

Forecasting Amazon Stock Prices: A Time-Series Approach

1. Introduction

Stock price forecasting is a critical aspect of financial analysis and investment decision making. In this project, I develop a time-series forecasting model to predict Amazon's stock prices using historical data. The primary focus is on predicting the Volume Weighted Average Price (VWAP), a key metric used by traders and analysts to evaluate stock price trends. By employing the ARIMA model, I aim to provide insights into future stock movements that can aid investors in making informed decisions.

2. Business Problem

The objective of this project is to predict Amazon's VWAP based on past stock market performance. This prediction is valuable to investors and financial analysts who rely on data-driven insights to optimize their trading strategies. The accuracy of stock forecasting plays a crucial role in minimizing financial risks and improving decision making.

3. Data Collection and Preparation

3.1 Data Source

- The dataset was obtained from [NASDAQ](#), containing daily historical data of Amazon's stock.
- The dataset includes the following key features:
 - Date: Trading date.
 - Close: Closing stock price.
 - Volume: Number of shares traded.
 - Open, High, Low: Stock price movements within a trading day.
 - VWAP: A calculated metric representing the average price based on volume and price movements.

3.2 Data Preprocessing

- Date Formatting: Converted the date column to a datetime format for time-series analysis.

- Feature Engineering: Computed rolling means and standard deviations using 3-day and 7-day windows for key variables such as Close, Volume, and High price.
- Train-Test Split: The dataset was divided into 80% training and 20% testing to evaluate model performance.

4. Modeling Approach

4.1 ARIMA Model

- Why ARIMA?
 - ARIMA is a widely used statistical model for time-series forecasting that captures trends, seasonality, and noise in historical data.
 - The Auto.ARIMA function in R was utilized to automatically select the best ARIMA configuration based on Akaike Information Criterion (AIC).
- Model Training & Forecasting:
 - The ARIMA model was trained on VWAP values and engineered features.
 - It generated a forecast for the test set, predicting future VWAP trends.

5. Model Evaluation

To assess model accuracy, I used two standard error metrics:

- Root Mean Squared Error (RMSE): 5.133
- Mean Absolute Error (MAE): 4.329

These metrics indicate that the predicted VWAP values deviate by approximately **4.33 to 5.13 units** from the actual VWAP, showing a reasonably good performance.

6. Challenges and Limitations

6.1 Challenges

- Initially, the project was attempted in Python, but issues related to datetime indexing errors led to difficulties in processing the time series.
- The switch to R allowed for smoother implementation and model training.

6.2 Limitations

- No External Market Factors Considered: The model only relies on historical stock data and does not incorporate macroeconomic events, earnings reports, or news sentiment.
- Limited to Historical Patterns: The model does not adapt well to sudden stock market shifts, which can be influenced by unforeseen external events.

7. Future Enhancements

To improve the accuracy and robustness of the model, we propose the following future directions:

1. Integrate External Variables: Market sentiment, economic indicators, and company earnings reports.
2. Explore Machine Learning Models: Use LSTMs (Long Short-Term Memory networks) for more advanced time-series predictions.
3. Expand to Other Stocks: Apply similar methodologies to predict the stock prices of other companies.
4. Automate Data Updates: Implement API-based real-time data retrieval for continuous model improvement.

8. Implementation Plan

1. Automate Data Collection using stock market APIs.
2. Model Deployment: Train and validate the model periodically.
3. Performance Monitoring: Regularly update and retrain the model as needed.

9. Ethical Considerations

- Transparency: Clearly communicate model limitations to prevent misleading financial decisions.
- Fairness: Ensure the model does not favor particular market participants.
- Responsible Use: Prevent misuse of forecasts in manipulative trading strategies.

10. Conclusion

The ARIMA model provided a reasonable forecast of Amazon's VWAP, offering valuable insights for investors. While the model performed well, further enhancements are necessary to improve accuracy by integrating additional data sources and exploring machine learning-based approaches. By continuously updating and refining the model, we can develop a more robust financial forecasting system.

References

- NASDAQ Stock Market Data: <https://www.nasdaq.com/market-activity/stocks/amzn/historical>