**Báo cáo LSTM cho Dự Báo Giá Cổ Phiếu Tesla**

**1. Lý do sử dụng mô hình LSTM:**

* **Xử lý dữ liệu chuỗi thời gian**: LSTM (Long Short-Term Memory) là một loại mạng nơ-ron hồi tiếp (RNN) đặc biệt hiệu quả cho dữ liệu chuỗi thời gian nhờ vào khả năng ghi nhớ thông tin trong thời gian dài và tránh vấn đề biến mất gradient.
* **Phù hợp với dữ liệu tài chính**: Mô hình này có thể xử lý sự phụ thuộc dài hạn trong dữ liệu giá cổ phiếu và các biến ngoại sinh như chỉ số S&P 500 và NASDAQ (IXIC), giúp cải thiện độ chính xác dự báo.

**2. Các kiểm thử cần thiết khi chạy mô hình:**

* **Phân chia dữ liệu**: Chia dữ liệu thành tập huấn luyện và kiểm tra (thường là 80-20) để đảm bảo mô hình được huấn luyện tốt và kiểm thử tính hiệu quả trên dữ liệu chưa từng thấy.
* **Kiểm thử hiệu suất**: Sử dụng các chỉ số như RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), và MAPE (Mean Absolute Percentage Error) để đánh giá hiệu suất mô hình trên tập kiểm tra.
* **Đánh giá tính ổn định**: Chạy nhiều lần với các tập dữ liệu khác nhau để đảm bảo tính ổn định của mô hình.

**3. Làm sao biết mô hình sẵn sàng để dự báo: ( file Test-LSTM)**

* **Đạt tiêu chí hiệu suất**: Mô hình cần có giá trị RMSE và MAPE đủ thấp trên tập kiểm tra (thường là < 10% cho MAPE được coi là tốt).   
  Test RMSE: 15.60%  
  Test MAPE: 5.32%  
  Test MAE: 11.08
* **Đường dự báo hợp lý**: Kiểm tra đồ thị của dữ liệu thực tế so với dự báo để đảm bảo rằng mô hình không quá khớp (overfitting) hoặc chưa khớp (underfitting).
* **Phân tích dư thừa (Residual Analysis)**: Dùng phân tích dư thừa để kiểm tra xem liệu các sai số có phân phối ngẫu nhiên và không còn mô hình nào khác có thể dự đoán được sai số đó.

A graph and a diagram

Description automatically generated A graph of a graph of a plot of residuals

Description automatically generated with medium confidence

**4. Các bước chuẩn bị trước khi dự báo:**

* **Tối ưu siêu tham số**: Sử dụng công cụ như Optuna để tối ưu hóa các siêu tham số như số lượng units LSTM, dropout rate, learning rate, số epochs, và batch size.

**Dùng file Run để làm 2 bước cuối**

* **Huấn luyện cuối cùng**: Huấn luyện lại mô hình với toàn bộ tập dữ liệu (kết hợp tập huấn luyện và kiểm tra) với các siêu tham số tối ưu đã tìm được.
* **Dự đoán thử**: Thực hiện dự đoán trên một đoạn dữ liệu có sẵn để xác nhận mô hình hoạt động đúng như mong đợi trước khi dự báo đoạn không có dữ liệu.  
  A graph with blue and orange lines

  Description automatically generated

Kết quả mô hình:

A graph showing the price of a stock market

Description automatically generated