

Assignment 8: Fitting a Model to Stock Data

Introduction

This report presents an analysis of historical stock price data for Alphabet Inc. (GOOGL). The objective is to predict future stock returns using a linear regression model and compute key financial metrics, including Annual Return, Sharpe Ratio, and Sortino Ratio. The dataset spans the past 2 years, with daily adjusted closing prices obtained via Yahoo Finance (yfinance).

Data Collection and Preprocessing

Data Collection

The dataset was collected using the yfinance package for the ticker symbol GOOGL, covering a 2-year period with a daily frequency.

Data Preprocessing

- **Daily Returns Computation:** Returns were calculated as the percentage change in the adjusted closing price.
- **Handling Missing Values:** There were no missing values in the adjusted closing price data.
- **Rolling Window Feature Engineering:** A **10-day rolling window** of past returns was used as features to predict the next day's return.
- **Train-Test Split:** The dataset was split into **80% training data** and **20% testing data**, ensuring time order was preserved (i.e., no shuffling).

Model Description

A Linear Regression Model was implemented to predict the next day's return based on the past 10 days of returns. The model was trained using the **scikit-learn LinearRegression** function, minimizing error between predicted and actual returns. The performance of the model was evaluated using:

- **Mean Absolute Error (MAE):** 0.015606
- **Root Mean Squared Error (RMSE):** 0.020311

These values suggest that the model's predictions have a relatively small error, making it useful for return forecasting.

Financial Metric Computation

Annual Return

The Annual Return was computed using the formula: $(1 + \text{cumulative_return})^{**} (252 / \text{num_days}) - 1$. Where cumulative_return is the total return, 252 is the approximate number of trading days in a year and num_days is amount of days. The computed **Annual Return is 31.75%**, indicating strong long-term growth.

Sharpe Ratio

The Sharpe Ratio is a risk-adjusted return measure: $((\text{daily_mean} - \text{risk_free_rate}) / \text{daily_std}) * \text{np.sqrt}(252)$

Where:

- daily_mean = Mean daily return
- risk_free_rate = Risk-free rate (4% annually, adjusted for daily returns)
- daily_std = Standard deviation of daily returns

The **Sharpe Ratio is 0.974**, indicating moderate risk-adjusted performance.

Sortino Ratio

The Sortino Ratio adjusts for downside risk only: $((\text{daily_mean} - \text{risk_free_rate}) / \text{downside_std}) * \text{np.sqrt}(252)$. Where downside_std is the standard deviation of negative returns. The **Sortino Ratio is 1.338**, showing that returns are relatively stable with limited downside volatility.

Insights and Conclusion

1. **Return Predictability:** The linear regression model demonstrated reasonable predictive accuracy with low error values (MAE: 0.0156, RMSE: 0.0203).
2. **Annual Growth:** The computed annual return of 31.75% suggests strong overall performance.
3. **Risk-Adjusted Performance:** A Sharpe Ratio of 0.974 indicates that the stock provides returns moderately higher than the risk-free rate, while the Sortino Ratio of 1.338 suggests a favorable risk-adjusted return, particularly in avoiding downside losses.

Recommendations

- Further improving predictions could involve testing more advanced time-series models like ARIMA or LSTMs.
- Portfolio optimization could benefit from analyzing the Sortino Ratio more closely to manage downside risk.
- Additional stocks could be analyzed for comparison to understand relative risk-adjusted performance.

Overall, this analysis provides insights into GOOGL's return predictability and risk-adjusted returns, useful for financial decision-making.