Stock Market Analysis and Prediction using Machine Learning

NAME – Naman Khard

College – Sr Group of Institution

COURSE – B-Tech (Data Science)

Date – 25/08/25

INTRODUCTION

This project aims to analyze stock market data and build machine learning models to predict future stock prices. The focus is on Apple (AAPL), using historical data and technical indicators. The project demonstrates skills in data analysis, feature engineering, and regression modeling. To analyze market trends and predict future market behavior using machine learning techniques.

DATASET DETAILS

- The Dataset was given by Unifed Mentor, it contain the following column:-
- List of column:- Date, Ticker, Open, High, Low, Close, Adj Close, Volume.
- Created column:- Daily Return (%), 7-Day Moving Average, 30-Day Moving Average, Volatility (30-day std), Lag_1, Lag_2, Lag_3.

Methodology

- Data Cleaning:- removed missing values due to lag features, converted Date to datetime.

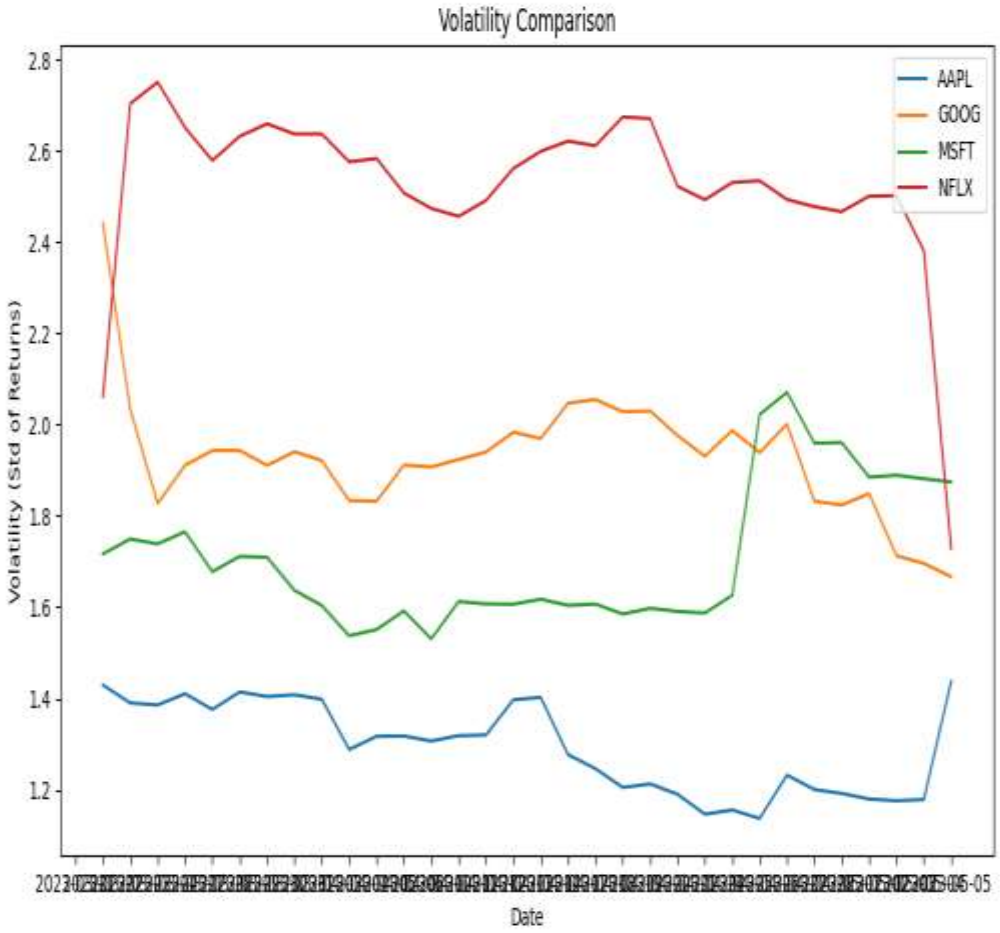
- Feature Engineering:- created lag features, moving averages, volatility.

- Modeling:- used Linear Regression & Random Forest.

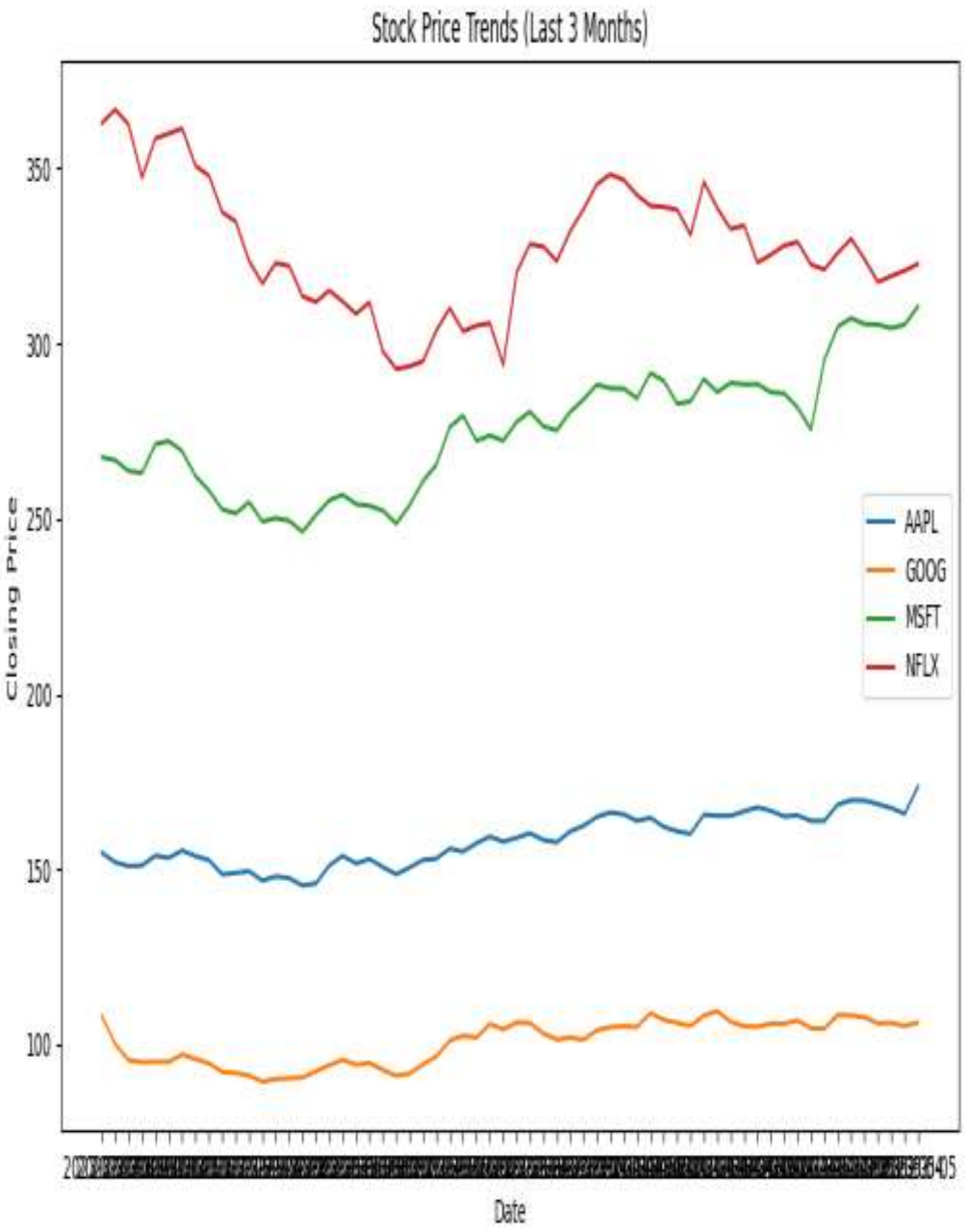
- Evaluation:- RMSE, MAE, R^2 used to compare models.

Exploratory Data Analysis (EDA)

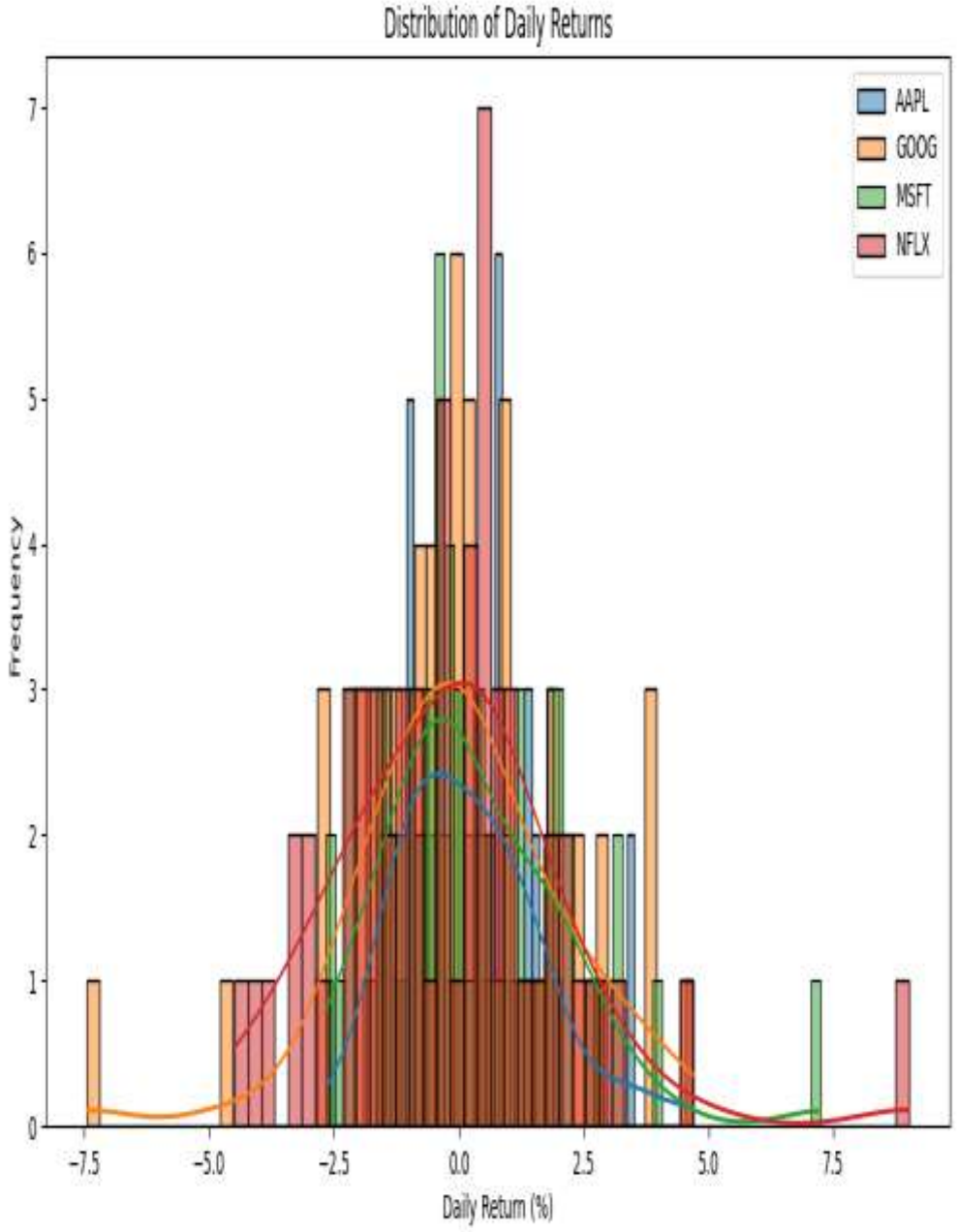
- Volatility Comparison:- Volatility comparison shows how different stocks fluctuate over time.



- Stock Price Trend (Last 3 Months):- Recent stock price trends provide insights into short-term movements.

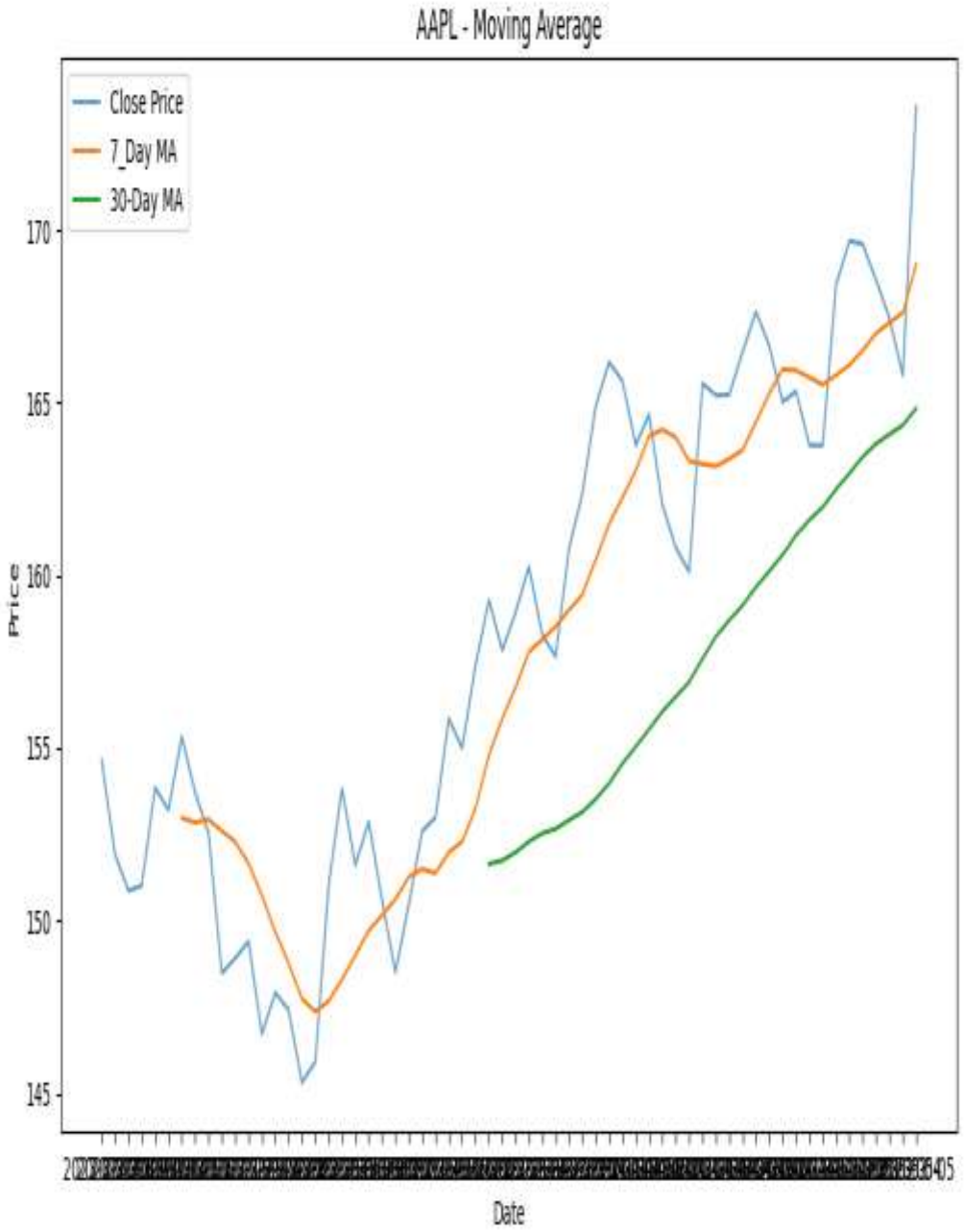


- Distribution of Daily Returns:- Distribution of daily returns highlights overall risk-return patterns.

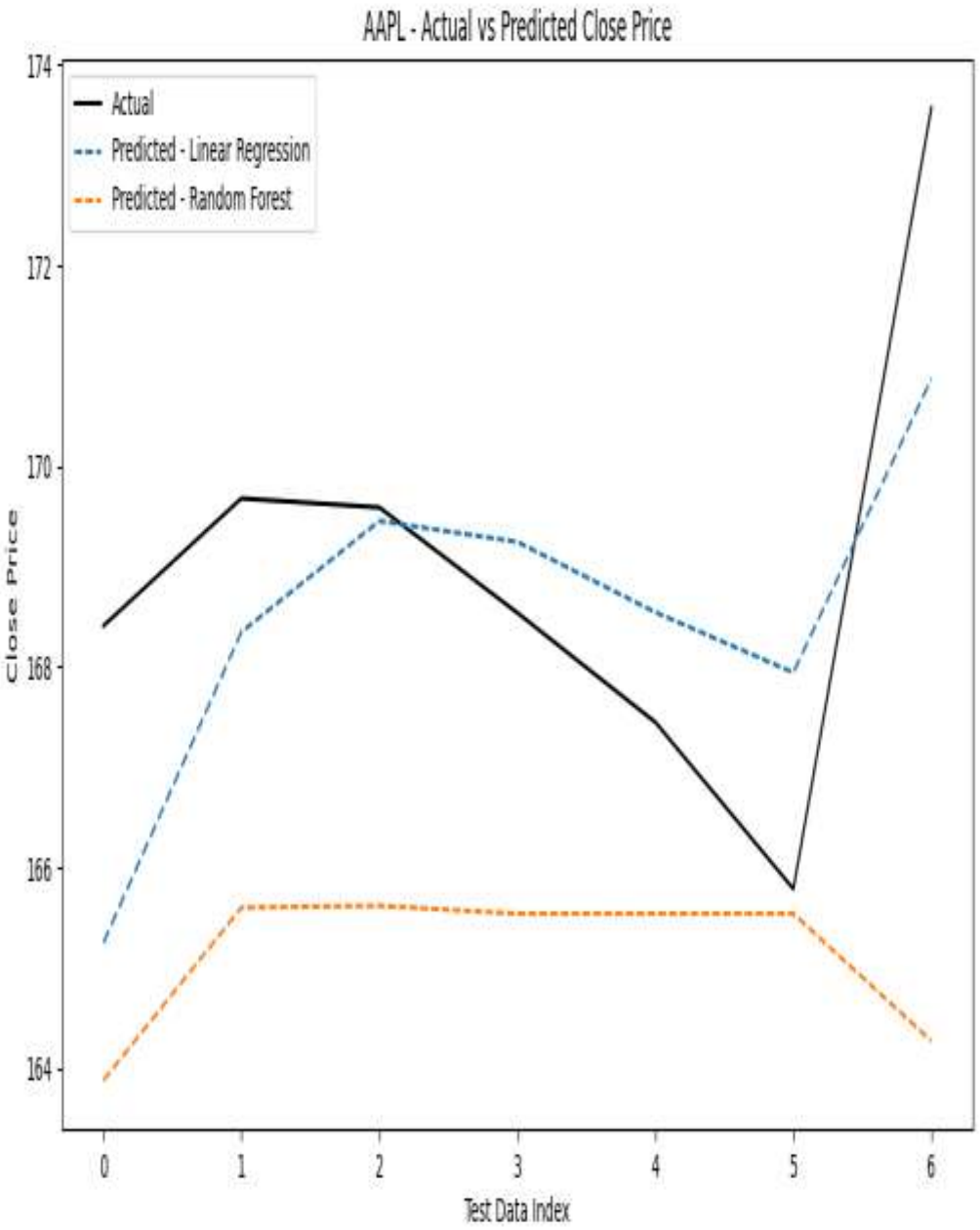


- Correlation between Stock Price:- Correlation heatmap shows how closely stock prices move together.

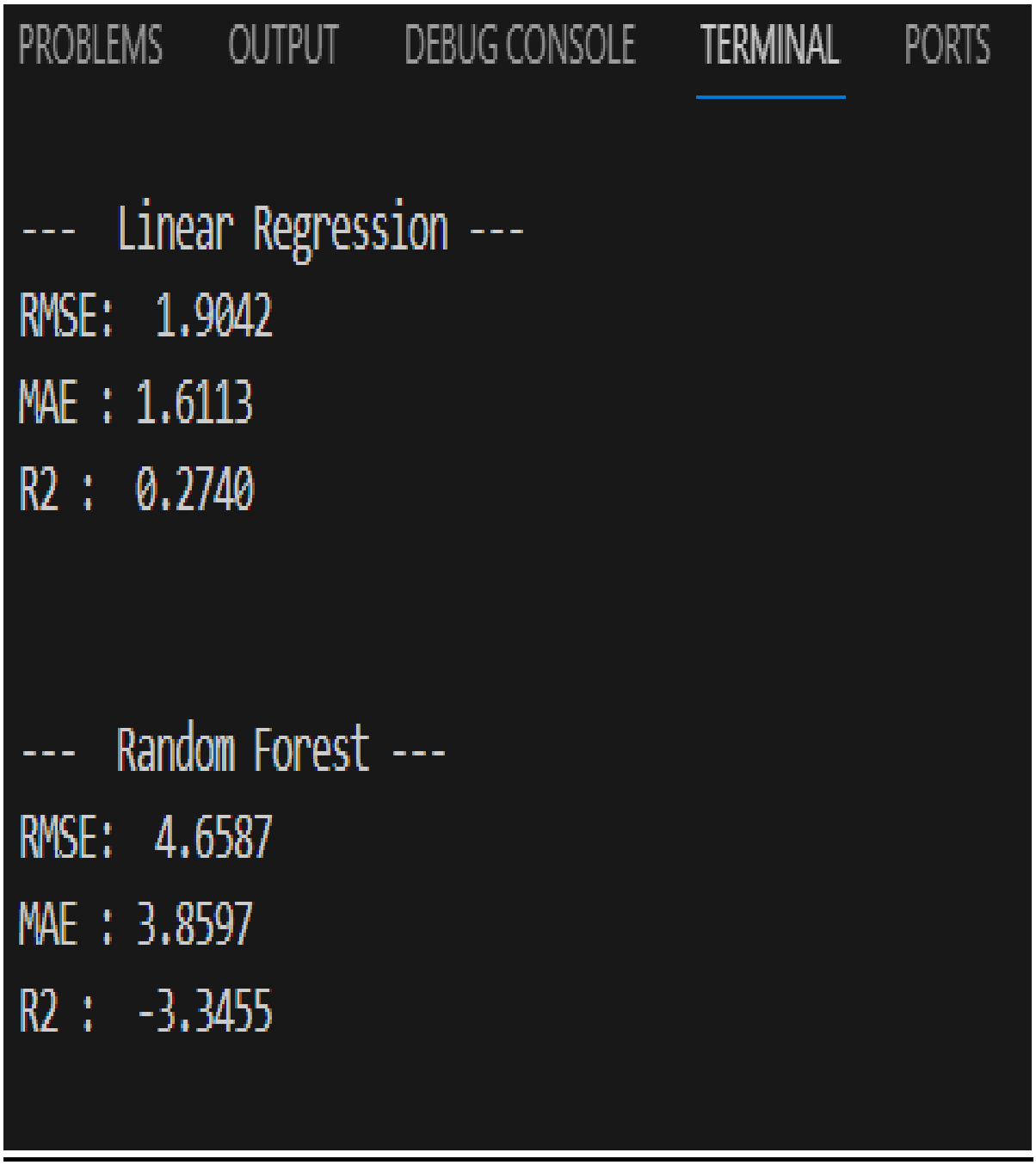
- AAPL Moving Average:- 7-day and 30-day moving averages help smooth price fluctuations and reveal trends.



- AAPL Actual vs Predicted Close Price:- Comparison between actual and predicted closing prices for AAPL stock.



- Evaluation metrics (RMSE, MAE, R^2):-



Insight & Conclusion

The analysis revealed several important findings about stock behavior and prediction:

- **AAPL's price prediction was reasonably accurate**, with Random Forest performing better than Linear Regression in capturing stock movement patterns.
- **Volatility levels differed across companies**, showing that some stocks fluctuate more heavily than others, making them riskier investments.
- **Correlation analysis showed strong positive relationships** among certain stocks, meaning they often move together in the market.

- **Moving averages (7-day and 30-day) provided useful trend signals**, smoothing daily fluctuations and helping identify long-term directions.

- **Distribution of daily returns showed clear risk-return profiles**, where some companies had higher potential gains but also higher risks.