Evaluating the Predictive Limits of Machine Learning in Daily Stock Return Forecasting: Evidence from Samsung Electronics

1.Introduction

Financial time-series forecasting remains one of the most challenging applications of machine learning. While stock prices exhibit long-term trends, short-term returns are widely believed to behave close to a random walk under the weak-form Efficient Market Hypothesis (EMH). This project evaluates whether modern machine learning techniques—specifically Random Forest and XGBoost—can extract meaningful predictive signals from historical price and volume data of Samsung Electronics. Rather than focusing solely on predictive performance, this study aims to critically assess the limits of daily stock return predictability and compare model performance against a naive persistence benchmark.

2.Data Description and Target Construction

The dataset consists of historical daily stock data for Samsung Electronics, including open, high, low, close, adjusted close prices, and trading volume. Adjusted close prices were used instead of raw closing prices because they account for stock splits, dividends, and corporate actions, ensuring economically consistent price representation over time.

To transform the non-stationary price series into a more suitable modeling target, daily log returns were computed:

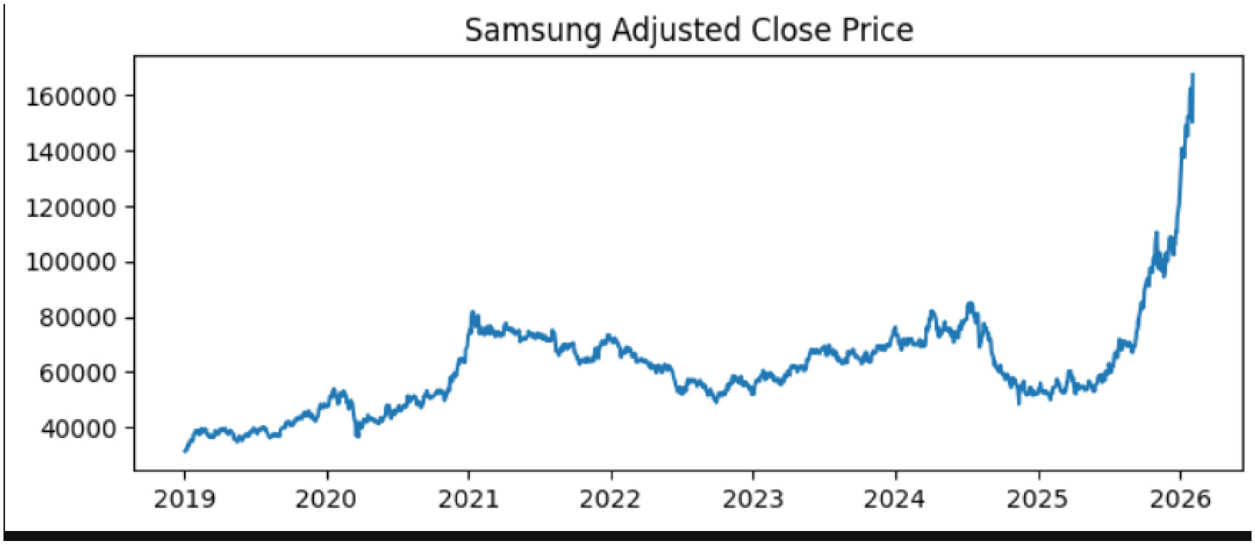
$$\text{Log\_Return}(t)=\log [P(t)/ P(t-1)]$$

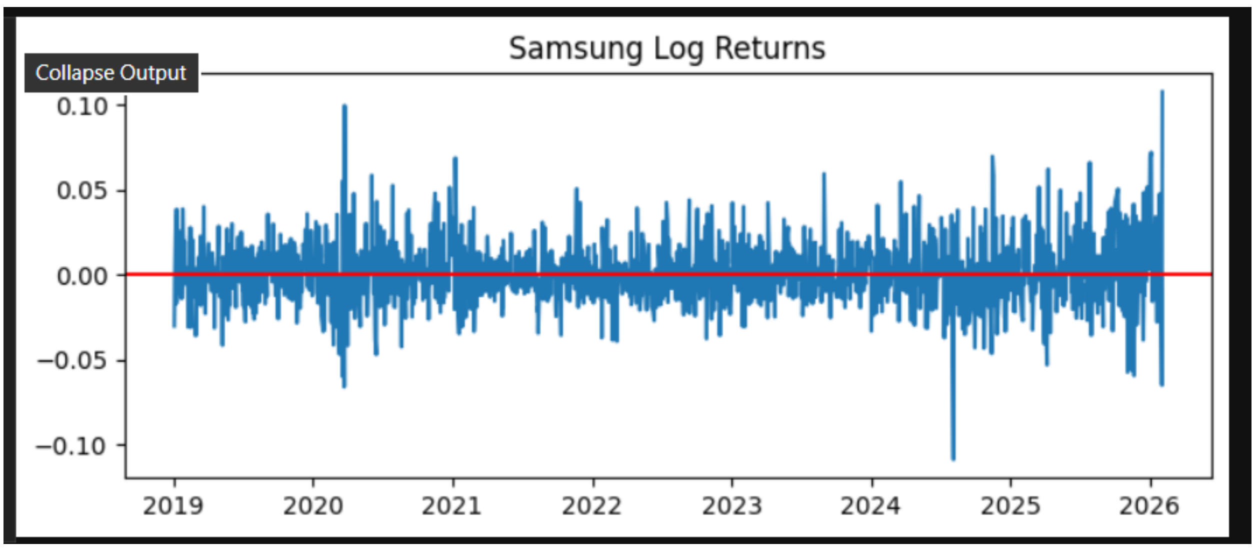
Log returns were selected because they remove deterministic trend components and are standard in financial econometrics. This transformation makes the series approximately stationary and suitable for predictive modeling.

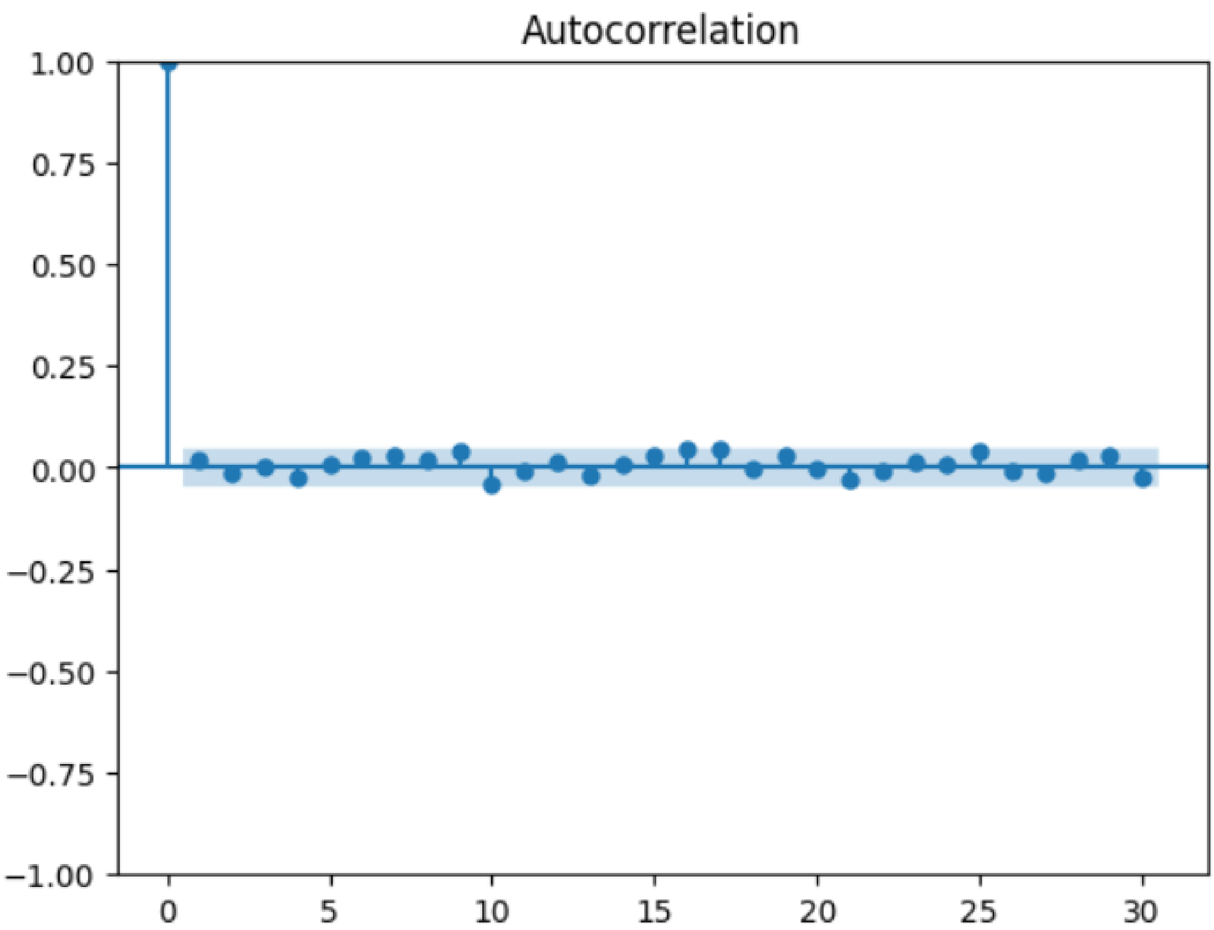
3.Exploratory Data Analysis (EDA)

Initial exploration revealed that adjusted closing prices follow a strong upward trend over time, confirming their non-stationary nature. In contrast, the log return series fluctuates around zero and exhibits volatility clustering, a well-documented characteristic of financial markets.

Autocorrelation analysis demonstrated very weak serial dependence in daily returns. The autocorrelation function (ACF) showed minimal statistically significant lags, suggesting that simple momentum-based strategies may not be effective at the daily frequency. These findings align with financial theory, which suggests that short-term returns contain limited predictable structure.







4.Feature Engineering

To investigate whether non-linear relationships or structural patterns exist beyond simple autocorrelation, a comprehensive set of time-series features was constructed. These included lagged returns (lag_1, lag_2, lag_5, etc.), rolling means, rolling standard deviations to capture volatility clustering, momentum indicators measuring multi-day price changes, moving average ratios to capture overbought or oversold conditions, and volume-based features reflecting trading intensity.

All features were appropriately shifted to ensure that only past information was used when predicting future returns, thereby preventing data leakage. The goal of this feature engineering process was to provide machine learning models with structured representations of trend, volatility, and momentum patterns that may not be captured through simple linear methods.

5.Modeling Strategy and Validation

A time-series-aware validation strategy was adopted using chronological train-test splitting and TimeSeriesSplit cross-validation. This approach preserves temporal ordering and prevents future information from leaking into the training process.

To contextualize model performance, a naive persistence model was implemented as a baseline:

$$\hat{R}_t = R_{t-1}$$

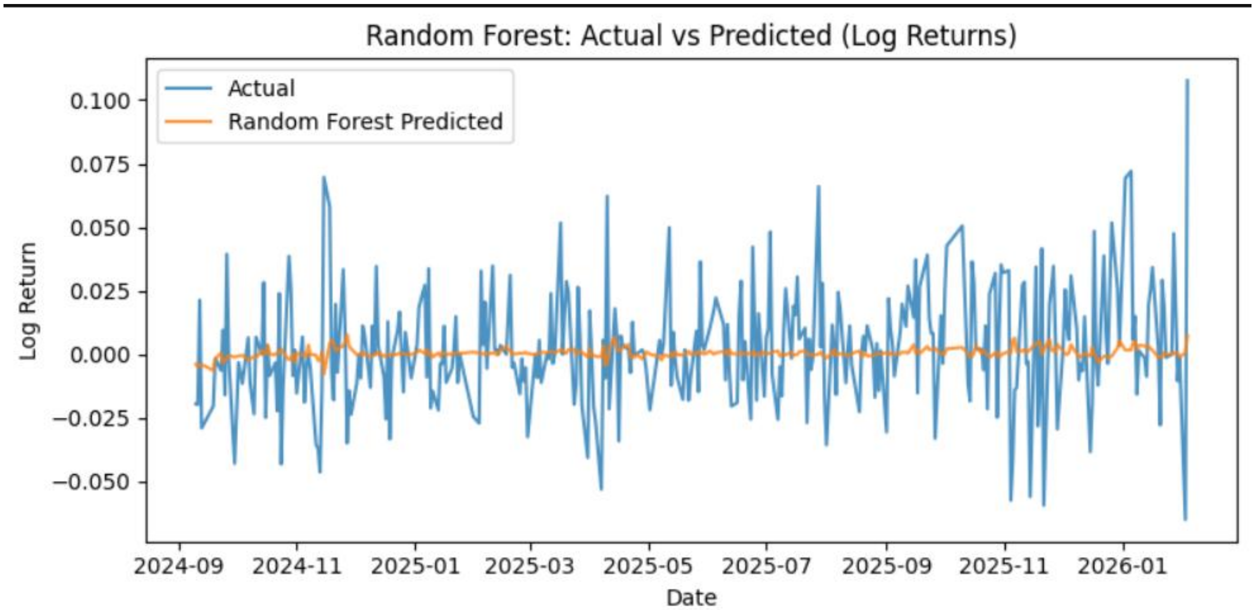
This model assumes that tomorrow's return equals today's return. The naive model performed poorly, with a directional accuracy of only 45.4% and substantially higher error metrics. This result indicates that daily returns do not exhibit strong first-order autocorrelation and that simple momentum assumptions are ineffective.

6.Random Forest Result

The Random Forest model was extensively tuned using randomized hyperparameter search across 600 candidate configurations. The final tuned model achieved a Mean Absolute Error (MAE) of 0.01724, a Root Mean Squared Error (RMSE) of 0.02317, and a directional accuracy of 52.52%. While the R^2 value remained close to zero (−0.0007), this outcome is expected in daily return forecasting, where variance explained is typically very small.

Compared to the naive baseline, Random Forest significantly reduced prediction error and improved directional accuracy, indicating that the engineered features contain limited but measurable predictive information.

```
Tuned RF Results:
MAE: 0.017244584469456137
RMSE: 0.023172425542176436
R²: -0.0007438012920375492
Directional Accuracy: 0.5252225519287834
```

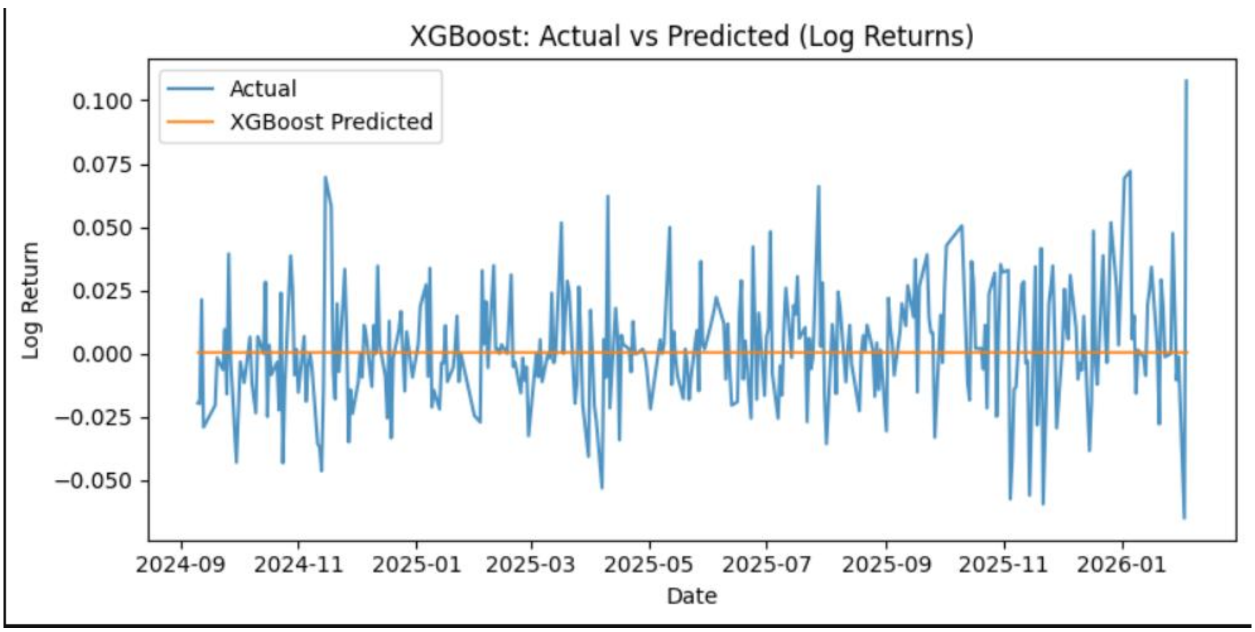


7.XGBoost Result

XGBoost was also tuned extensively using randomized hyperparameter optimization. The tuned model achieved performance metrics similar to Random Forest, with MAE of 0.01735, RMSE of 0.02328, and directional accuracy of 50.7%. Although XGBoost slightly improved upon the naive baseline, it did not outperform Random Forest in this case.

The similarity in performance across models suggests that the predictive ceiling of this dataset may have been reached, and additional complexity does not necessarily translate into substantial performance gains.

```
Tuned XGBoost Results:
MAE: 0.017353987579446315
RMSE: 0.023282420456883363
R²: -0.010267015431339788
Directional Accuracy: 0.5074183976261127
```



8.Interpretation and Discussion

The key insight from this study is that while machine learning models outperform naive persistence strategies, their predictive power remains modest. The near-zero R^2 values indicate that daily return variance is largely unexplained by historical price and volume information alone. However, the consistent improvement in directional accuracy above random guessing suggests that small, non-linear predictive structures may exist.

These findings are consistent with the weak-form Efficient Market Hypothesis, which states that past price information is largely reflected in current prices. The results demonstrate the practical difficulty of extracting economically meaningful predictive signals from high-frequency financial data.

9.Conclusion

This study evaluated the predictive limits of machine learning models in daily stock return forecasting using Samsung Electronics data. Through rigorous feature engineering, proper time-series validation, and extensive hyperparameter tuning, Random Forest and XGBoost models were assessed against a naive benchmark.

While the models achieved measurable improvements over the naive baseline, overall predictive power remained limited. These results highlight both the strengths and constraints of machine learning in financial markets and reinforce the importance of realistic evaluation frameworks.

Rather than presenting exaggerated predictive performance, this project emphasizes methodological rigor, critical interpretation, and alignment with established financial theory.

10.Future Research Directions

Future extensions of this work could include predicting weekly returns, incorporating macroeconomic indicators, integrating sentiment analysis features, converting the task to directional classification, or implementing full trading strategy backtesting to evaluate economic profitability rather than purely statistical accuracy.