DS301 Project

Stock Prediction with Multi-Scale Transformer

Yukang Fan, Rongjian Wang, Jialuo Huang

Executive Summary

- **Problem Statement:** Predict stock prices effectively using high-frequency data.
- **Goal:** Capture short-term and long-term patterns for accurate predictions.
- **Technical Challenges:** Handling multiple time scales. Modeling volatility and long-term dependencies.
- **Solution Approach:** Multi-scale Transformer model.
- **Benefit of Transformer:** Improved accuracy and stability over traditional models.

Multi-Scale Transformer

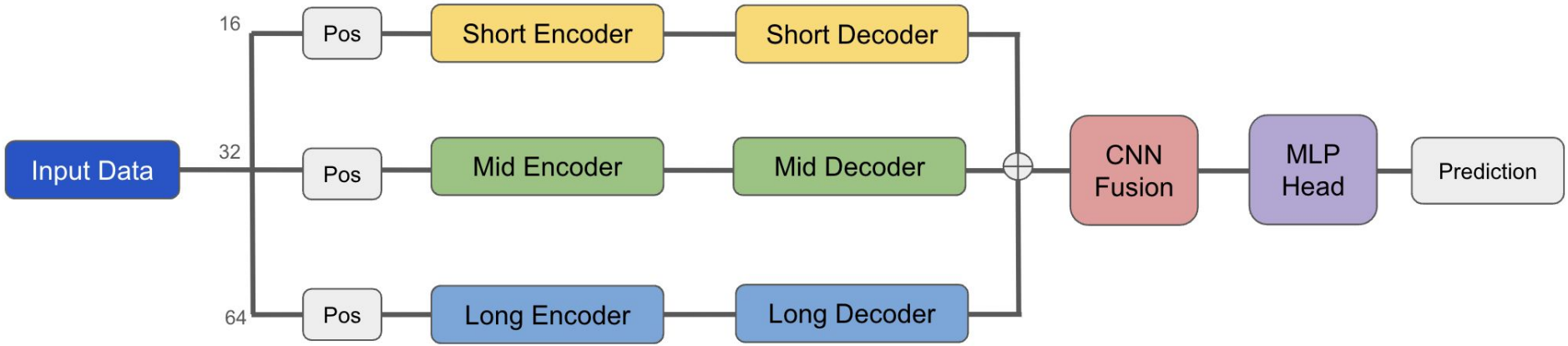
•**Solution Architecture:**

-Multi-Scale Transformer with sequence length 16, 32, 64.

-Separate Encoder-Decoder pairs for each time scale.

-Feature fusion using Conv1D and MLP.

-Captures short-term, intermediate, and long-term dependencies.



Data Overview

- **Data:** NASDAQ index, Historical Data of Individual Stocks.

  (2000/01/01~2010/01/01)

- **Features**: Open, High, Low, Close, and Volume (for individual stocks) (minute by minute).

- **Preprocessing:**

- Resampling into 15-min, 30-min, 60-min intervals.

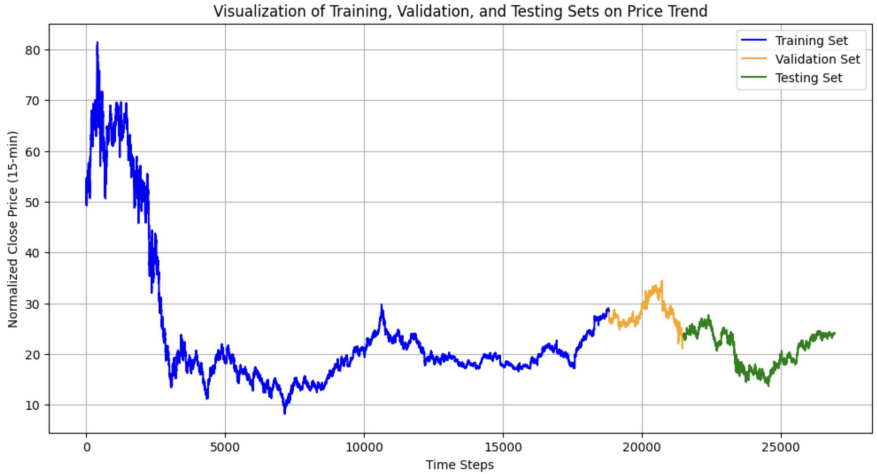
- Handling missing values with forward filling.

- Normalization (MinMax)

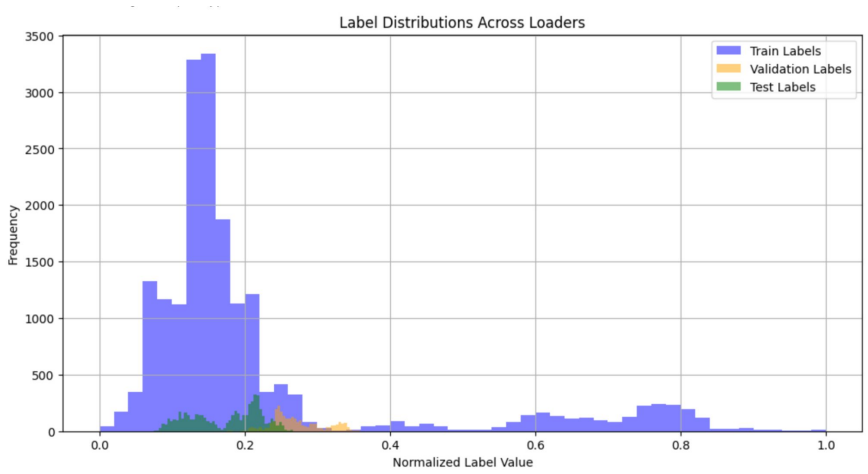
- **Sample:** [short_seq(16), mid_seq(32), long_seq(64), target_seq, label]

Experiment

- Train: 0.7, Validation: 0.1, Test: 0.2
- Hyperparameters:
- LR: 0.0001, Batch Size: 16, Optimizer: Adam, Loss: MSE, Epochs: 10



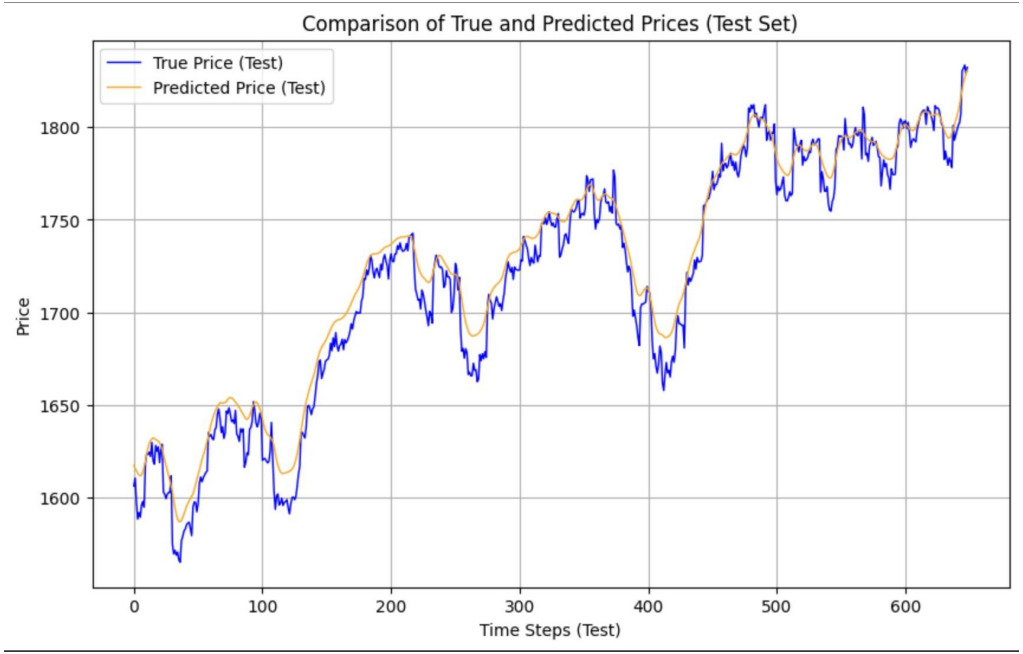
Data Split Example(CSCO)



Data Distribution Example(CSCO)

Result

- **Mean Squared Error (MSE)**: 0.3581
- **Root Mean Squared Error (RMSE)**: 0.5951
- **Mean Absolute Error (MAE)**: 0.6720
- **R^2 Score**: 0.9697

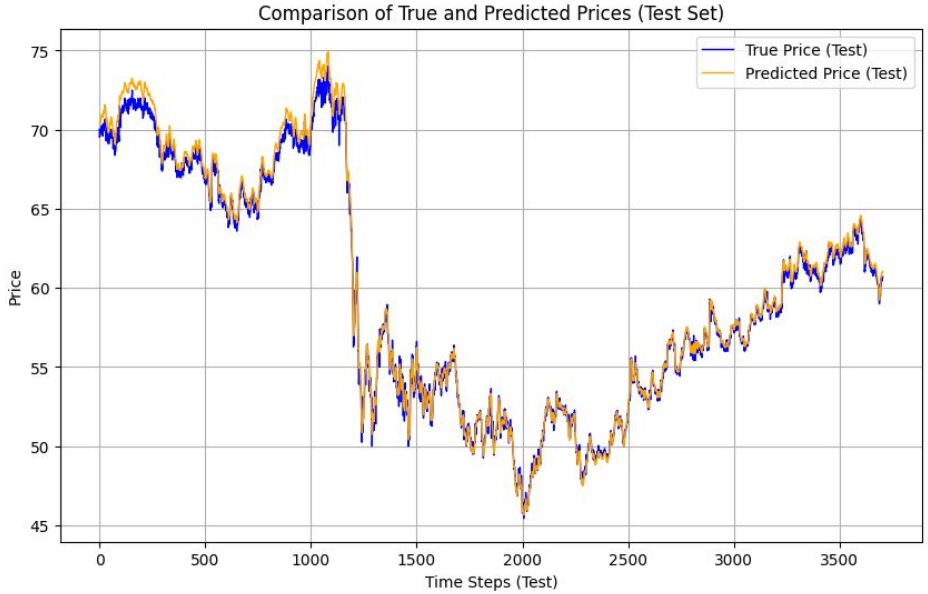


Inverse-Transformed Prediction Result Visualization (NASDAQ)

Testing Results

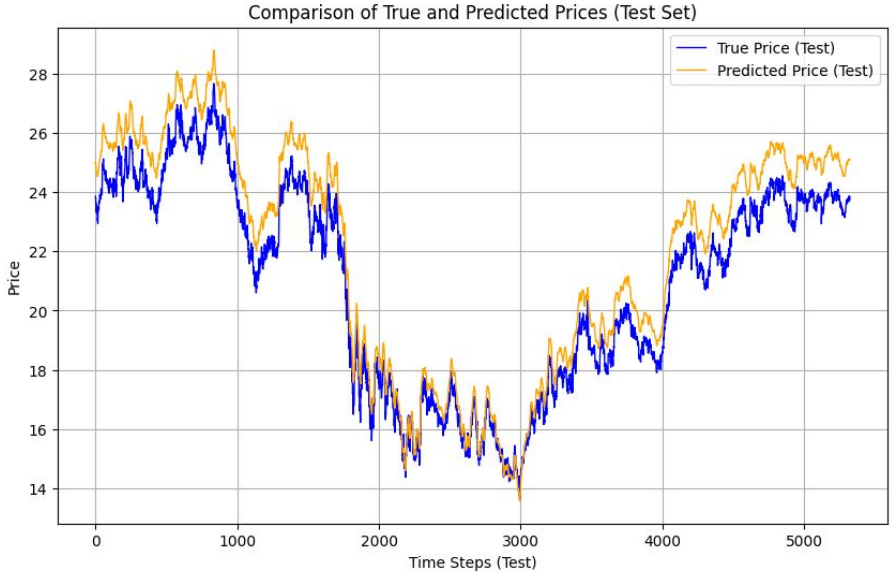
- PEP (Pepsi)

```
Evaluation Results on Denormalized Scale (Test Set):
MSE: 0.2778
RMSE: 0.5271
MAE: 0.3920
R²: 0.9949
```



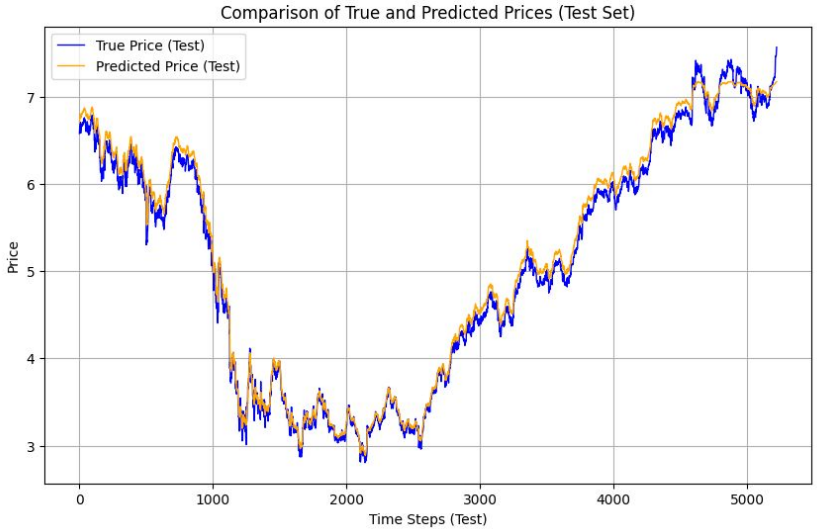
- CSCO (Cisco)

```
Evaluation Results on Denormalized Scale (Test Set):
MSE: 1.1526
RMSE: 1.0736
MAE: 0.9881
R²: 0.9054
```



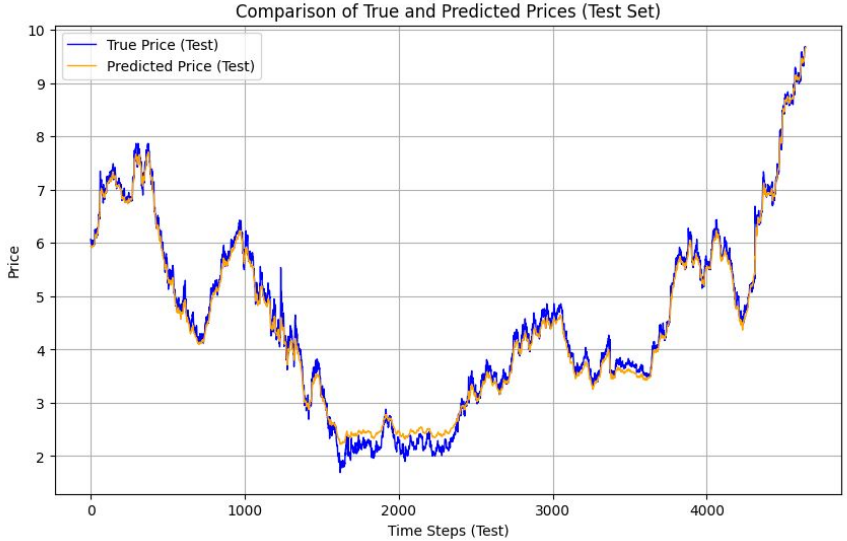
- APPL (Apple)

```
Evaluation Results on Denormalized Scale (Test Set):
MSE: 0.0123
RMSE: 0.1109
MAE: 0.0968
R²: 0.9936
```



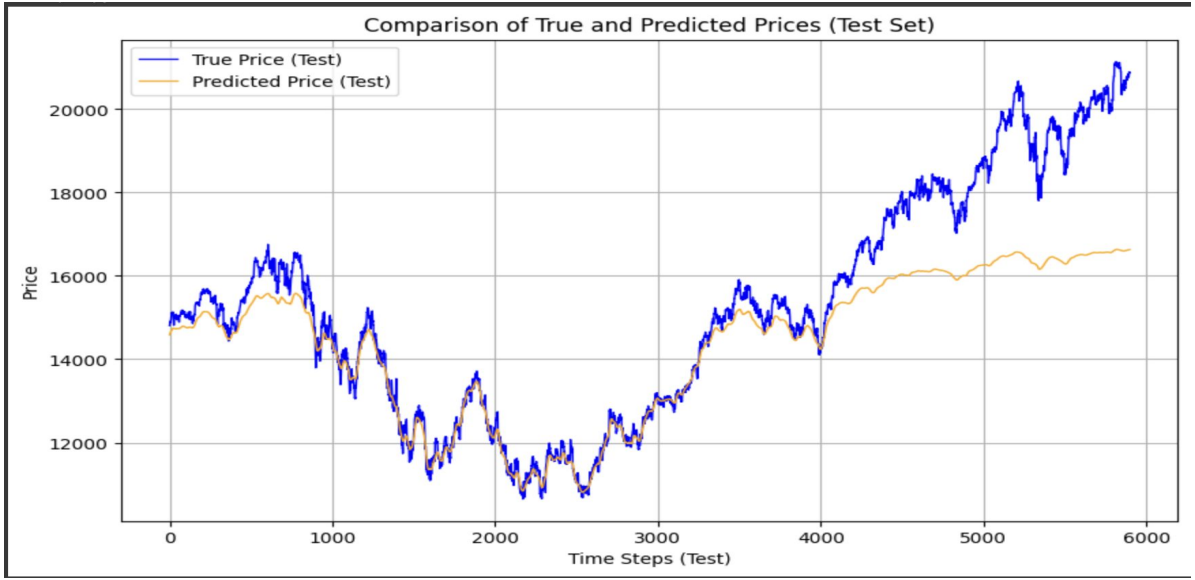
- AMD

```
Evaluation Results on Denormalized Scale (Test Set):
MSE: 0.0252
RMSE: 0.1587
MAE: 0.1280
R²: 0.9916
```



Combining with traditional Machine Learning Techniques

- **High Volatility Nature of Equity Market**
- **The model may perform bad when the market hikes very high**



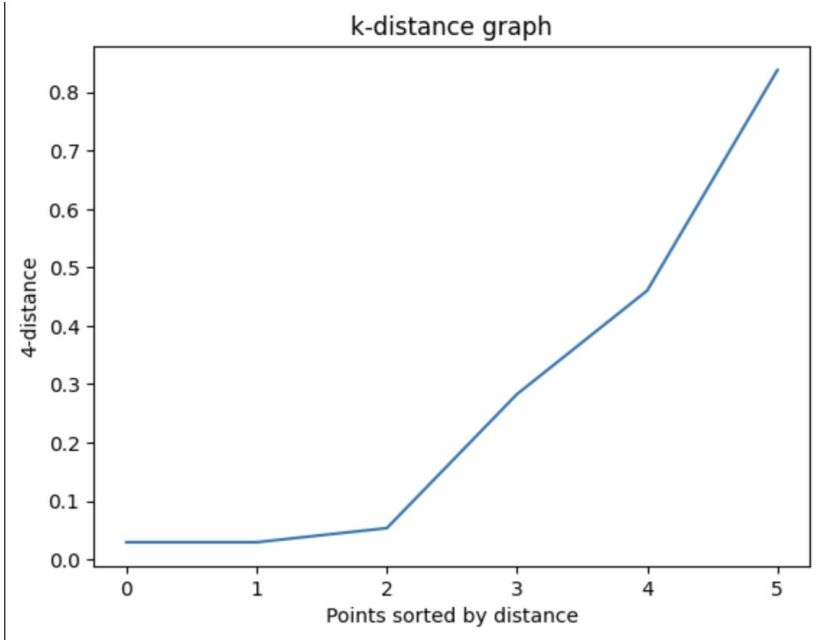
Nasdaq 100 index（2015-2024）

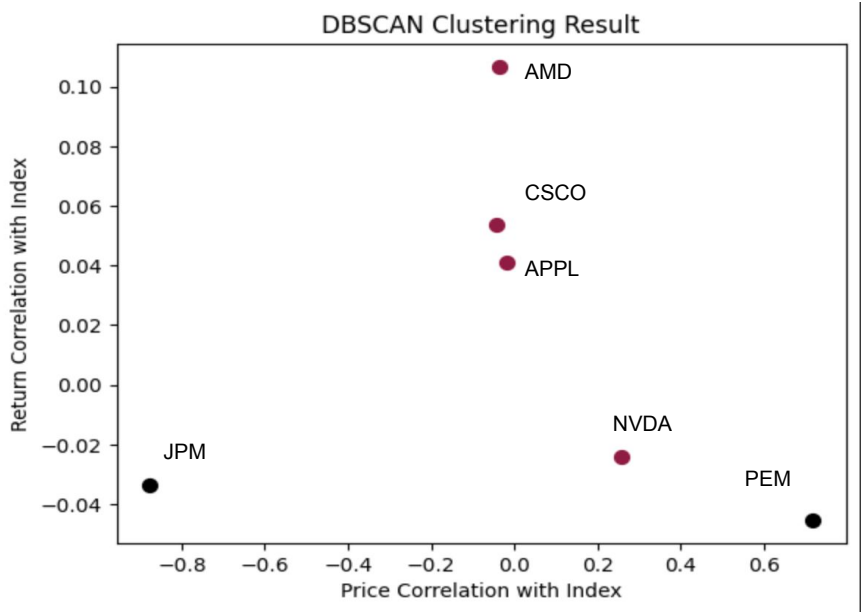
Evaluation Results on Denormalized Scale (Test Set):

MSE: 2146876.1757
RMSE: 1465.2222
MAE: 900.5547
R^2: 0.7094

Combining with traditional Machine Learning Techniques

- **Unsupervised Clustering Analysis:** Identifying deviations in individual stock and index movements





Conclusion and Future Work

- In the data time scale, the multi-scale transformer model effectively captures historical pattern and make the prediction.

- we could further collect more individual stock data, potentially the clustering analysis can become more robust.

- In addition to the deep learning evaluation system, this algorithm can be implemented in the market: predicting next-day rises to buy the day before and declines to sell the day before, enabling PNL calculation