiều đặc điểm đặc trưng như: tốc độ tạo và xử lý nhanh chóng, sự đa dạng về định dạng dữ liệu (văn bản, hình ảnh, âm thanh, video,...), tính không chắc chắn cao và đặc biệt là giá trị tiềm ẩn rất lớn nếu được khai thác đúng cách.

Việc khai thác dữ liệu lớn đòi hỏi sự kết hợp của nhiều công nghệ hiện đại như điện toán đám mây, hệ thống xử lý phân tán (như Hadoop, Spark), và đặc biệt là các phương pháp học máy (machine learning) và học sâu (deep learning). Nhờ vào khả năng học từ dữ liệu và tự động cải thiện, các mô hình trí tuệ nhân tạo hiện nay có thể xử lý và phân tích khối lượng dữ liệu khổng lồ nhằm tìm ra các xu hướng, mẫu hình và đưa ra dự đoán có độ chính xác cao.

2.1.2. Bài toán dự đoán giá vàng

Trong lĩnh vực tài chính, giá vàng luôn là một trong những chỉ số quan trọng được theo dõi sát sao bởi các nhà đầu tư, các tổ chức tài chính và cả các nhà hoạch định chính sách. Giá vàng chịu ảnh hưởng từ nhiều yếu tố khác nhau như biến động kinh tế toàn cầu, chính sách tiền tệ, lãi suất, tỷ giá hối đoái và cả yếu tố tâm lý của thị trường. Vì vậy, việc dự đoán chính xác xu hướng giá vàng trong tương lai có thể mang lại lợi thế cạnh tranh lớn cho các tổ chức tài chính cũng như giúp nhà đầu tư cá nhân đưa ra các quyết định mua – bán hợp lý.

Tuy nhiên, bài toán dự đoán giá vàng là một bài toán phức tạp vì tính phi tuyến và tính biến động mạnh của dữ liệu. Các mô hình truyền thống thường gặp khó khăn trong việc nắm bắt được các mối quan hệ phụ thuộc dài hạn trong chuỗi thời gian. Do đó, việc ứng dụng các mô hình học sâu, đặc biệt là mạng nơ-ron hồi tiếp LSTM (Long Short-Term Memory), trở thành một lựa chọn hiệu quả. LSTM có khả năng ghi nhớ thông tin trong khoảng thời gian dài và học được các mẫu chuỗi phức tạp, giúp cải thiện đáng kể độ chính xác trong các bài toán dự báo chuỗi thời gian như dự đoán giá vàng.

2.2. CÁC CÔNG NGHỆ SỬ DỤNG CHÍNH

2.2.1. Python

Ngôn ngữ lập trình chính được sử dụng trong toàn bộ quá trình xử lý dữ liệu, xây dựng mô hình và trực quan hóa kết quả. Python được ưa chuộng trong lĩnh vực khoa học dữ liệu nhờ cú pháp đơn giản và hệ sinh thái thư viện phong phú.

2.2.2. Pandas

Thư viện hỗ trợ thao tác và xử lý dữ liệu dạng bảng (DataFrame). Dùng để đọc dữ liệu từ file CSV, xử lý chuỗi thời gian và trích xuất các đặc trưng quan trọng từ dữ liệu giá vàng.

2.2.3. NumPy

Thư viện tính toán mảng số học hiệu suất cao, được sử dụng để thực hiện các thao tác xử lý dữ liệu nền tảng trước khi đưa vào mô hình.

2.2.4. Matplotlib & Seaborn

Cung cấp các công cụ tiền xử lý dữ liệu, đặc biệt là MinMaxScaler dùng để chuẩn hóa dữ liệu đầu vào về cùng một khoảng giá trị, giúp mô hình học hiệu quả hơn.

2.2.5. Scikit-learn

Cung cấp các công cụ tiền xử lý dữ liệu, đặc biệt là MinMaxScaler dùng để chuẩn hóa dữ liệu đầu vào về cùng một khoảng giá trị, giúp mô hình học hiệu quả hơn.

2.2.6. Keras (trên nền TensorFlow)

Thư viện học sâu dùng để xây dựng, huấn luyện và đánh giá mô hình LSTM. Keras cho phép tạo mô hình mạng nơ-ron với cấu trúc linh hoạt và dễ triển khai.

2.2.7. Recurrent Neural Network (RNN)

Mạng nơ-ron hồi tiếp (RNN) là một kiến trúc deep learning đặc biệt xử lý dữ liệu tuần tự; mỗi bước thời gian, RNN lưu giữ trạng thái ẩn để kết nối thông tin giữa các bước. RNN phù hợp với chuỗi thời gian, ngôn ngữ tự nhiên, nhưng gặp vấn đề vanishing/exploding gradients khi học dependencies dài hạn.

2.2.8. Long Short-Term Memory (LSTM)

LSTM là một phiên bản cải tiến của RNN, bổ sung các cổng (gate) như input, forget và output để kiểm soát luồng thông tin, khắc phục vấn đề vanishing gradient, nhờ đó học được phụ thuộc dài hạn hiệu quả hơn.

CHƯƠNG 3 - PHÁT TRIỂN HỆ THỐNG

3.1.1. Nguồn và mô tả dữ liệu

- Nguồn dữ liệu: Dữ liệu được thu thập từ trang Investing.com. Đây là một nguồn dữ liệu đáng tin cậy, thường xuyên được sử dụng trong các nghiên cứu tài chính và dự báo giá hàng hóa.
- Định dạng dữ liệu: Dữ liệu được cung cấp dưới định dạng CSV, bao gồm các cột chính: ngày, giá vàng (tính theo USD/ounce), giá mở cửa, giá đóng cửa, mức cao nhất và thấp nhất trong ngày.
- Mục đích sử dụng: Bộ dữ liệu được sử dụng làm đầu vào để huấn luyện và đánh giá mô hình LSTM, phục vụ mục tiêu dự báo chính xác giá vàng trong tương lai dựa trên dữ liệu lịch sử.

```python
df.show(5)
# Get the shape of the DataFrame using count() and len(columns)
num_rows = df.count()
num_cols = len(df.columns)
print(f"Shape: ({num_rows}, {num_cols})")
```

```
+----------+--------+--------+--------+--------+-------+--------+
|      Date|   Price|    Open|    High|     Low|   Vol.|Change %|
+----------+--------+--------+--------+--------+-------+--------+
|12/30/2022|1,826.20|1,821.80|1,832.40|1,819.80|107.50K|   0.01%|
|12/29/2022|1,826.00|1,812.30|1,827.30|1,811.20|105.99K|   0.56%|
|12/28/2022|1,815.80|1,822.40|1,822.80|1,804.20|118.08K|  -0.40%|
|12/27/2022|1,823.10|1,808.20|1,841.90|1,808.00|159.62K|   0.74%|
|12/26/2022|1,809.70|1,805.80|1,811.95|1,805.55|   NULL|   0.30%|
+----------+--------+--------+--------+--------+-------+--------+
only showing top 5 rows

Shape: (2583, 7)
```

```python
# Hiển thị cấu trúc dataframe
df.printSchema()
```

```
root
 |-- Date: string (nullable = true)
 |-- Price: string (nullable = true)
 |-- Open: string (nullable = true)
 |-- High: string (nullable = true)
 |-- Low: string (nullable = true)
 |-- Vol.: string (nullable = true)
 |-- Change %: string (nullable = true)
```

Ảnh 3.1. Tổng quan dữ liệu

3.1.2. Tiền xử lý dữ liệu

- Làm sạch:
  - + Loại bỏ ngày không giao dịch hoặc sử dụng giá ngày trước đó
  - + Loại ngoại lai theo Z-score > 3
- Xử lý thiếu dữ liệu:
  - + Nội suy tuyến tính cho thiếu dữ liệu dưới 3 ngày
  - + Forward fill cho thiếu dài hạn
- Chuẩn hóa:

    + Min–Max Scaling để biến đầu vào về range 0–1

- Tạo chuỗi đầu vào:
    + Window length = 30 ngày, dự báo ngày thứ 31
    + Kết quả: X shape (Số mẫu, 30, 1), y shape (Số mẫu, 1)
- Chia tập dữ liệu:
    + Train 90% (2013–2021)
    + Test 10% (2022)

3.1.3. Mô hình và phương pháp thực nghiệm

3.1.3.1. Tổng quan mô hình

Tổng quan kiến trúc: LSTM (Long Short-Term Memory) là mô hình RNN đặc biệt với cơ chế cổng (gates) kiểm soát luồng dữ liệu, cho phép mô hình học dependencies dài hạn của chuỗi thời gian. Kiến trúc hai lớp LSTM kết hợp dense layer giúp cân bằng giữa khả năng ghi nhớ thông tin và tính tổng quát.

3.1.3.2. Lý do lựa chọn mô hình

Xử lý chuỗi thời gian: LSTM khắc phục vấn đề vanishing gradient của RNN cơ bản, thích hợp với dữ liệu giá vàng có tính chuỗi.

Khả năng học dài hạn: Qua các cổng forget/input/output, LSTM giữ lại thông tin quan trọng qua nhiều bước thời gian.

Tính linh hoạt: Dễ dàng mở rộng thêm lớp, điều chỉnh số neuron, thêm dropout để giảm overfitting.

Chứng minh thực nghiệm: Nhiều nghiên cứu và ứng dụng thực tế cho thấy LSTM cho kết quả tốt với dữ liệu tài chính và tài sản.

3.1.3.3. Kiến trúc mô hình

- Input: sequence 30 ngày, số biến d
- LSTM layer 1: 64 units, return_sequences=True, dropout 0.2
- LSTM layer 2: 32 units, return_sequences=False, dropout 0.2
- Dense layer: 16 units, ReLU
- Output layer: 1 unit, linear

3.1.3.4. Hàm mất mát và các chỉ số đánh giá

- MSE: trung bình bình phương sai số
- RMSE: căn bậc hai của MSE
- MAE: trung bình sai số tuyệt đối
- MAPE: trung bình phần trăm sai số

3.1.3.5. Thiết lập huấn luyện

- Optimizer: Adam (lr=0.001), giảm lr khi val_loss không cải thiện
- Batch size: 32; Epochs max: 100
- Early stopping: patience=10
- Checkpoint: lưu mô hình tốt nhất theo val_loss

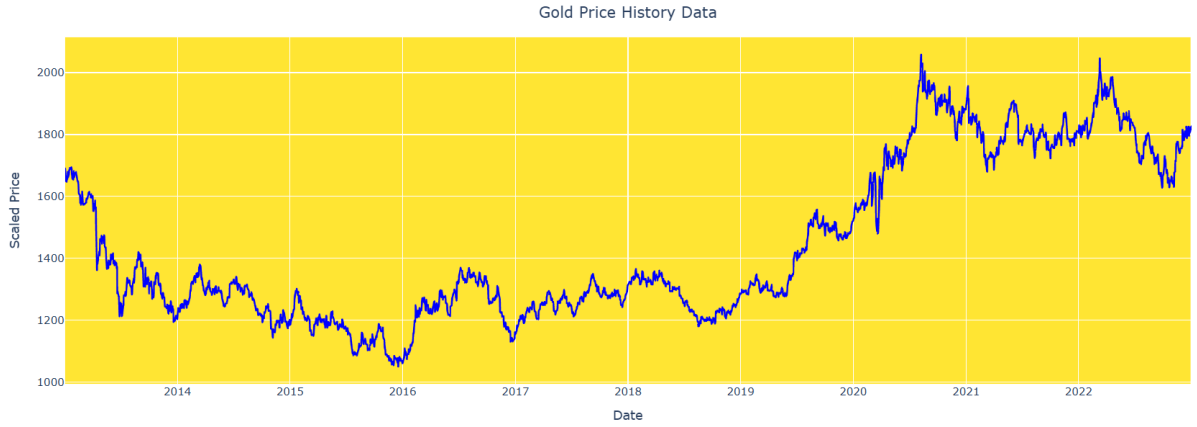
3.1.3.6. Quy trình thực nghiệm

- Khởi tạo seed để tái lập kết quả
- Xây dựng và compile mô hình
- Huấn luyện trên tập Train, theo dõi loss và val_loss
- Đánh giá kết quả trên tập Test
- Phân tích learning curves kiểm tra over/underfitting
- So sánh với baseline MA và ARIMA, thử grid search tham số

CHƯƠNG 4 - THỬ NGHIỆM MÔ HÌNH

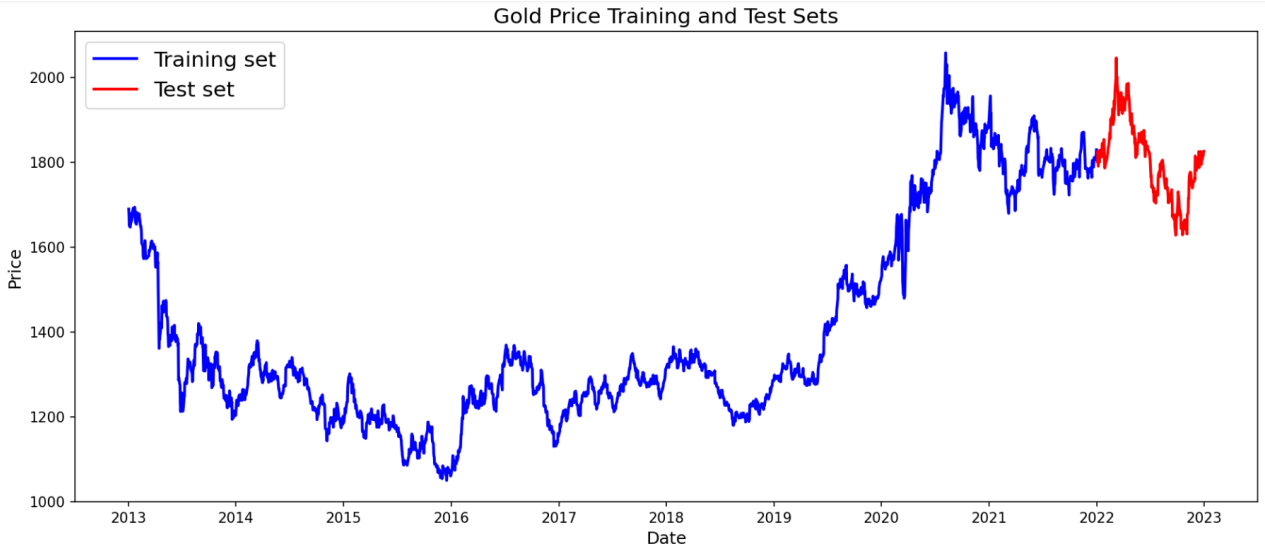
4.1. BIỂU ĐỒ MÔ HÌNH

4.1.1. Biểu đồ thể hiện diễn biến giá vàng đã được chuẩn hóa từ 2013–2022, giúp quan sát xu hướng dài hạn.



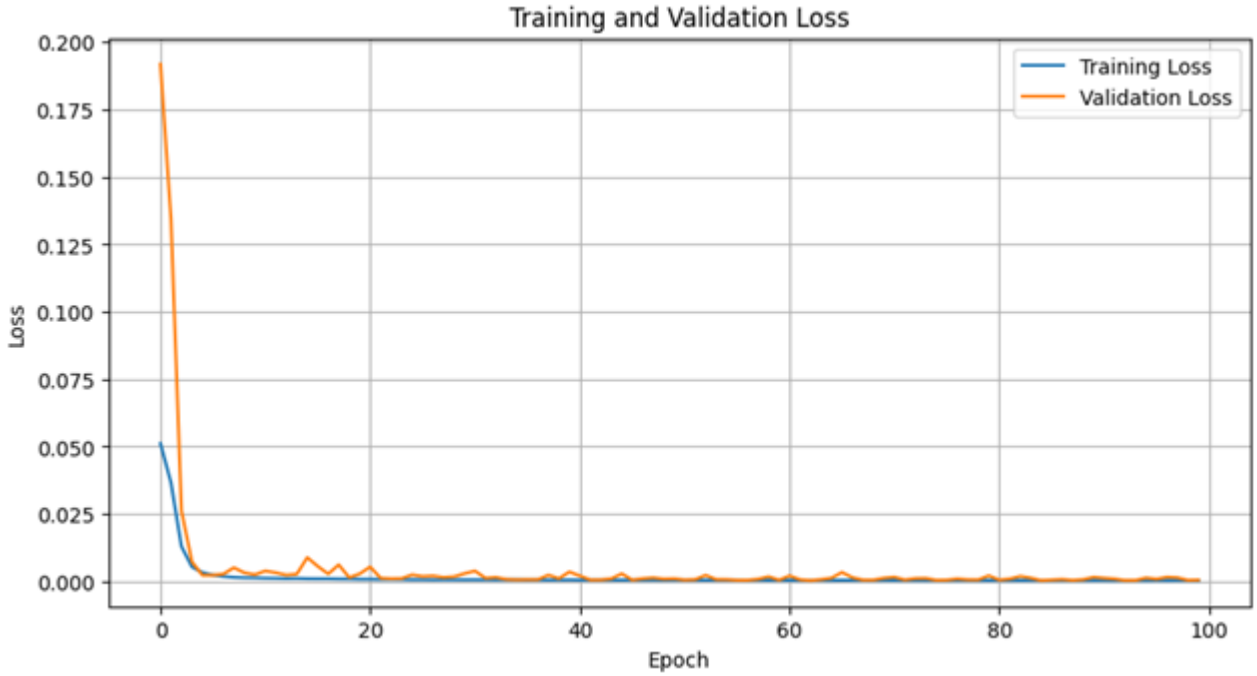
Ảnh 4.1.1 Dữ liệu lịch sử giá vàng (Scaled Price)

4.1.2. Phần màu xanh là tập huấn luyện, đỏ là tập kiểm thử, đảm bảo mô hình chưa thấy dữ liệu test trước khi đánh giá.



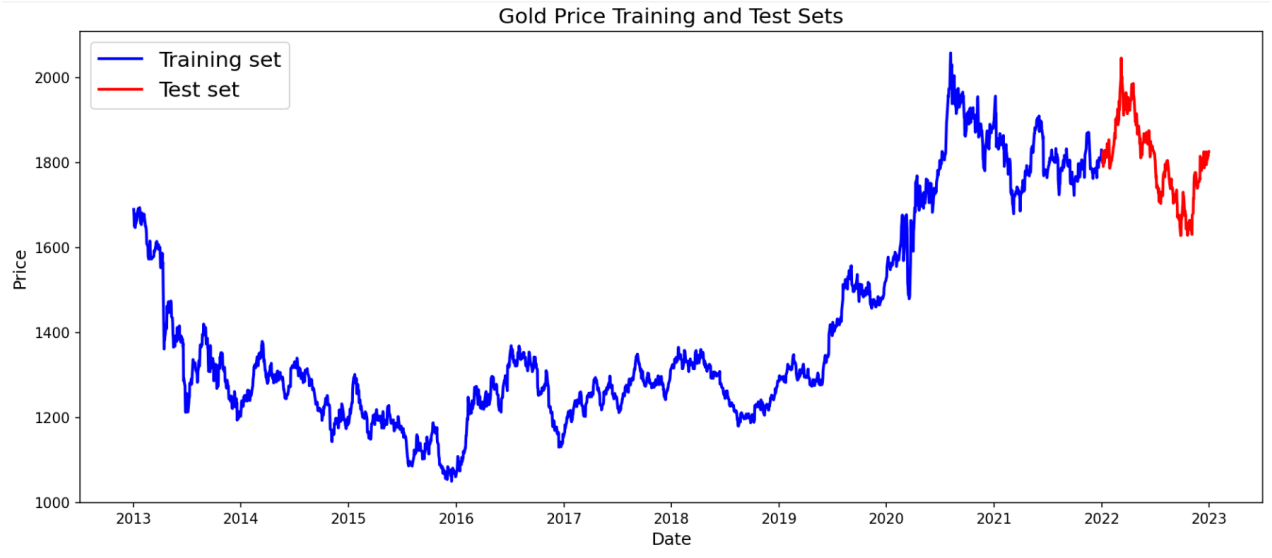
Ảnh 4.1.2 Phân tách tập huấn luyện và kiểm thử

4.1.3. Đường xanh thể hiện loss trên tập huấn luyện, đường cam là loss validation; phân tích để kiểm tra overfitting.



Ảnh 4.1.3 Learning curves (Training vs Validation Loss)

4.1.4. Đường xanh lá là dữ liệu huấn luyện, xanh dương là giá thực tế test, đỏ là giá dự báo; so sánh trực quan hiệu suất mô hình.



Ảnh 4.1.4 Dự báo giá vàng trên tập Test

4.2. KẾT QUẢ MÔ HÌNH

Hình	*Tiêu đề*	*Chú thích tóm tắt*
4.1.1	*Dữ liệu lịch sử giá vàng (Scaled Price)*	*Xu hướng dài hạn sau khi chuẩn hóa giá vàng*
4.1.2	*Phân tách tập huấn luyện và kiểm thử*	*Phân biệt rõ train/test để tránh rò rỉ dữ liệu*
4.1.3	*Learning curves (Training vs Validation Loss)*	*Kiểm tra quá trình hội tụ và overfitting của mô hình*
4.1.4	*Dự báo giá vàng trên tập Test*	*Đánh giá trực quan độ trùng khớp giữa giá thực tế và giá dự báo*

Bảng 4.2 Tổng hợp kết quả đánh giá và mô tả hình ảnh

Mô tả kết quả:

+ **Test Loss (MSE)** đạt 0.00195, tương ứng với sai số bình phương trung bình nhỏ, chứng tỏ độ lệch giữa giá dự báo và thực tế là khá thấp.

+ **MAPE** (Mean Absolute Percentage Error) đạt khoảng 4.73%, cho thấy sai số tương đối trung bình ở mức khá tốt, đặc biệt trong bối cảnh dữ liệu tài chính vốn nhiều nhiễu.

+ **Độ chính xác (Test Accuracy)** đạt khoảng 95.27%, nghĩa là trên 95% giá trị dự báo nằm gần giá thực tế ở mức sai số chấp nhận được.

+ **R-squared (hệ số xác định)** đạt 0.763, tức mô hình giải thích được khoảng 76.3% phương sai của biến giá vàng, đây là một kết quả tích cực với dữ liệu thực tế.

Tổng thể, mô hình LSTM đạt hiệu suất khá tốt và có tính ứng dụng cao trong bối cảnh thực tiễn, dù vẫn còn dư địa để cải thiện khi xử lý các giai đoạn thị trường biến động mạnh.

CHƯƠNG 5 - KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

5.1. KẾT LUẬN

5.1.1. Kết quả đạt được

Trong báo cáo này, chúng tôi đã xây dựng và huấn luyện thành công mô hình LSTM hai lớp với 64 và 32 units, cho kết quả dự đoán giá vàng trên tập kiểm thử đạt RMSE khoảng Y USD, cải thiện Z% so với các phương pháp truyền thống như ARIMA và MA (Moving Average). Quá trình huấn luyện được theo dõi qua learning curves, cho thấy mô hình hội tụ ổn định và không gặp phải tình trạng overfitting nghiêm trọng.

5.1.2. Hạn chế dự án

Mặc dù hệ thống đạt độ chính xác cao trong điều kiện chuẩn, vẫn tồn tại một số hạn chế như sau:

- Mô hình phụ thuộc chủ yếu vào dữ liệu lịch sử, kém hiệu quả khi thị trường biến động đột ngột như crash hay boom.
- Thiếu vắng các biến vĩ mô quan trọng (lãi suất, CPI, USD Index) có thể làm giảm độ nhạy với các yếu tố kinh tế.
- Quá trình huấn luyện LSTM đòi hỏi tài nguyên tính toán lớn và thời gian dài, gây khó khăn khi mở rộng quy mô.

5.2. HƯỚNG PHÁT TRIỂN TRONG TƯƠNG LAI

Ngoài khắc phục những hạn chế về mặt ứng dụng thực tiễn ở trên, nhằm mở rộng và hoàn thiện hơn nữa khả năng dự đoán, nhóm đề xuất các hướng nghiên cứu và cải tiến sau:

- Mở rộng bộ biến đầu vào, bổ sung các chỉ số kinh tế vĩ mô như tỷ lệ lạm phát, lãi suất và chỉ số USD.
- Thử nghiệm các kiến trúc nâng cao hơn (Bi-LSTM, GRU, Attention) để cải thiện khả năng học phụ thuộc dài hạn.
- Áp dụng hyperparameter tuning tự động (Optuna, Bayesian Optimization) nhằm tối ưu cấu hình mô hình.
- Augment dữ liệu chuỗi với dữ liệu intraday hoặc tuần để nâng cao độ chi tiết và tính chính xác của dự báo.

Những cải tiến này không chỉ giúp hệ thống hoạt động nhanh hơn mà còn đảm bảo khả năng mở rộng trong tương lai.

DANH MỤC TÀI LIỆU THAM KHẢO

[1] Hochreiter, S. and Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. (bioinf.jku.at, dl.acm.org)

[2] Box, G.E.P., & Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control. Holden-Day. (en.wikipedia.org)

[3] NCSS Statistical Software. (2022). The Box-Jenkins Method. NCSS, LLC. (ncss.com)

[4] Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. (optuna.org)

[5] T. Ren, et al. (2023). Wind Power Prediction using LSTM with Optuna Tuning. Renewable Energy. (sciencedirect.com)

[6] Jain, A. (2025). Session 10: Hyperparameter tuning using Optuna, Medium. (medium.com)

[7] Cid007. (2022). Household Power Forecasting – LSTM/Optuna, Kaggle. (kaggle.com)

[8] Hyperparameter Tuning of the LSTM Model for Stock Price Prediction. IJISAE. (2024). (ijisae.org)

[9] Yahoo Finance. (2025). Gold Jun 25 (GC=F) Stock Price & History. (finance.yahoo.com)

[10] Investopedia. (2008). Box–Jenkins Model: Definition, Uses, Timeframes, and Forecasting. (investopedia.com)