Transformer-Based Multivariate Stock Forecasting
with Random Forest and AdaBoost Comparison

1. Introduction

Forecasting stock market movements is a notoriously difficult challenge due to volatility, noise, and nonlinear interactions among financial instruments. Traditional models often fail to capture long-term dependencies and cross-asset relationships. To address these challenges, this project implements a simplified Transformer architecture along with two classical ensemble models—Random Forest and AdaBoost—for performance benchmarking.

The aim is to predict one-step-ahead log percent changes for ExxonMobil (XOM) using six multivariate oil-industry time-series inputs from Yahoo Finance. A simple cumulative prediction backtest is also included.

2. Methodology

2.1 Implemented Methods

2.1.1 Data Pipeline

- Data Source: Yahoo Finance API
- Tickers: XOM, CVX, COP, BP, PBR, EOG
- Target Variable: XOM
- Metric: LogPercentChange = ln(Close / Open + 1)
- Lookback Window: 20 days
- Task: One-step-ahead regression forecasting

2.1.2 Transformer Regressor

A PyTorch-based Transformer architecture was implemented with the following components:
- Linear input projection to d_model
- Two-layer Transformer encoder with multi-head attention

- Adaptive pooling layer
- Fully connected regression output head

2.1.3 Random Forest Regressor

A 200-estimator Random Forest was trained on flattened inputs ($20 \times 6 = 120$ features).

2.1.4 AdaBoost Regressor

A 200-estimator AdaBoost model was used. AdaBoost reduces bias through iterative boosting of weak learners.

2.2 Evaluation Metrics and Visualizations

The following evaluation metrics were used:
- Mean Squared Error (MSE)
- Mean Absolute Error (MAE)
- Cumulative predicted returns (proxy for signal quality)

3. Results

3.1 Quantitative Results

The following table summarizes the model performance on the test set:

Model	Test MSE	Test MAE
Transformer	4.95e-05	0.00515
Random Forest	4.41e-05	0.00477
AdaBoost	4.16e-05	0.00475

3.2 Observations

Random Forest and AdaBoost significantly outperform the Transformer model on both MSE and MAE. Despite its sophistication, the Transformer requires more hyperparameter tuning, larger datasets, and optimized training procedures to perform competitively.

3.3 Visual Analysis

Cumulative predicted return curves show that ensemble models produce smoother and more stable signals. Meanwhile, the Transformer predictions exhibit higher variance, suggesting potential overfitting.

4. Discussion

The simplified Transformer architecture captures sequential dependencies but struggles with: limited data volume, training instability, and lack of advanced positional or temporal encoding. Tree-based ensemble models perform strongly due to their robustness and suitability for small datasets.

Random Forest and AdaBoost excel in scenarios where temporal relationships are less critical, and strong non-linear tabular patterns exist.

5. Conclusion

This project demonstrates a full multivariate time-series forecasting pipeline and compares Transformer, Random Forest, and AdaBoost performance. Both classical ensemble models outperform the Transformer on this dataset. However, the Transformer shows potential for improvement with additional optimization.

Future Work

- Integrate additional features: technical indicators, macroeconomic data, sector indices
- Use higher-frequency data (hourly)
- Implement Informer-style sparse attention
- Hyperparameter tuning using Optuna/Ray Tune
- Develop full trading backtests including transaction costs