

Mugangasia /  
Time-Series-Analysis---Predicting-Microsoft-Stock-Prices-

<> Code

Issues 4

Pull requests

Actions

Projects

Wiki

Security

0 stars

2 forks

2 watching

Activity

Public repository

main

Branches Tags

Mugangasia

...

24 minutes ago

View code

README.md



## Project Overview

The project is aimed at analyzing and predicting Microsoft's stock prices. The aim of the project is to use predictive models (ARIMA, LSTM, and FB Prophet) to forecast Microsoft's stock prices, improve investment decision-making, optimize trading strategies, provide actionable insights, and enhance decision support tools.

The data used in the project is acquired from Yahoo Finance, containing Microsoft stock market information from January 1, 2010, to June 21, 2023. The dataset includes columns such as Date, Open, High, Low, Close, Volume, and Adj Close, providing insights into the stock's performance and trading activity.

The success of the project will be measured based on prediction accuracy, stability, and robustness of the models, as well as the generation of meaningful insights and interpretations regarding the factors influencing Microsoft's stock prices. Statistical measures such as mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) will be used to assess prediction accuracy.

Overall, the project aims to provide reliable predictions, optimize decision-making processes, and enhance trading strategies in the context of Microsoft's stock prices.

## Background Information

---

The stock market encompasses a series of markets and exchanges wherein the buying, selling, and issuance of publicly traded company shares occur. These financial activities are conducted through institutionalized formal exchanges or over-the-counter (OTC) marketplaces, which adhere to established regulations. Within a country or region, there may exist numerous stock trading venues that facilitate transactions involving stocks and various other types of securities.

## Problem Statement

---

Misinformation and obsolete data in making analysis in the money markets across the world has been and continue to be a problem. The required information to make informed decision, at times becomes too technical for some players in the stock market industry despite the industry players craving for information to accurately make predictions. Our client, a an asset management firm, needs proper analysis on prices on Microsoft's Stock to make informed investment decisions and optimize their trading strategies as they carry out their operations as well as forecast for strategic directions and decisions.

## Objectives

---

- To use predictive models (ARIMA, LSTM, and FB Prophet) to forecast Microsoft's stock prices between (01/01/2010 - 06/21/2023). Our desired outcome is to have a model with a Mean Absolute Error of less than  $10$ .
- Compare the performance of the predictive model against a baseline model(ARIMA), Mean Absolute Error MAE , Root Mean Squared Error (RMSE) , and Mean Absolute Percentage Error (MAPE) . Aim for a significant improvement over the baseline model.
- Optimize trading strategies by incorporating predicted stock prices, specifically using the Moving Averages Bouncing strategy.

- **Provide actionable insights:** Extract meaningful insights from the analysis of historical stock market data and communicate them effectively to stakeholders.
- **Enhance decision support tools:** To create user-friendly interfaces to facilitate easy access to predictions and insights.

## Data Understanding

---

The dataset used in this project contains Microsoft stock market information from 01/01/2010 through 06/21/2023. It is acquired from Yahoo Finance and consists of several columns that provide valuable insights into the stock's performance and trading activity. The key columns in the dataset are as follows:

- **Date:** This column represents the date in the format yy-mm-dd, indicating the trading day for which the stock information is recorded.
- **open:** The "Open" column denotes the price of Microsoft stock at the market open, reflecting the initial trading price for the day.
- **High:** The "High" column indicates the highest price reached by Microsoft stock during the trading day, capturing the peak value achieved.
- **Low:** The "Low" column represents the lowest price reached by Microsoft stock during the trading day, providing insight into the minimum value observed.
- **close:** The "Close" column signifies the final price of Microsoft stock at the market close, reflecting the last traded price of the day.
- **volume:** The "Volume" column denotes the number of shares traded for Microsoft stock on a given trading day, indicating the level of market activity.
- **Adj Close:** The 'Adj Close' column refers to the adjusted closing price of a stock. It is a modified version of the closing price that takes into account various factors such as dividends, stock splits, and other corporate actions.

By analyzing this dataset, we can gain a comprehensive understanding of Microsoft's stock market performance, including the opening and closing prices, the range of prices throughout the day (high and low), and the trading volume. These features are crucial for identifying patterns, trends, and factors that may influence the stock's price movements.

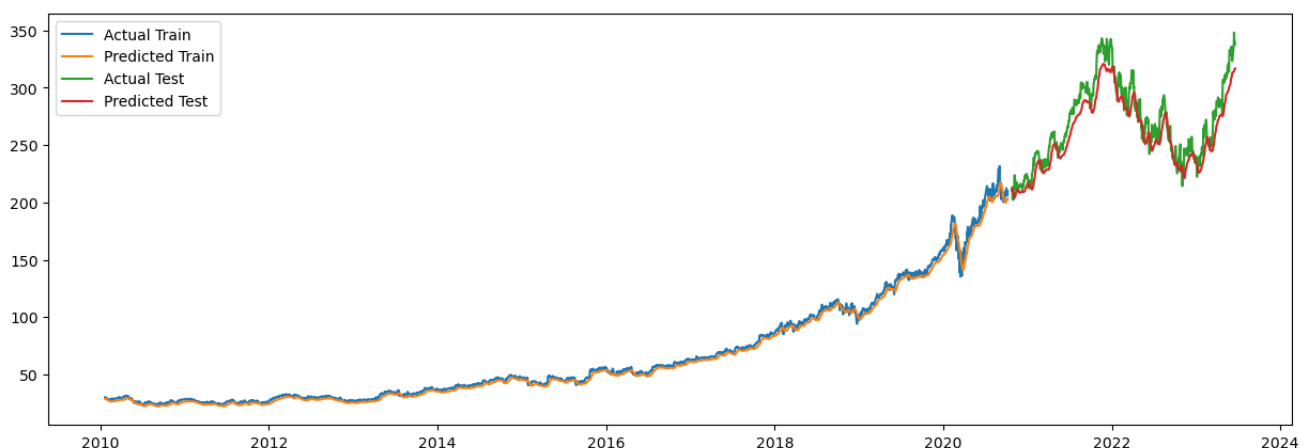
The dataset covers a significant time period, allowing for the exploration of long-term trends and capturing various market conditions. It provides a valuable resource for conducting time series analysis and developing predictive models to forecast future stock prices.

## Sucess Metrics

---

- **Prediction Accuracy:** The model's ability to accurately forecast Microsoft stock prices will be assessed using appropriate statistical measures such as mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE). The lower these values, the higher the prediction accuracy.
- **Stability and Robustness:** The model should demonstrate stability and robustness by consistently providing accurate predictions across different time periods, including both normal and volatile market conditions.
- **Insights and Interpretability:** The project's success also lies in the generation of meaningful insights and interpretations regarding the factors influencing
- Microsoft's stock prices. The ability to identify key patterns and relationships can contribute to a better understanding of the stock market dynamics and aid decision-making processes.

## Results



## Summary of Analysis

### *Inference*

- The LSTM model has a training mean squared error (MSE) of 3.4 , root mean squared error (RMSE) of 1.86 , and mean absolute error (MAE) of 1.484 . These metrics evaluate the performance of the model on the training set. The lower the values, the better the model's accuracy in predicting the target variable.
- The model's testing phase results in a mean squared error (MSE) of 59.239 , root mean squared error (RMSE) of 7.697 , and mean absolute error (MAE) of 5.898 . These metrics measure the model's performance on the test set, providing an indication of its generalization capabilities. Ideally, these values should be similar to the training metrics.

- The relatively higher values for the testing metrics compared to the training metrics suggest that the model may be overfitting the training data. Overfitting occurs when the model learns specific patterns in the training set that do not generalize well to unseen data.
- The plot of actual vs. predicted values for the training set shows that the model is able to capture the general trend and patterns in the data. However, it is unable to capture the extreme values and fluctuations in the data, resulting in a significant deviation from the actual values.

Overall, the model exhibits decent performance on the training set, but its performance decreases when applied to the test set. Further analysis and potential adjustments to the model may be necessary to improve its generalization capabilities.

## Conclusion

---

The following were the key findings from the project:

- The LSTM model performed better than the other models with a Mean Absolute Error of 5.898 . The model was then used to forecast the stock prices for the next 2 years. The forecasted prices were used to optimize trading strategies using the Moving Averages Bouncing strategy. The project also involved the extraction of meaningful insights from the explanatory analysis. The insights include the following:
- The stock price of Microsoft has been increasing over the years.
- The average prices are concentrated between 23.0 and 50. Investors can buy the stock at a lower price and sell it at a higher price between this price ranges for long term investment.

## Limitations

---

- The dataset used for this project was limited to the Microsoft stock prices from 2010 to 2023. This limited the scope of the analysis to the Microsoft stock prices only. The analysis could have been more comprehensive if it included other companies in the same industry.

## Recommendations

---

- Consider incorporating advanced deep learning models as ( NLP ) that can provide insights from company financials, industry trends, news events, and market sentiment to make well-informed investment decisions. This factors change over time even at daily basis that will affect stationarity within the data.

- Regularly Monitor and Update Models: The stock market is dynamic, and market conditions can change rapidly. Continuously monitor the performance of our model and update with new data to ensure effectiveness and relevance. Regularly reevaluate and refining strategies based on market conditions and new information.
- Implement mitigation measures for short term investment to help avoid losses due to volatility and market fluctuations. This can be achieved by using the Moving Averages Bouncing strategy to optimize trading strategies.
- For long term investment the fb prophet model is the best model to use because it has the lowest RMSE value of 71.1654 . This model can be used to predict the stock prices for the next 2 years. The forecasted prices can be used to optimize trading strategies using the Moving Averages Bouncing strategy.

## Project Dependancies

---

- IDE - Google Colab/vscode/sublime text
- Python Libraries - Pandas, Numpy, Matplotlib, Seaborn, Scikit-Learn, fbprophet, statsmodels, tensorflow, pmdarima, plotly

## Collaborators

---

- Marwa Osman - [marwaosman9975@gmail.com](mailto:marwaosman9975@gmail.com)
- Victoria Nabea - [nabeamvictoria@gmail.com](mailto:nabeamvictoria@gmail.com)
- Kennedy Juma - [kenjuma96@gmail.com](mailto:kenjuma96@gmail.com)
- Grace Nekesa - [gracenekesa87@gmail.com](mailto:gracenekesa87@gmail.com)
- Paul Machau - [ndungupaul16@gmail.com](mailto:ndungupaul16@gmail.com)
- Josiah Okumu - [josiahopenji@gmail.com](mailto:josiahopenji@gmail.com)
- Mugangasia Bravin - [bravinmugangasia@gmail.com](mailto:bravinmugangasia@gmail.com)

---

### Releases

No releases published

[Create a new release](#)

---

### Packages

No packages published

[Publish your first package](#)

---

### Contributors 2



Mugangasia Bravin Mugangasia



kenjuma

---

Languages

● Jupyter Notebook 100.0%