

Boston University  
Electrical & Computer Engineering

EC 463 Senior Design Project

# AI Trading Platform

By: Team 6

First Prototype Test Report

Adrian Pawlowski | [ajp12@bu.edu](mailto:ajp12@bu.edu)  
Gabriel Brown | [gabrown@bu.edu](mailto:gabrown@bu.edu)  
Yagiz Idlman | [ykaan@bu.edu](mailto:ykaan@bu.edu)  
Odilon Quevillon | [odilon@bu.edu](mailto:odilon@bu.edu)

# 1. Introduction

This report evaluates the performance of LSTM and XGBoost models for predicting stock prices of TSLA (high-volatility), KO (medium-volatility), and MSFT (low-volatility) stocks. The objective is to assess their accuracy and robustness through predefined metrics such as Mean Squared Loss, RMSE, and Sharpe Ratio, among others.

# 2. Equipment and Setup

## Required Materials

- **Models:** LSTM and XGBoost.
- **Scripts:**
  - lstm\_model.py, train\_lstm.py, lstm\_model\_training\_prototype.ipynb, lstm\_model\_backtesting\_prototype.ipynb.
  - train\_boost.py, xgboost\_model\_training\_prototype.ipynb, xgboost\_backtesting\_prototype.ipynb.
- **Data Processing:** data\_processing.py.
- **Optimization Tool:** genetic\_algorithm.py.
- **Environment File:** Environment.yml.

## Setup

1. All scripts are run on a local machine within an Anaconda environment containing necessary Python packages.
2. Training data (January 1, 2008 – December 31, 2021) and backtesting data (January 1, 2021 – January 1, 2024) are fetched using Yahoo Finance.

# 3. Methodology

## Testing Procedure

1. Initialize the Anaconda environment with pre-downloaded packages.
2. Use the fetch\_stock\_data function to retrieve the desired ticker symbols and date ranges.
3. Execute:
  - a. LSTM scripts to evaluate predictive metrics.
  - b. XGBoost scripts to analyze classification metrics.
4. Evaluate models for TSLA, KO, and MSFT across specified metrics.

## 4. Measurable Criteria and Results

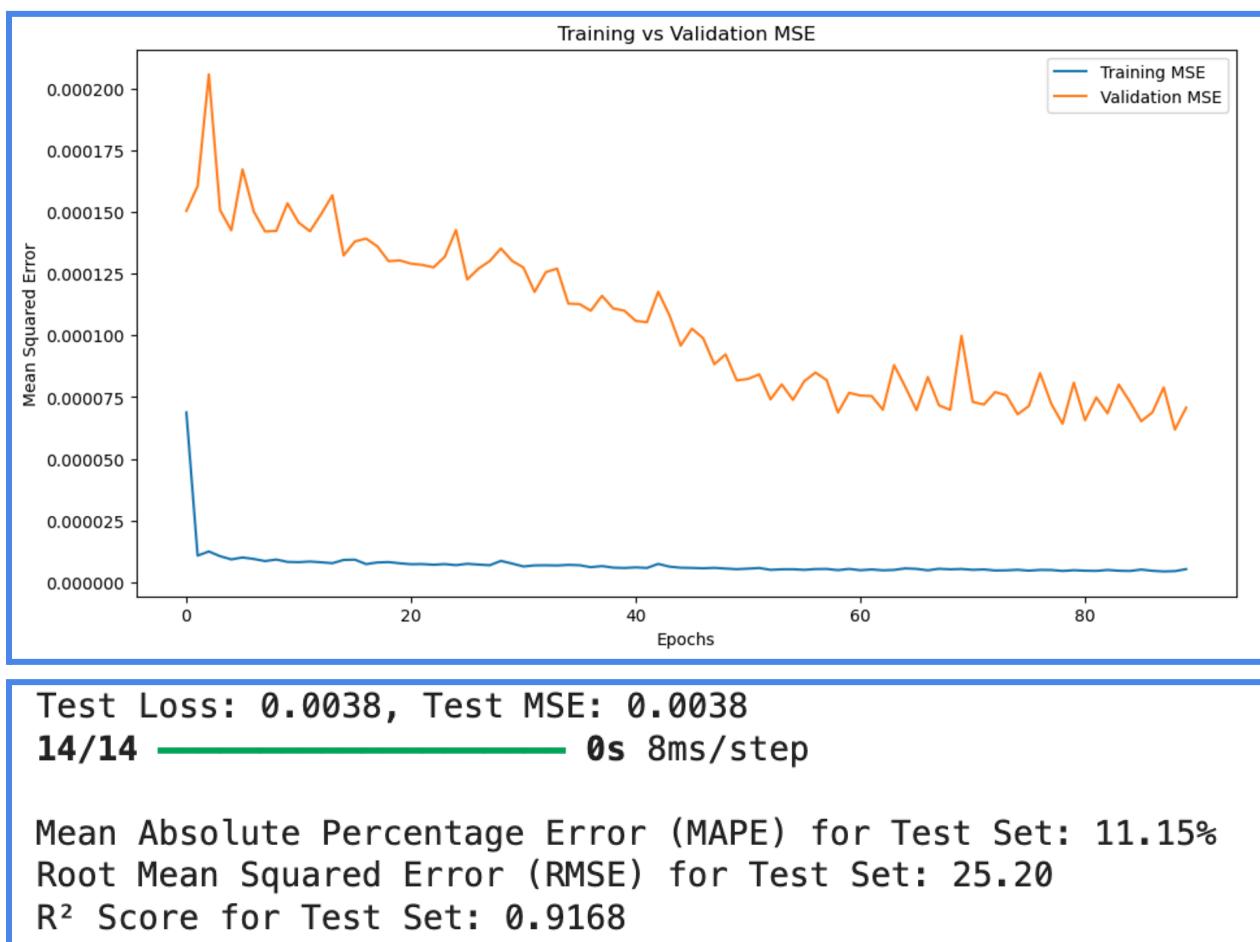
### 4.1 LSTM Model Results

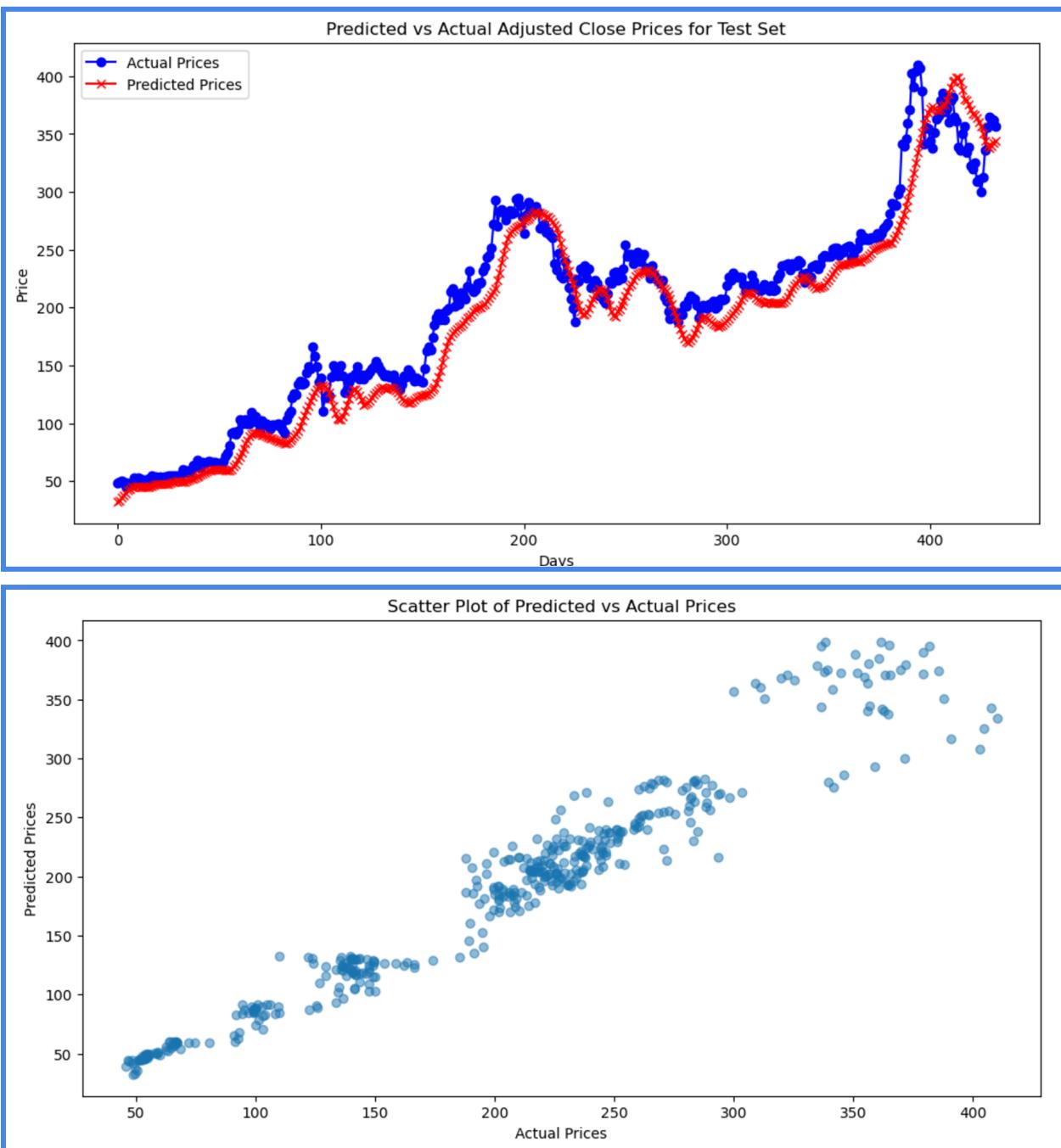
#### Metrics Evaluated:

- Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)
- Mean Absolute Percentage Error (MAPE)
- R-squared ( $R^2$ )
- Total Profit and Loss (P/L)
- Sharpe Ratio
- Maximum Drawdown

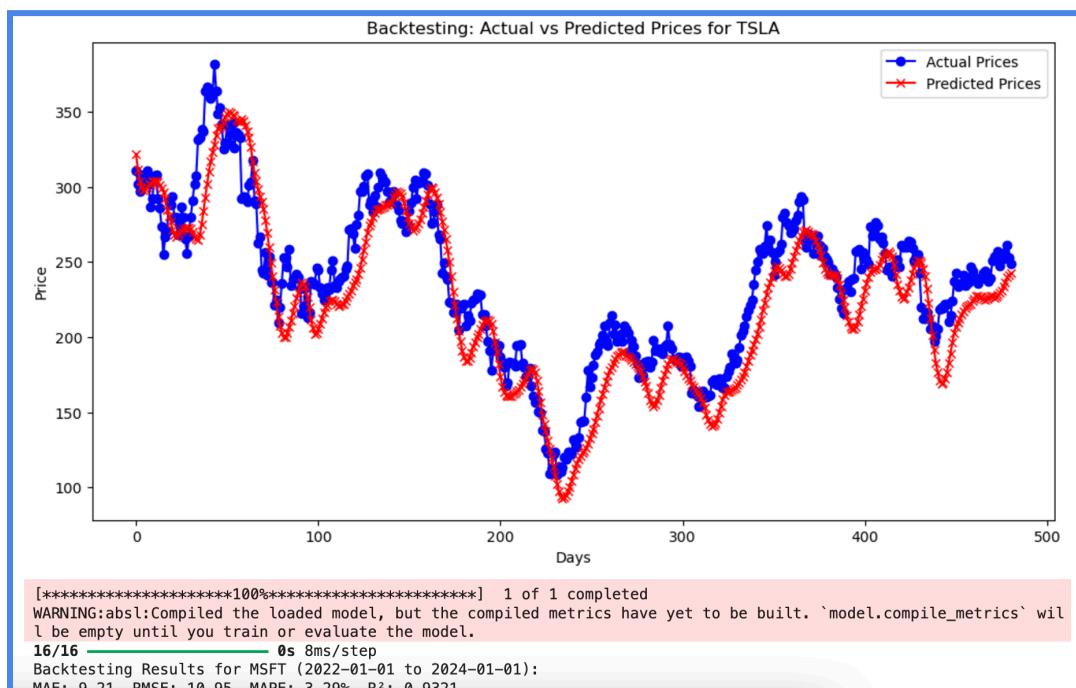
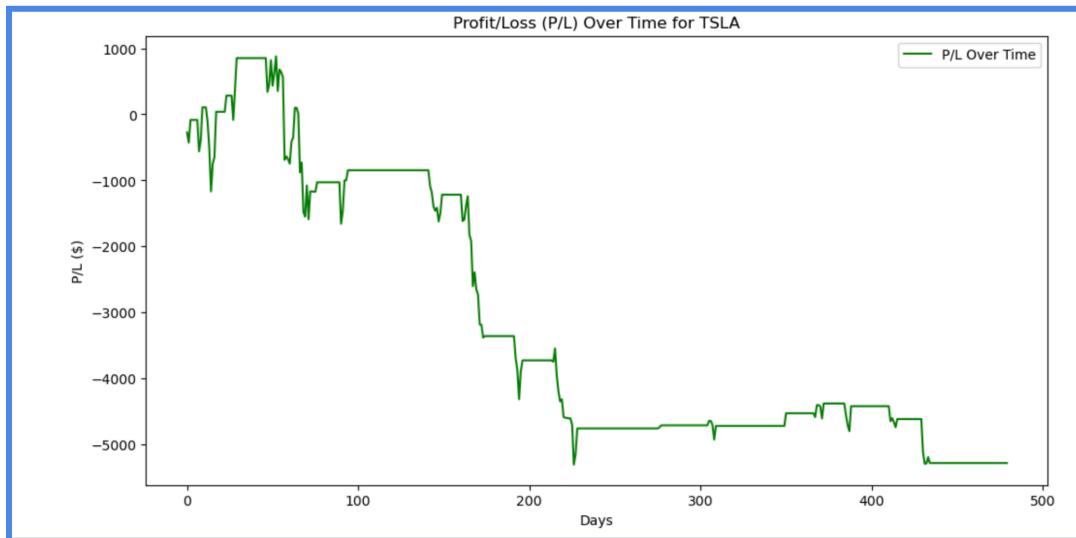
Tesla (TSLA)

#### TRAINING RESULTS:





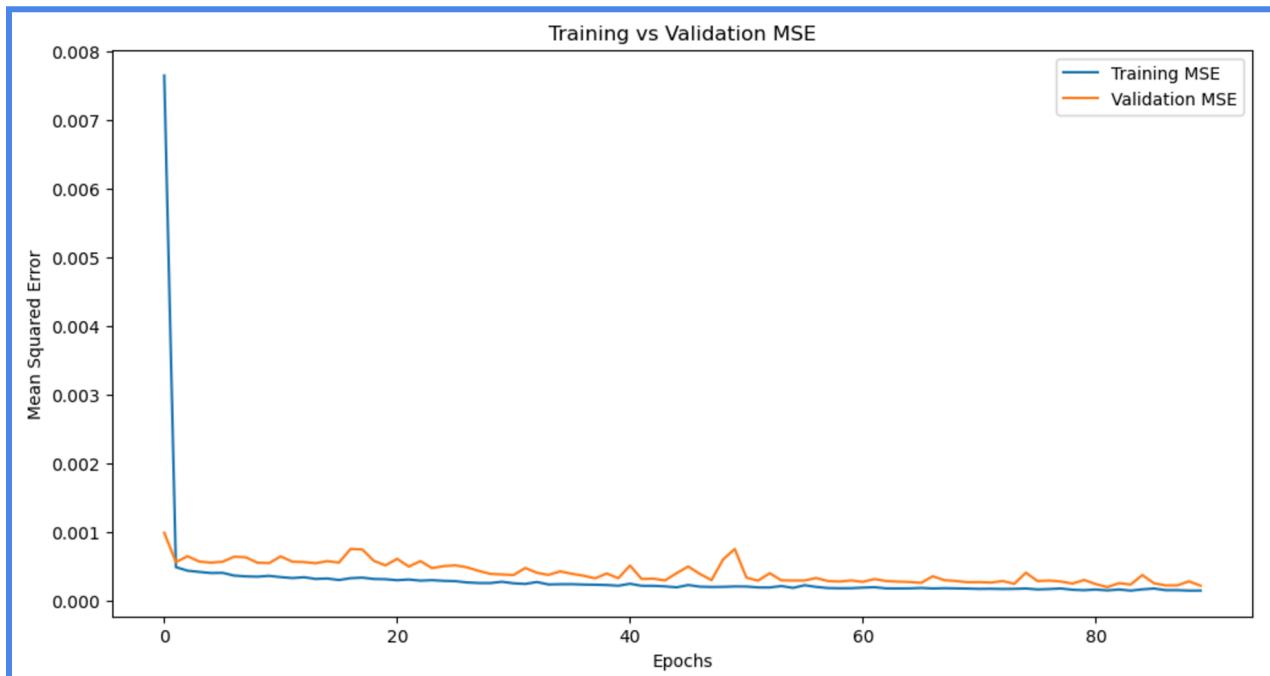
## BACKTESTING RESULTS:



Backtesting Results for TSLA (2022-01-01 to 2024-01-01):  
MAE: 19.99, RMSE: 24.28, MAPE: 8.67%, R<sup>2</sup>: 0.7853  
Total P/L: \$-5279.61, Sharpe Ratio: -1.57, Max Drawdown: \$-5302.89

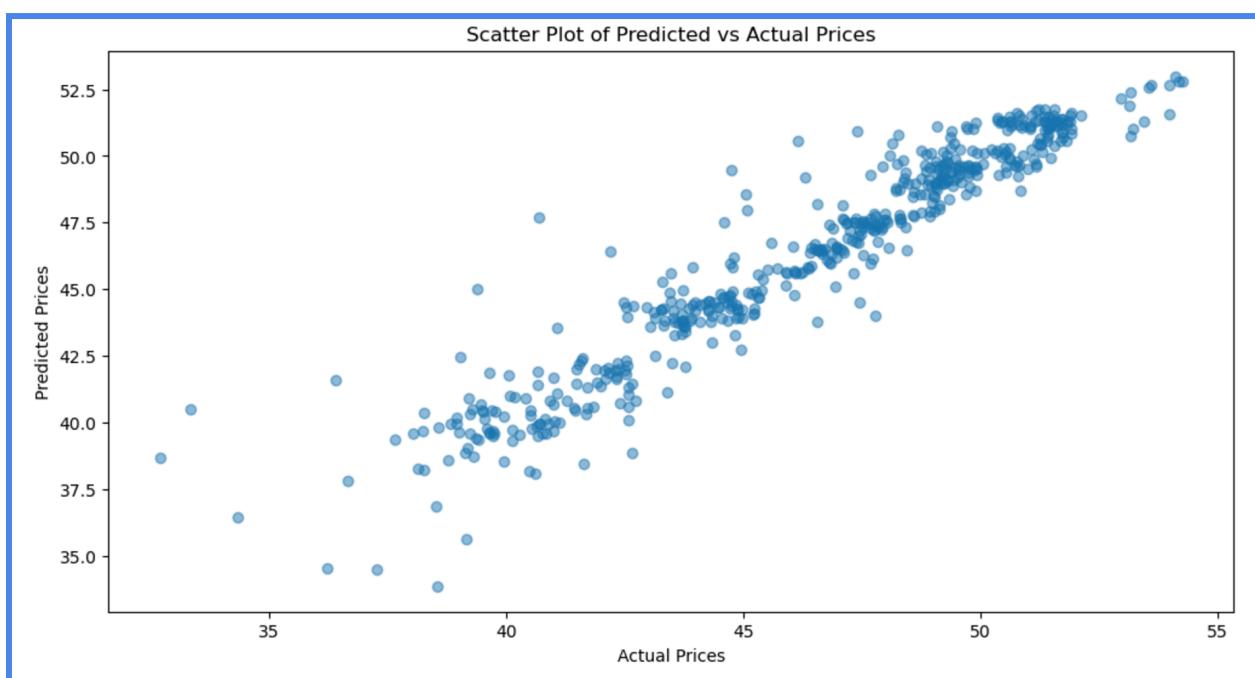
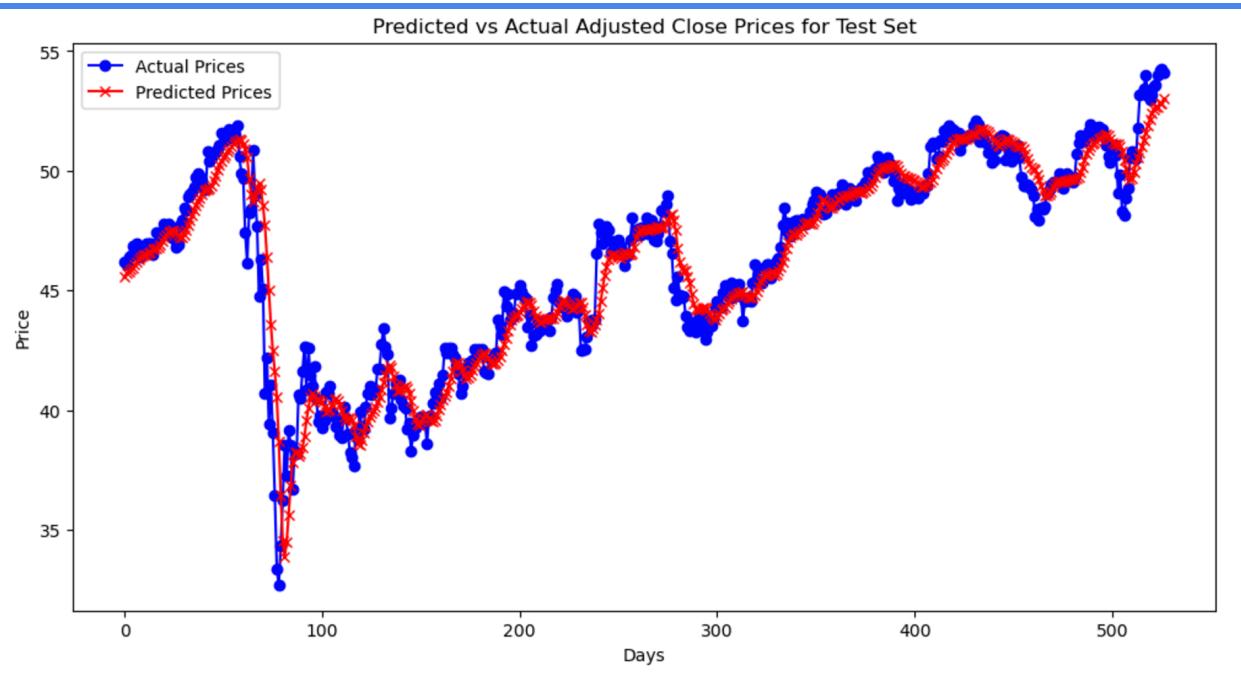
## The Coca-Cola Company (KO)

### TRAINING RESULTS:

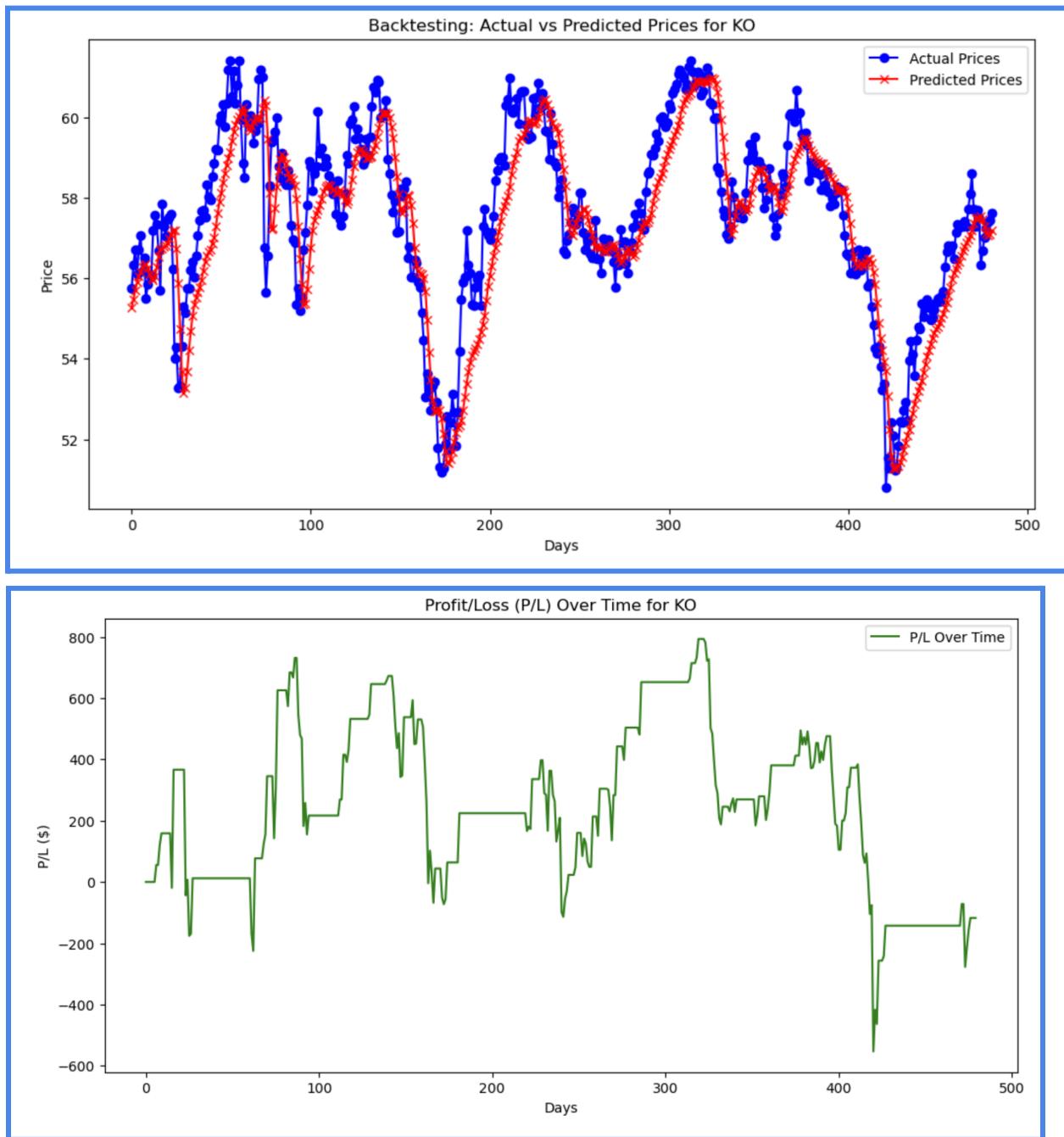


Test Loss: 0.0008, Test MSE: 0.0008  
17/17 ————— 0s 7ms/step

Mean Absolute Percentage Error (MAPE) for Test Set: 1.87%  
Root Mean Squared Error (RMSE) for Test Set: 1.24  
R<sup>2</sup> Score for Test Set: 0.9109



## BACKTESTING RESULTS:



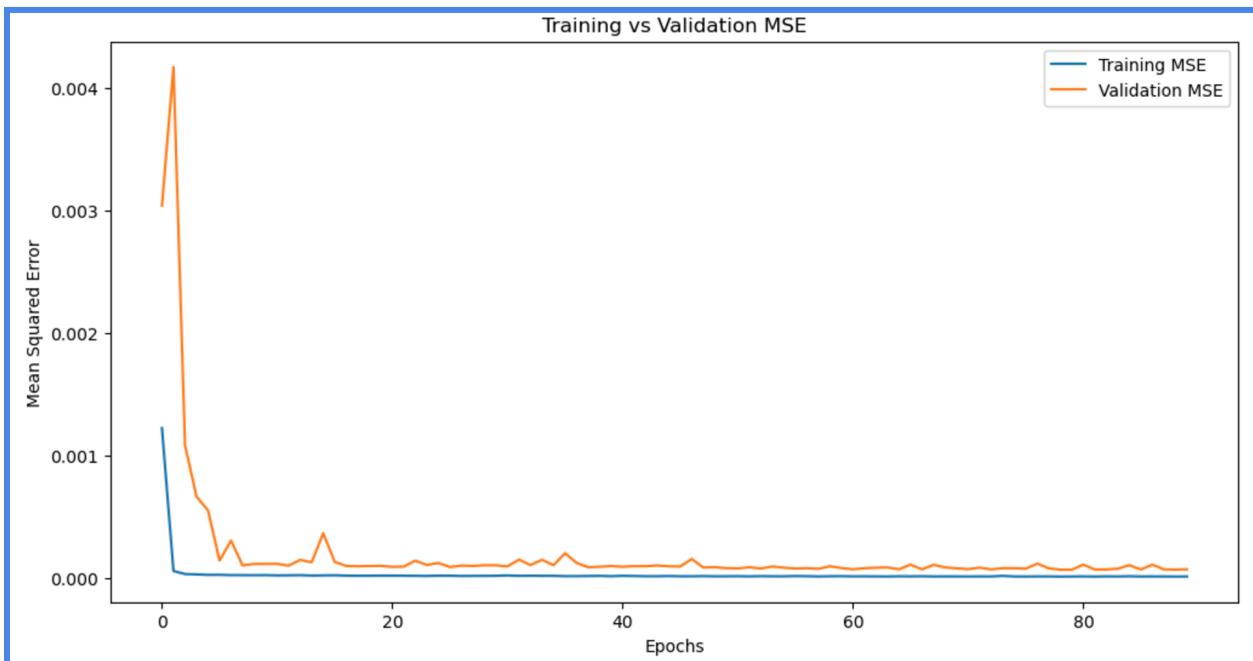
Backtesting Results for KO (2022-01-01 to 2024-01-01):

MAE: 0.91, RMSE: 1.15, MAPE: 1.59%, R<sup>2</sup>: 0.7418

Total P/L: \$-117.43, Sharpe Ratio: 0.94, Max Drawdown: \$-553.01

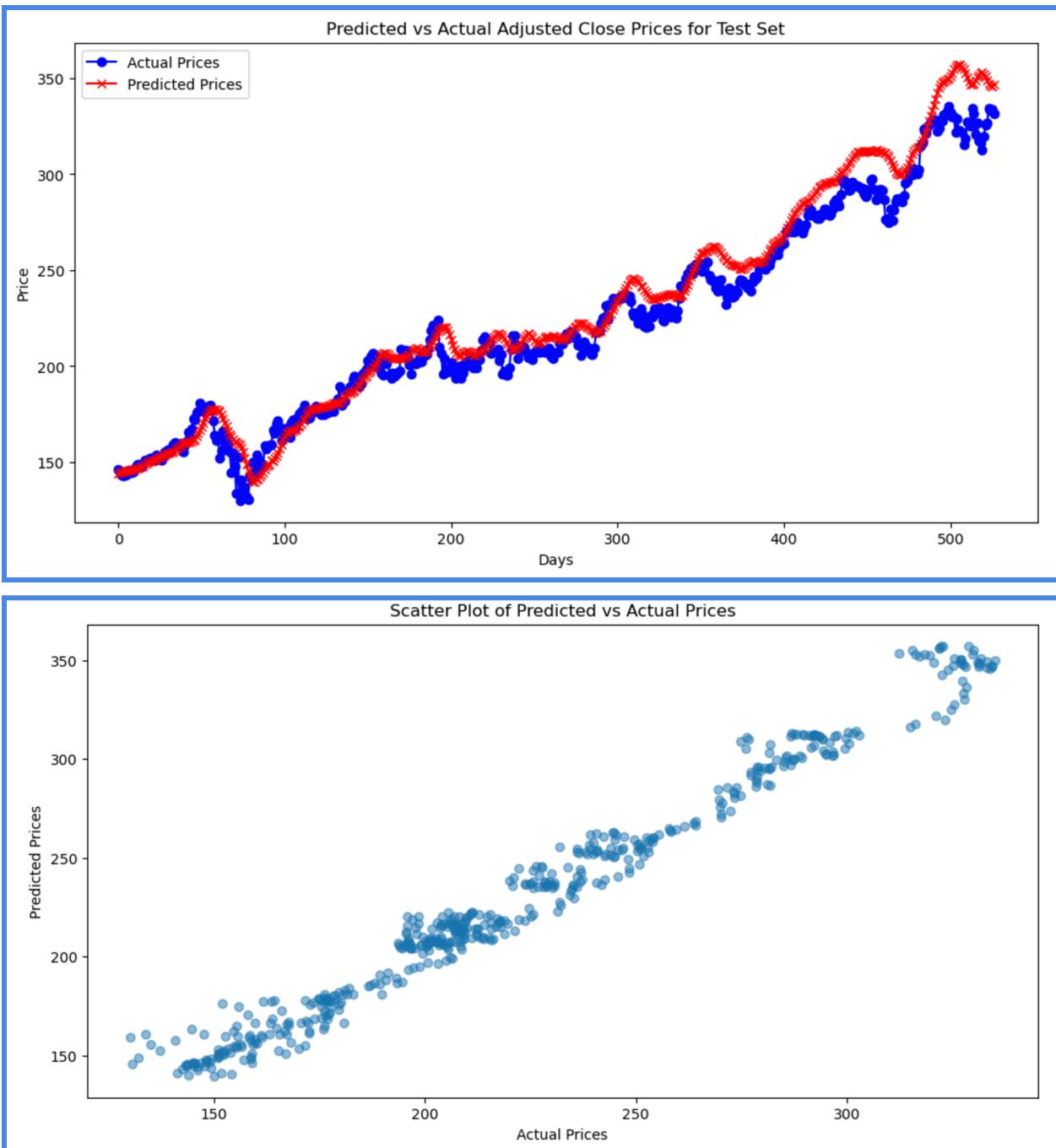
## Microsoft (MSFT)

### TRAINING RESULTS:

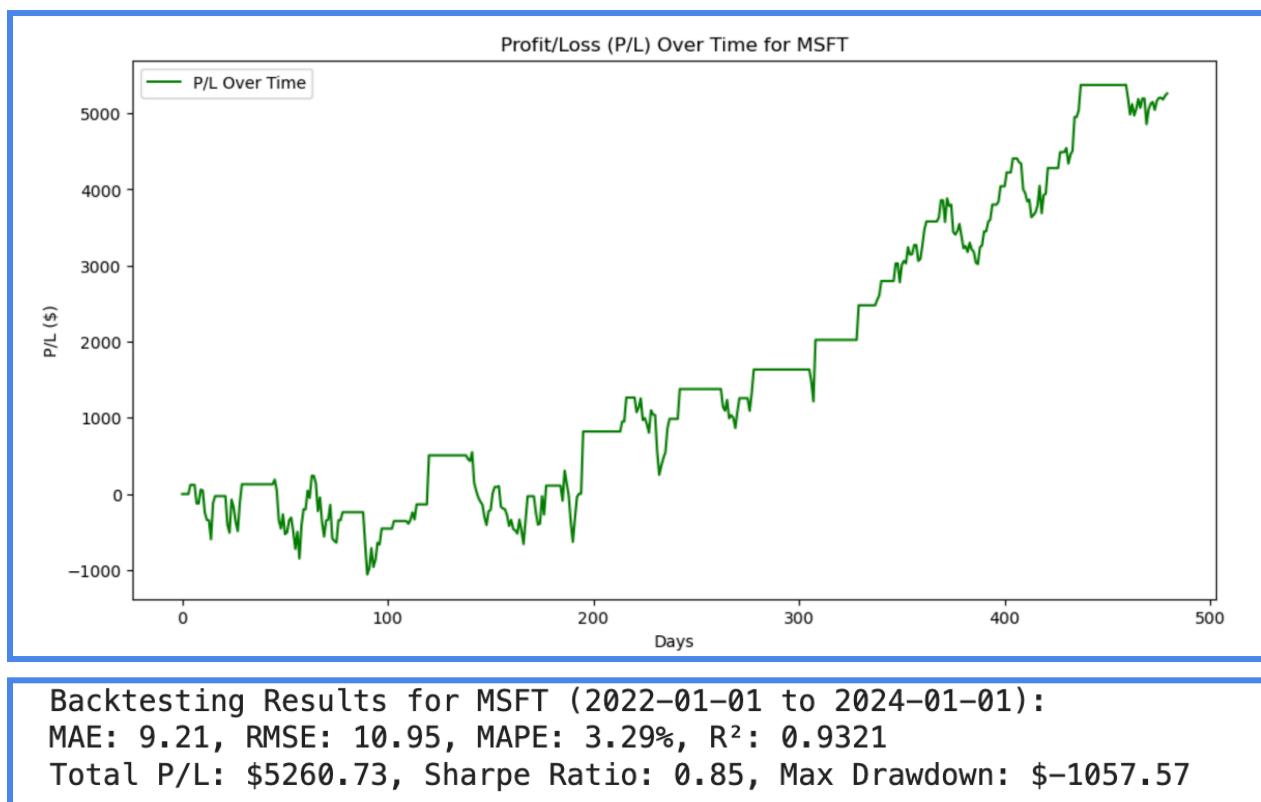
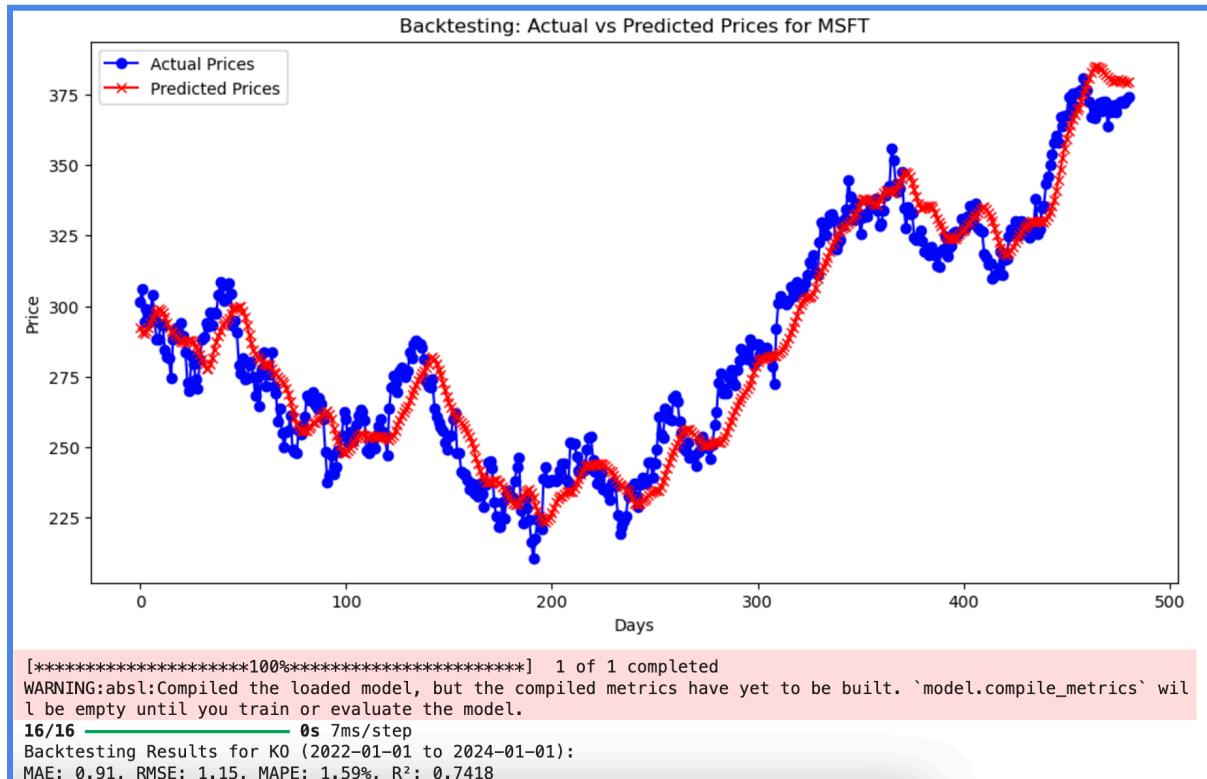


Test Loss: 0.0014, Test MSE: 0.0014  
17/17   0s 7ms/step

Mean Absolute Percentage Error (MAPE) for Test Set: 3.97%  
Root Mean Squared Error (RMSE) for Test Set: 11.98  
R<sup>2</sup> Score for Test Set: 0.9479



## BACKTESTING RESULTS:



## 4.2 XGBoost Model Results

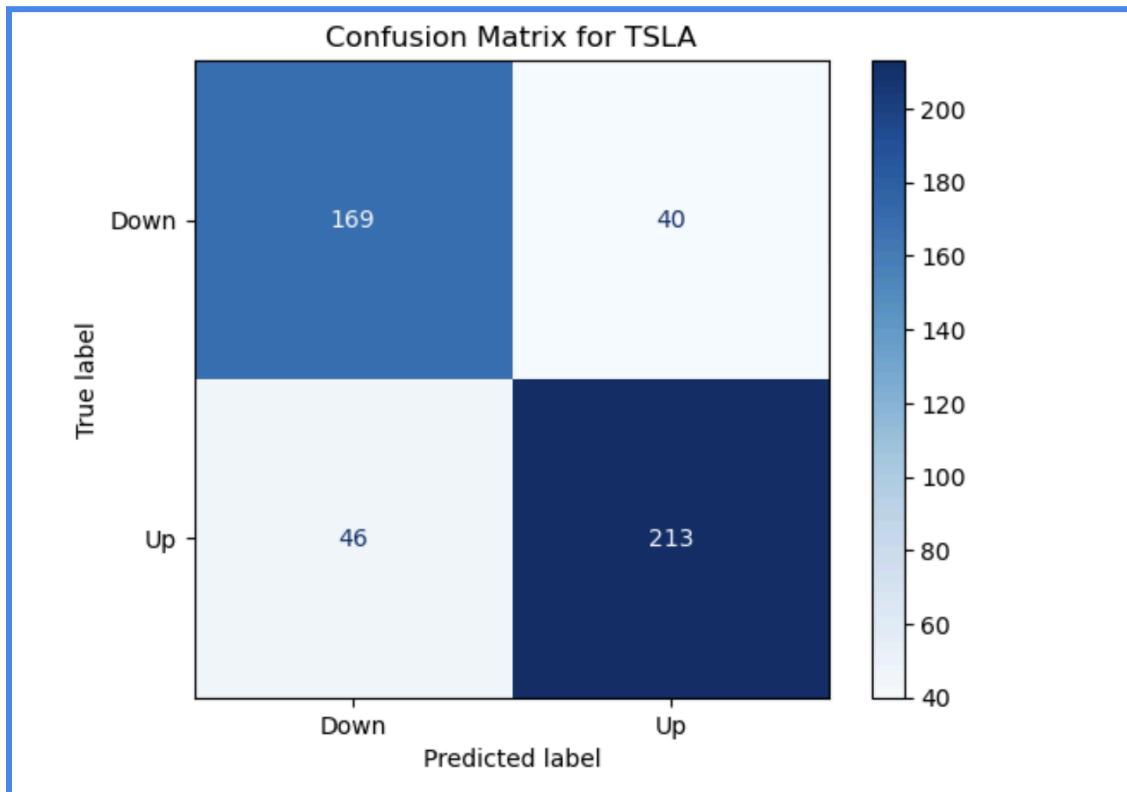
### Metrics Evaluated:

- Confusion Matrix
- Cumulative Profit and Loss (P/L)
- Accuracy
- Precision, Recall
- ROC\_AUC and F1 Score

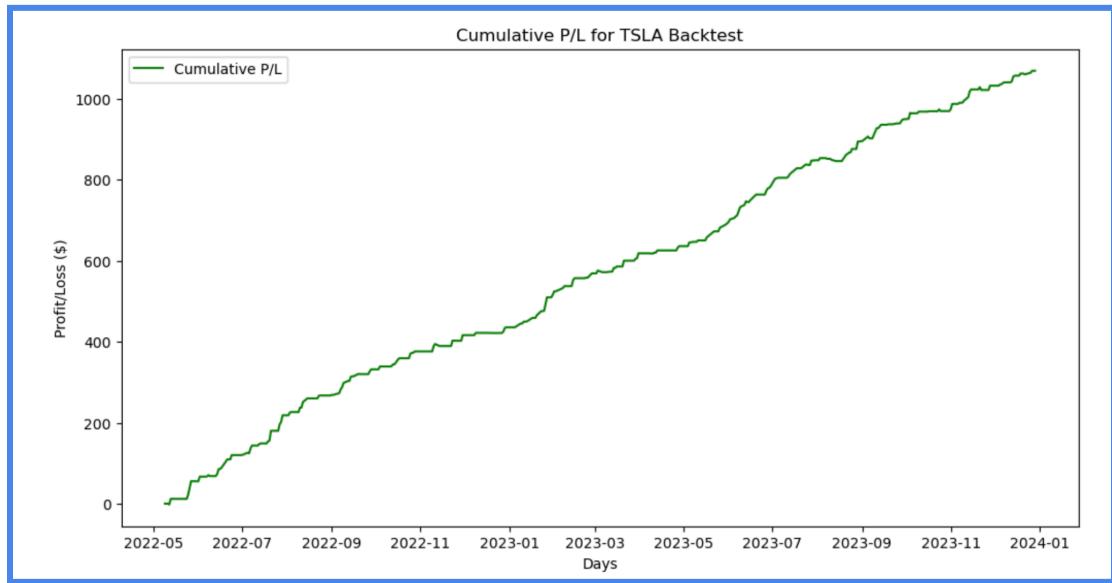
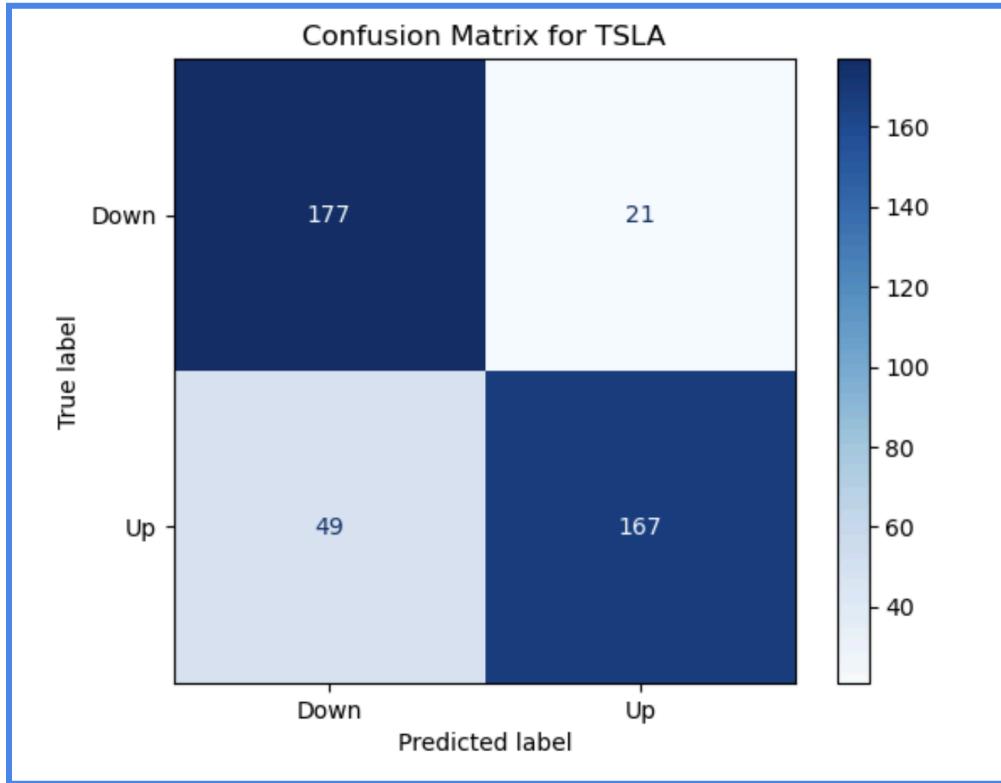
Tesla (TSLA)

- Threshold: Mean Squared Loss < 20%

### TRAINING RESULTS:



## BACKTESTING RESULTS:

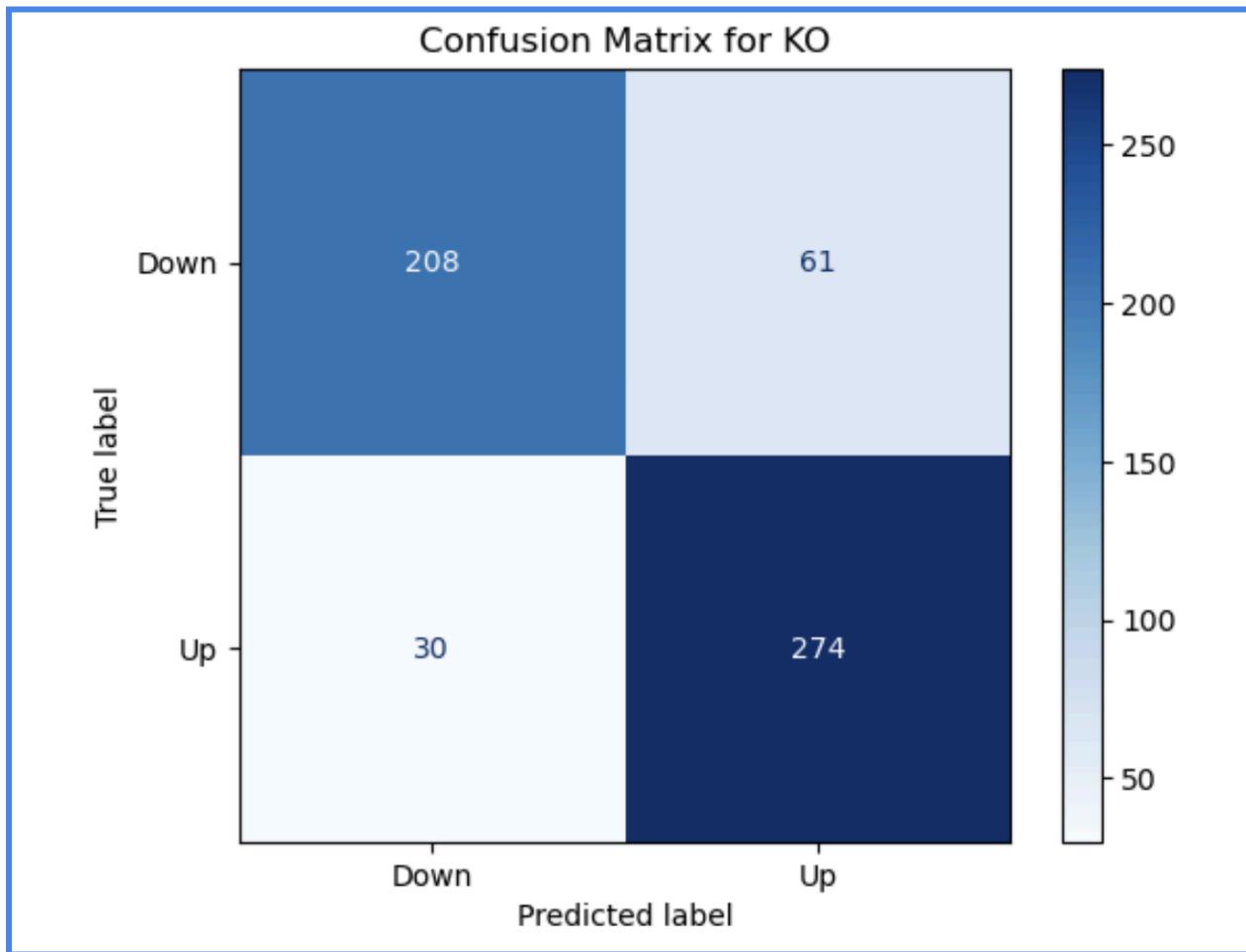


```
Backtesting Metrics: {'accuracy': 0.8309178743961353, 'precision': 0.8882978723404256, 'recall': 0.7731481481481481, 'f1_score': 0.8267326732673267, 'roc_auc': 0.9118967452300786, 'total_pnl': 1069.023338317871, 'avg_daily_pnl': 2.582181976613215, 'max_drawdown': -7.3300018310546875}
```

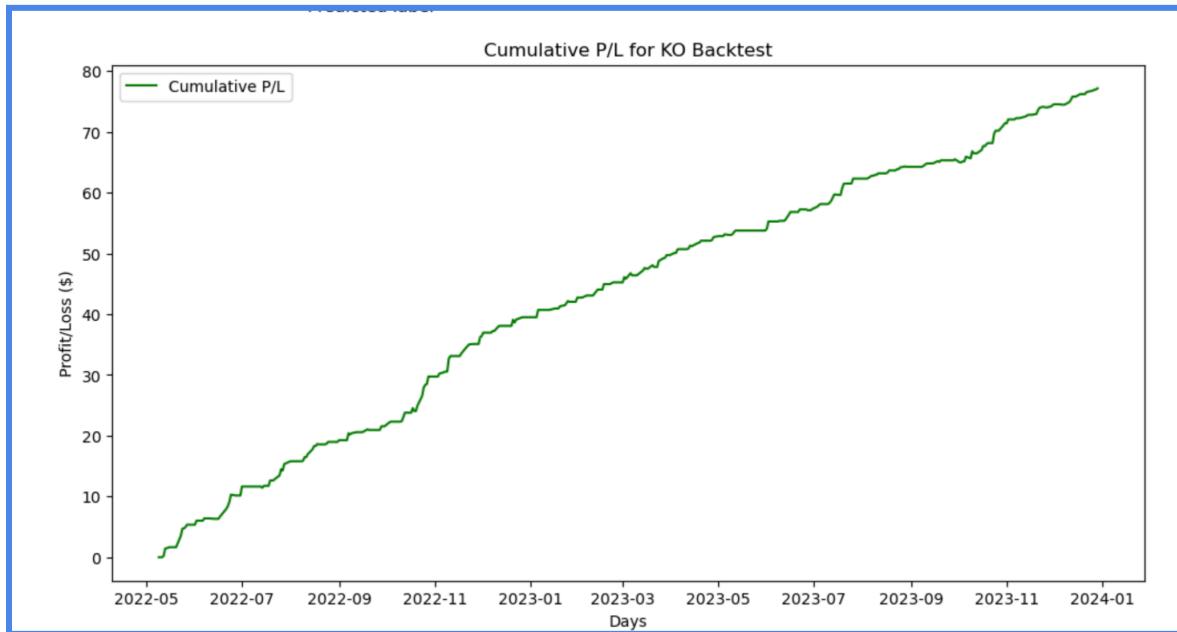
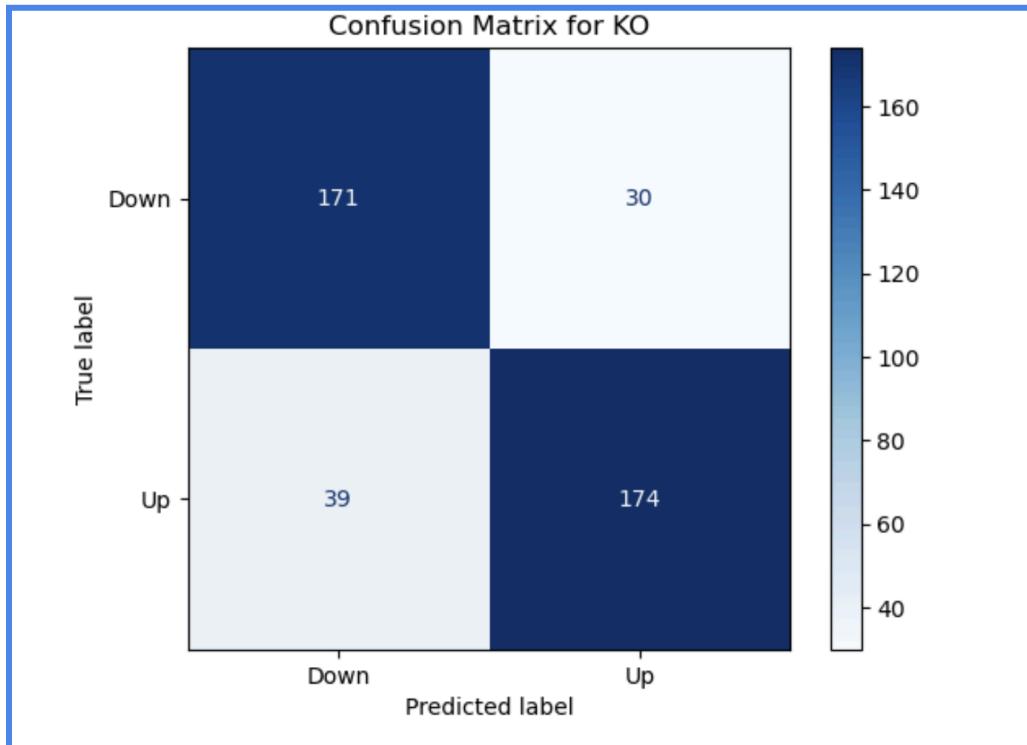
## The Coca-Cola Company (KO)

- Threshold: Mean Squared Loss < 17%

### TRAINING RESULTS:



## BACKTESTING RESULTS:

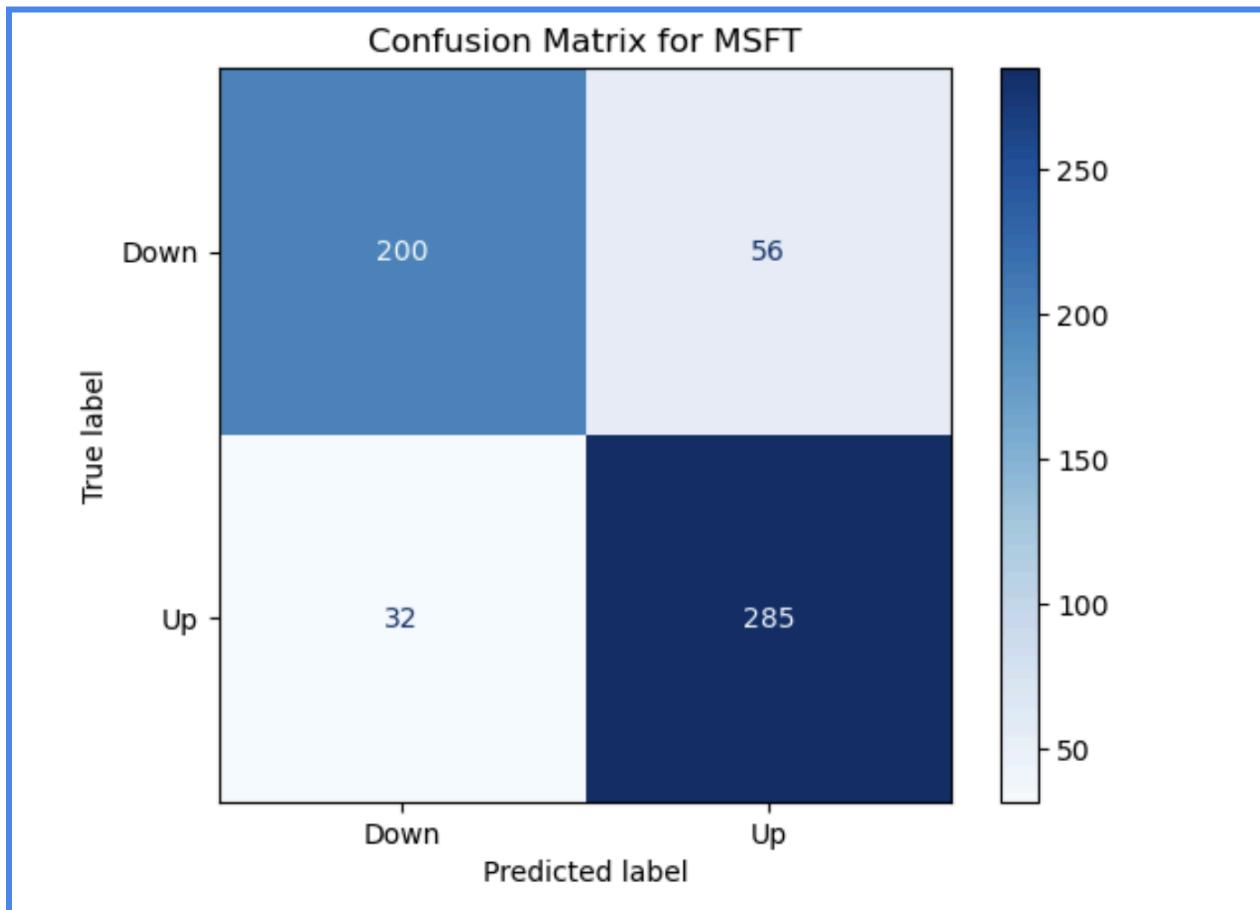


Backtesting Metrics: {'accuracy': 0.8333333333333334, 'precision': 0.8529411764705882, 'recall': 0.8169014084507042, 'f1\_score': 0.8345323741007195, 'roc\_auc': 0.9228038212692405, 'total\_pnl': 77.14997482299805, 'avg\_daily\_pnl': 0.18635259619081654, 'max\_drawdown': -0.5}

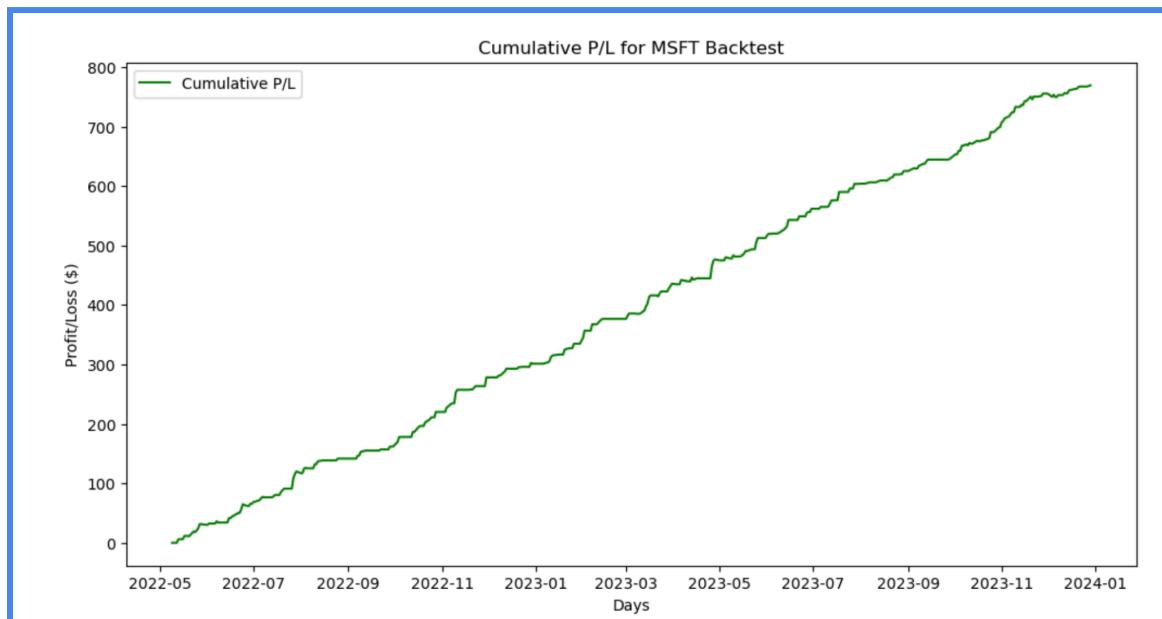
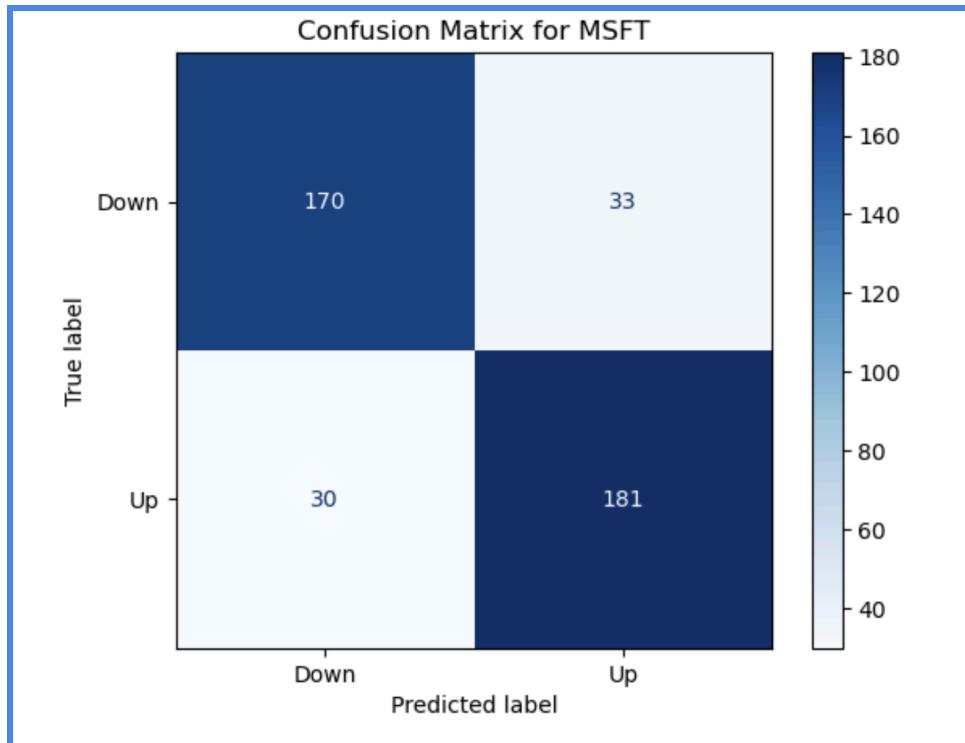
## Microsoft (MSFT)

- Threshold: Mean Squared Loss < 18%

### TRAINING RESULTS:



### BACKTESTING RESULTS:

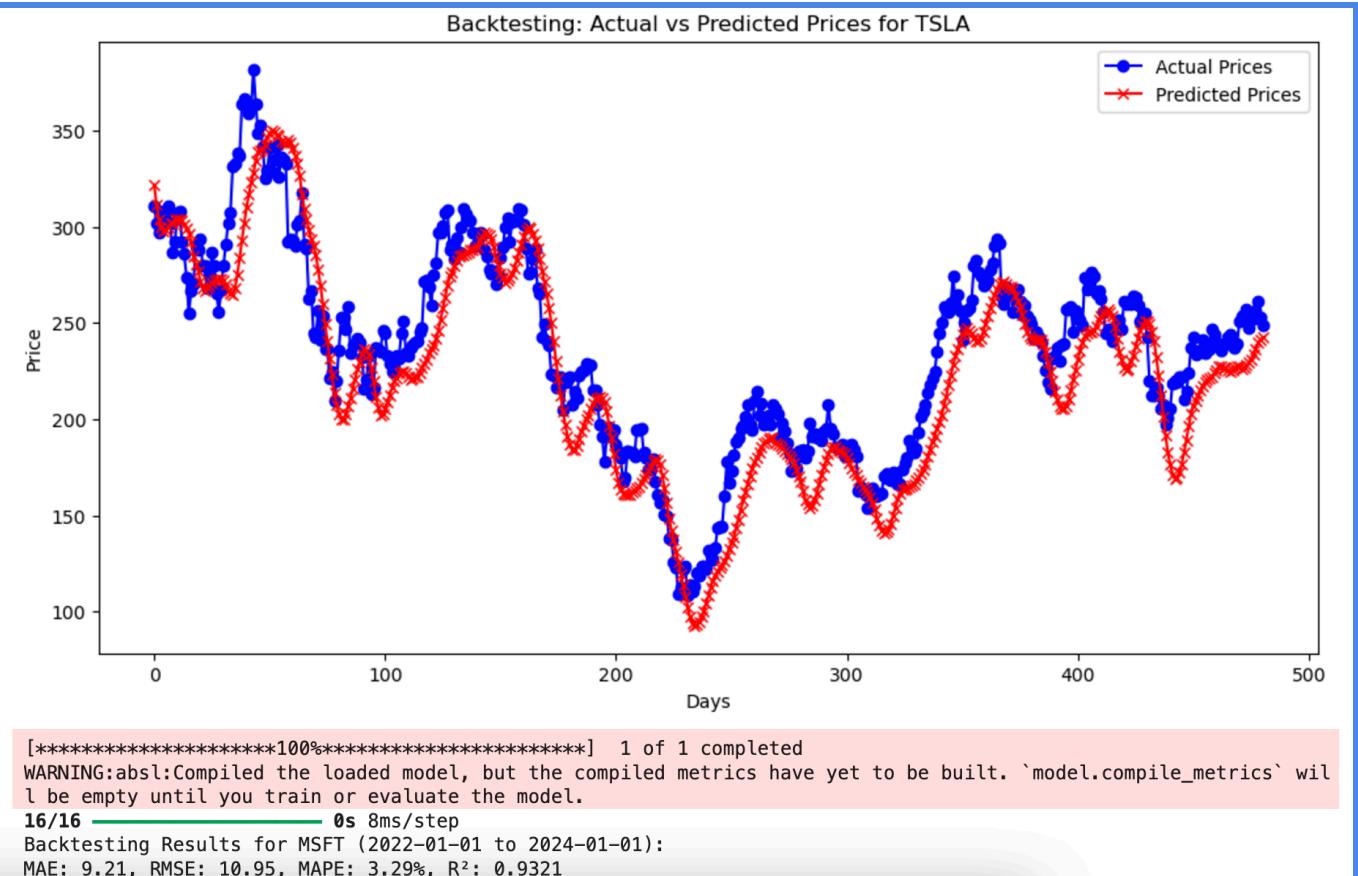


```
Backtesting Metrics: {'accuracy': 0.8478260869565217, 'precision': 0.8457943925233645, 'recall': 0.8578199052132701, 'f1_score': 0.851764705882353, 'roc_auc': 0.9151238531039152, 'total_pnl': 769.3001251220703, 'avg_daily_pnl': 1.8582128626136964, 'max_drawdown': -5.71002197265625}
```

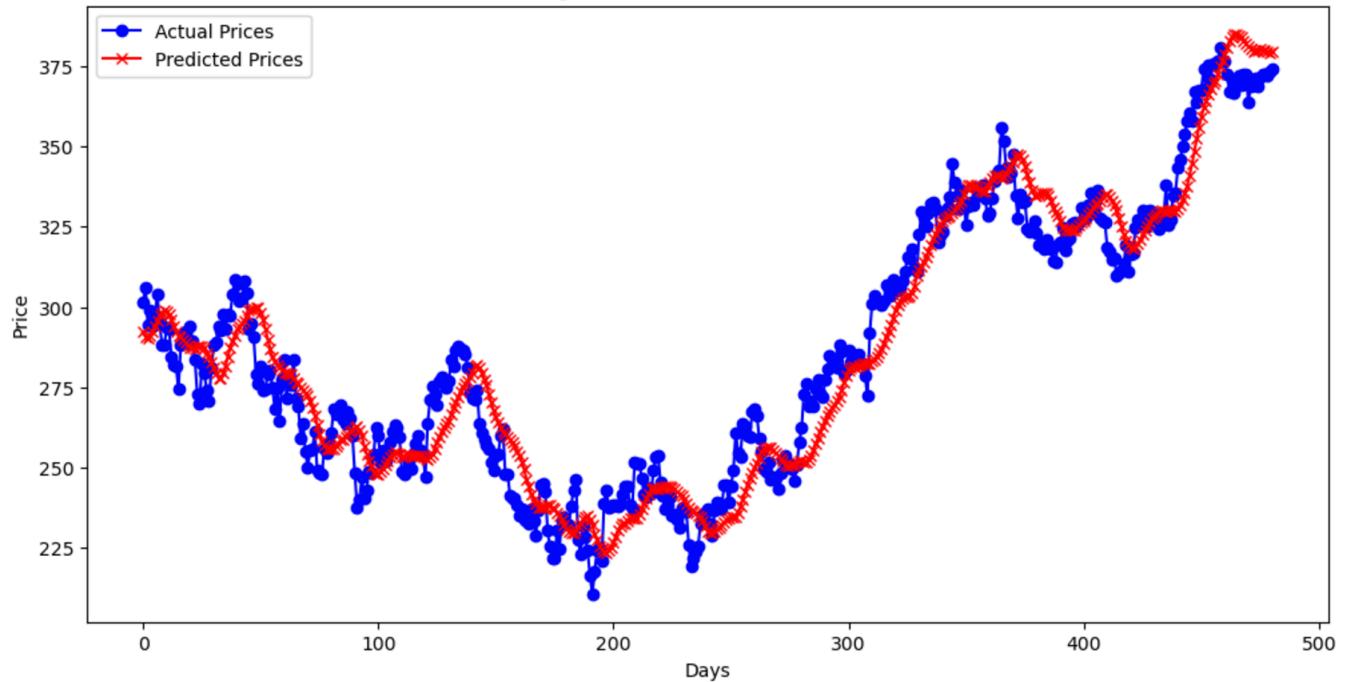
## 5. Discussion and Analysis

### Comparison of Models

- **LSTM:** The model demonstrates strong predictive capabilities, achieving relatively low error metrics, including MAE, RMSE, and MAPE, alongside an acceptable R<sup>2</sup> score, indicating a good fit between predicted and actual prices. When integrated into a simple trading strategy, the model effectively generates profit/loss (P/L) insights, showcasing its ability to identify price trends. The P/L analysis shows distinct outcomes: TSLA delivers higher differential P/L due to volatility but faces larger drawdowns with a negative ending balance. MSFT offers steady growth with moderate P/L and better risk-adjusted returns with a respectable ending balance. KO generates the most stable P/L with minimal drawdowns but lower overall gains. The model performs well but varies based on stock volatility. However, the Sharpe Ratio and Max Drawdown highlight the need for improved risk-adjusted performance and drawdown control, suggesting the trading strategy could benefit from optimization or additional features like stop-loss mechanisms. Overall, the model performs well in forecasting and demonstrates potential for real-world application with further refinement.
- **XGBoost:** The P/L analysis for the XGBoost model highlights notable differences across the three stocks due to variations in price behavior and volatility. TSLA demonstrates the highest total P/L, with profits increasing steadily over time. This reflects the model's ability to capitalize on TSLA's dynamic price trends while avoiding major drawdowns, showcasing a stable growth trajectory despite the stock's inherent volatility. MSFT offers a balanced outcome, generating steady P/L with moderate gains and lower drawdowns. The model performs effectively in capturing trends for this relatively stable, growth-oriented stock. MSFT's lower volatility allows for more accurate predictions and consistent returns, resulting in superior risk-adjusted performance compared to TSLA. KO exhibits the most stable performance, with minimal drawdowns and consistent, though lower, total P/L. As a defensive stock with limited price swings, KO aligns well with the model's predictions, making it an ideal choice for conservative strategies. However, the trade-off between stability and profitability is evident, as its returns are less pronounced. Overall, the XGBoost model demonstrates strong adaptability to each stock's characteristics. It performs effectively across TSLA, MSFT, and KO, capturing stable growth in TSLA, consistent gains in MSFT, and reliable performance in KO. This flexibility underscores the model's potential across a range of market conditions, provided appropriate risk management strategies are in place.



Backtesting: Actual vs Predicted Prices for MSFT



```
[*****100%*****] 1 of 1 completed
WARNING:absl:Compiled the loaded model, but the compiled metrics have yet to be built. `model.compile_metrics` will be empty until you train or evaluate the model.
16/16 ━━━━━━ 0s 7ms/step
Backtesting Results for K0 (2022-01-01 to 2024-01-01):
MAE: 0.91, RMSE: 1.15, MAPE: 1.59%, R2: 0.7418
```