Final Project - Machine Learning

Stock Price Prediction Using MLP and LSTM Neural Networks

By Fahad Almheiri - fb2160

By Fahad Almheiri - fb2160

Problem Definition

Goal: Predict the next day's closing stock price.

Why?: Useful for trading insights, risk assessment.

Input: A sequence of the previous lookback of 7 days' closing prices.

Output: The predicted closing price for the next day.

Task Type: Regression

This is a popular problem in finance and machine learning, known to be very challenging due to the often unpredictable and complex nature of stock market movements.

By Fahad Almheiri - fb2160

Dataset & Pre-Processing

Dataset

Dataset Source: Historical daily stock data downloaded using yfinance library.

Data Size: Depending on the stock ticker chosen for this presentation we will be using NIKE stock which has 11191 rows goes back to December 2, 1980!

Price	Close	Close(t-1)	Close(t-2)	Close(t-3)	Close(t-4)	Close(t-5)	Close(t-6)	Close(t-7)
Ticker	NKE							
Date								
2025-04-21	56.119999	55.759998	53.549999	54.830002	55.410000	54.389999	54.400002	59.320000
2025-04-22	57.060001	56.119999	55.759998	53.549999	54.830002	55.410000	54.389999	54.400002
2025-04-23	57.389999	57.060001	56.119999	55.759998	53.549999	54.830002	55.410000	54.389999
2025-04-24	58.480000	57.389999	57.060001	56.119999	55.759998	53.549999	54.830002	55.410000
2025-04-25	57.619999	58.480000	57.389999	57.060001	56.119999	55.759998	53.549999	54.830002

Preprocessing Steps

Selected 'Close' price only.

Created lagged features (used lookback=7 days).

Scaled data to [-1, 1] range using MinMaxScaler

By Fahad Almheiri - fb2160

Model Design

Backbones Chosen: MLP (simple) and LSTM (complex).

Justification:

MLP

Standard feedforward network; serves as a comparison to see if sequence modeling adds value with this feature set.

LSTM

Good for capturing temporal dependencies and patterns in sequential data like stock prices.

Loss function: The loss function used for training both the LSTM and the MLP models is Mean Squared Error (MSE).

Optimizer: Adam - Common adaptive learning rate optimizer.

By Fahad Almheiri - fb2160

Model Training

MLP

MLP Params: Hidden Layers=[32, 16], LR=0.001, Batch Size=16, Input Features=7 (Lookback)

LSTM

LSTM Initial Params: Hidden Size=4, Layers=1, LR=0.001, Dropout=0.0, Batch Size=16, Lookback=7

```
MLP Training Epoch: 1/10, Loss: 0.01570
MLP Training Epoch: 2/10, Loss: 0.00023
MLP Training Epoch: 3/10, Loss: 0.00018
MLP Training Epoch: 4/10, Loss: 0.00015
MLP Training Epoch: 5/10, Loss: 0.00013
MLP Training Epoch: 6/10, Loss: 0.00012
MLP Training Epoch: 7/10, Loss: 0.00012
MLP Training Epoch: 8/10, Loss: 0.00012
MLP Training Epoch: 9/10, Loss: 0.00011
MLP Training Epoch: 10/10, Loss: 0.00012
```

```
LSTM Training Epoch: 1/10, Loss: 0.02707
LSTM Training Epoch: 2/10, Loss: 0.00018
LSTM Training Epoch: 3/10, Loss: 0.00015
LSTM Training Epoch: 4/10, Loss: 0.00015
LSTM Training Epoch: 5/10, Loss: 0.00014
LSTM Training Epoch: 6/10, Loss: 0.00013
LSTM Training Epoch: 7/10, Loss: 0.00013
LSTM Training Epoch: 8/10, Loss: 0.00012
LSTM Training Epoch: 9/10, Loss: 0.00012
LSTM Training Epoch: 10/10, Loss: 0.00012
```

By Fahad Almheiri - fb2160

Improving the LSTM - Hyperparameter Tuning

Method

- Used Grid Search to systematically test different hyperparameter combinations for the LSTM model.
  - Tested 24 combinations in total.
  - Ran each combination for 5 epochs to evaluate validation loss fast.

```
Testing 24 combinations over 5 epochs each...
```

Regularization techniques

- **Method Explored:** Tested Dropout as a regularization technique.
- **Values Tested:** Included dropout probability of 0.0 (no dropout) vs. 0.2 (20% dropout) in the grid search.

```
==================================================
Grid Search Complete.
Best Validation Loss achieved: 0.00050
Best Parameters found: {'hidden_size': 4, 'num_stacked_layers': 1, 'learning_rate': 0.01, 'dropout_prob': 0.0}
==================================================
```

Parameters Tuned & Ranges Explored

- Hidden Size (hidden_size): [4, 8, 16]
- Stacked Layers (num_stacked_layers): [1, 2]
- Learning Rate (learning_rate): [0.01, 0.001]
- Dropout Probability (dropout_prob): [0.0, 0.2]

```python
# Define Hyperparameter Grid
param_grid = {
    'hidden_size': [4, 8, 16],
    'num_stacked_layers': [1, 2],
    'learning_rate': [0.01, 0.001],
    'dropout_prob': [0.0, 0.2]
}
num_search_epochs = 5
```

By Fahad Almheiri - fb2160

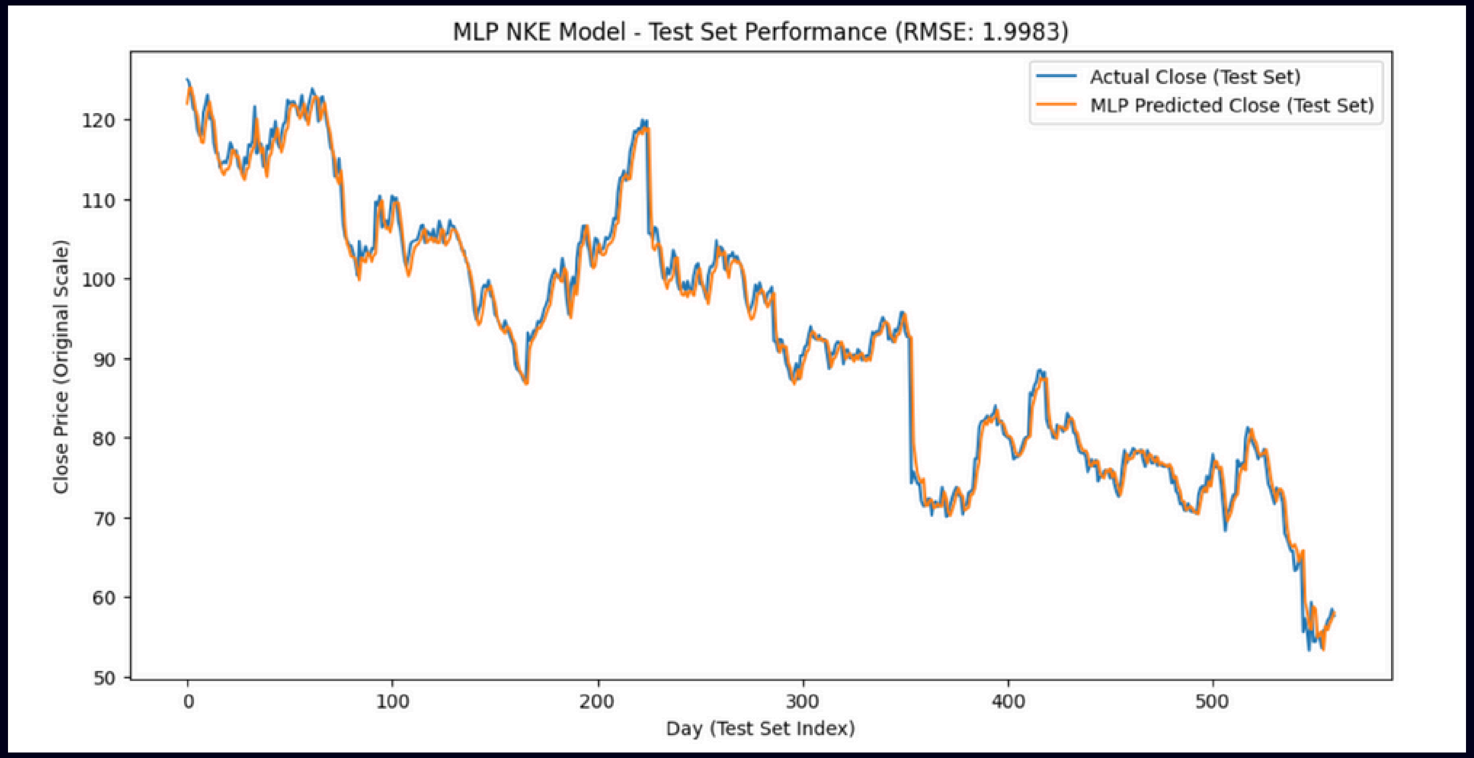
Evaluation and Comparison

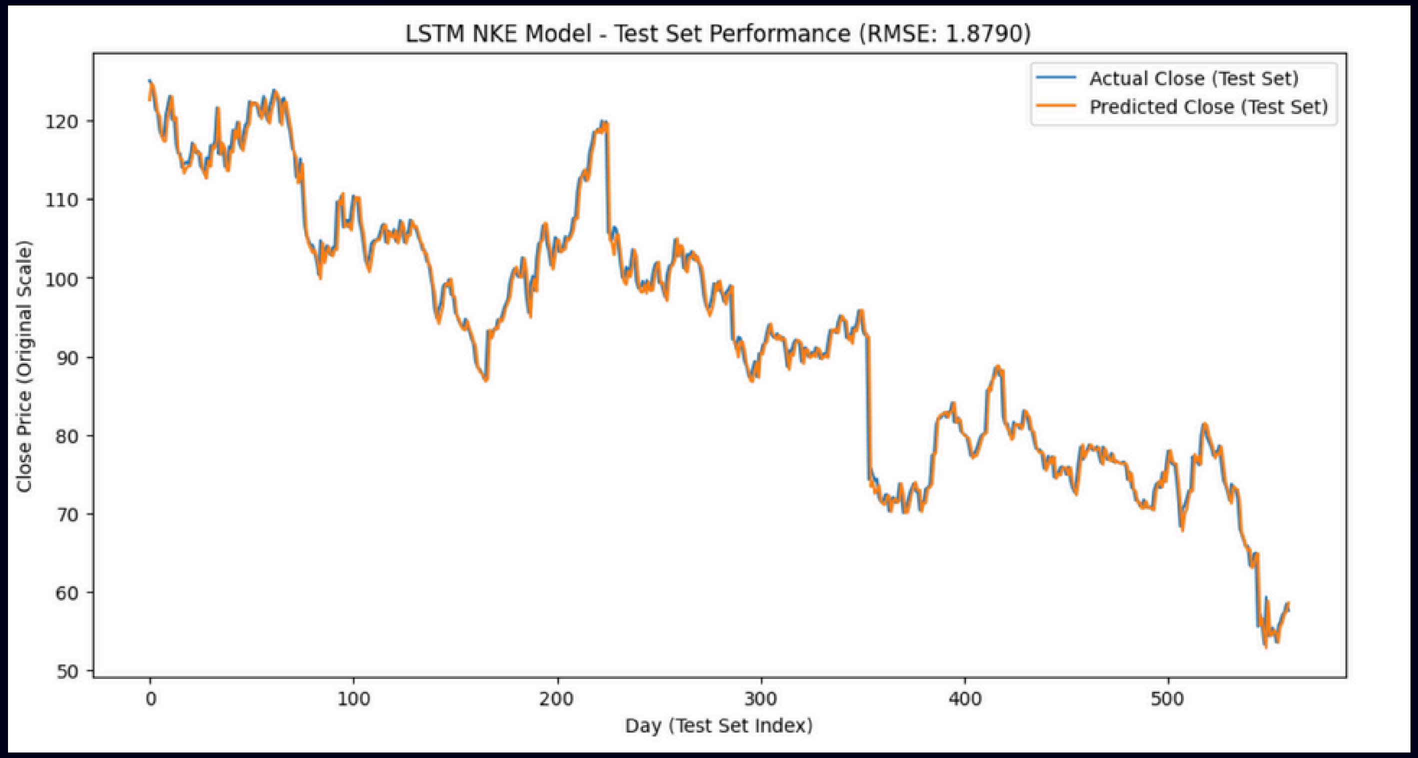
Performance Metric

- Used Root Mean Squared Error (RMSE) to evaluate final model performance.
- Calculated on the test set after inverse-transforming predictions and actual values back to the original dollar scale.
- Lower RMSE indicates better performance (smaller average prediction error).

Comparison of Backbones (NKE Ticker)

- LSTM Test RMSE: 1.8790
- MLP Test RMSE: 1.9983
- Result: For the NKE ticker with the current feature set (past 7 closing prices), the tuned LSTM model achieved a lower RMSE, indicating slightly better performance on the test data compared to the simpler MLP model.





Conclusion and Future Work

By Fahad Almheiri - fb2160

Conclusion

Best Model & Result

The tuned LSTM slightly outperformed the MLP (Test RMSE: 1.9983) on the unseen test data for the NKE ticker.

Future Work

Feature Engineering: Incorporate richer input features beyond just closing price

- Open, High, Low prices, and Volume (OHLCV).
- Calculated technical indicators (Moving Averages, RSI, MACD, etc.).
- News Sentiment

Improved Training Strategy:

- Utilize a dedicated validation set (split from training data).
- Train for more epochs combined with Early Stopping based on validation loss to prevent overfitting and find optimal training time.

Advanced Hyperparameter Tuning:

- Explore a wider range of parameters.
- More complex techniques.

Final Project - Machine Learning

THANK YOU.

By Fahad Almheiri - fb2160