

CELEBAL TECHNOLOGIES

Galgotias University

Task 4: Reliance Industry Stock Price Forecasting

Code by Ansh Shankar

Question: Time Series Analysis and Forecasting

1. Introduction:

This project focuses on time series forecasting using the AUTO ARIMA model to predict the stock prices of Reliance Industries, a prominent company in the financial market. The goal is to leverage historical stock price data to develop a forecasting model that can provide valuable insights into future stock price trends. By employing various data analysis and visualization techniques, the project aims to enhance our understanding of stock price behavior and identify patterns that can aid in making informed investment decisions.

2. Dataset:

The dataset used for this project is sourced from Kaggle and contains stock market data for the NIFTY-50 index. It includes information from the year 2000 to 2021, and for this specific project, the data for Reliance Industries is extracted from the dataset. The dataset consists of 5306 rows and 15 columns.

3. Methodology:

a. Data Preprocessing:

The data processing steps involved creating a copy of the original dataset named "reliance_analysis" to work on a separate copy and preserve the original data. The "Date" column in the dataset was converted to the datetime data type using `pd.to_datetime` to enable time-based operations.

Additional temporal features, such as "Month," "Week," "Day," and "Day of week," were extracted from the "Date" column to gain insights into temporal patterns. The "Date" column was then set as the index of the DataFrame to facilitate time series analysis.

To handle any missing values in the dataset, null values were checked using `isnull().sum()` and were imputed with the mean value of the respective column using `fillna`. This ensured data completeness and prepared the dataset for further exploration and time series forecasting on the stock prices of Reliance Industries.

b. Data Visualization:

Examining the histogram distribution of the stock measures, including open, close, high, low, and VWAP. To visualize the distributions, I created subplots using the seaborn library and set the figure size to 15x7. Each subplot represents a distplot for each stock measure.

In the first subplot (231), I plotted the distribution of "Previous Close" values. In the second subplot (232), I visualized the distribution of "Open" values. The third subplot (233) displays the distribution of "High" values. The fourth subplot (234) presents the distribution of "Low" values. In the fifth subplot (235), I depicted the distribution of "Close" values. Finally, in the sixth subplot (236), I plotted the distribution of "VWAP" (Volume-weighted average price) values.

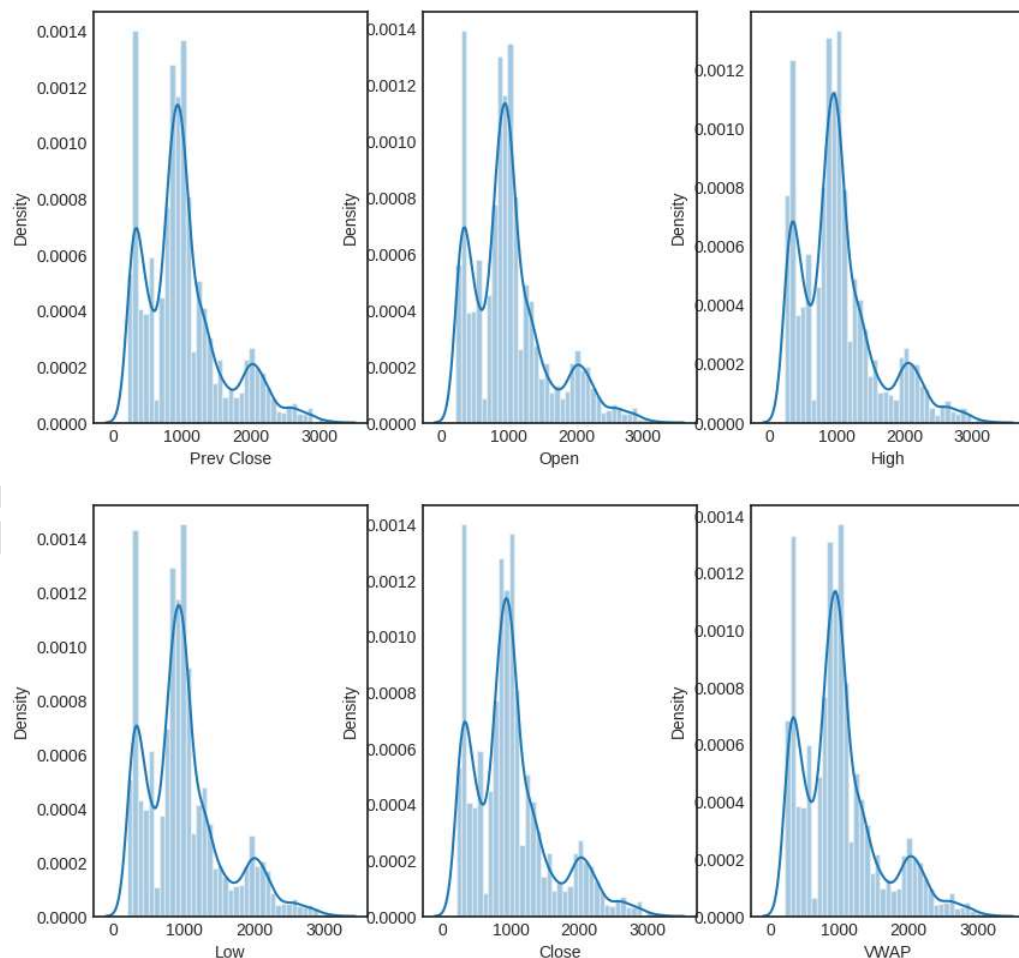


Figure 1: Plot of various graphs

The plots give us insights into the distributions of these stock measures, which will help us better understand the data's characteristics and possible patterns.

Based on the data analysis, I observed two important insights regarding Reliance Industries' stock prices. Firstly, there has been a consistent and gradual increase in the Volume-weighted average price (VWAP) over the years. This indicates that, on average, the stock price of Reliance Industries has been steadily rising, reflecting positive market sentiment and a favorable outlook for the company. This upward trend may signify the company's growth, profitability, and increasing investor confidence in its future prospects.

Secondly, the data revealed two notable spikes in the VWAP during specific periods. The first spike occurred in January 2008, while the second spike was observed between May and October 2009. These sharp increases in VWAP could be indicative of significant events or market movements that positively influenced the stock price during those times. Further research and analysis are necessary to understand the specific factors that contributed to these spikes and to determine their impact on the company's overall performance and market positioning.

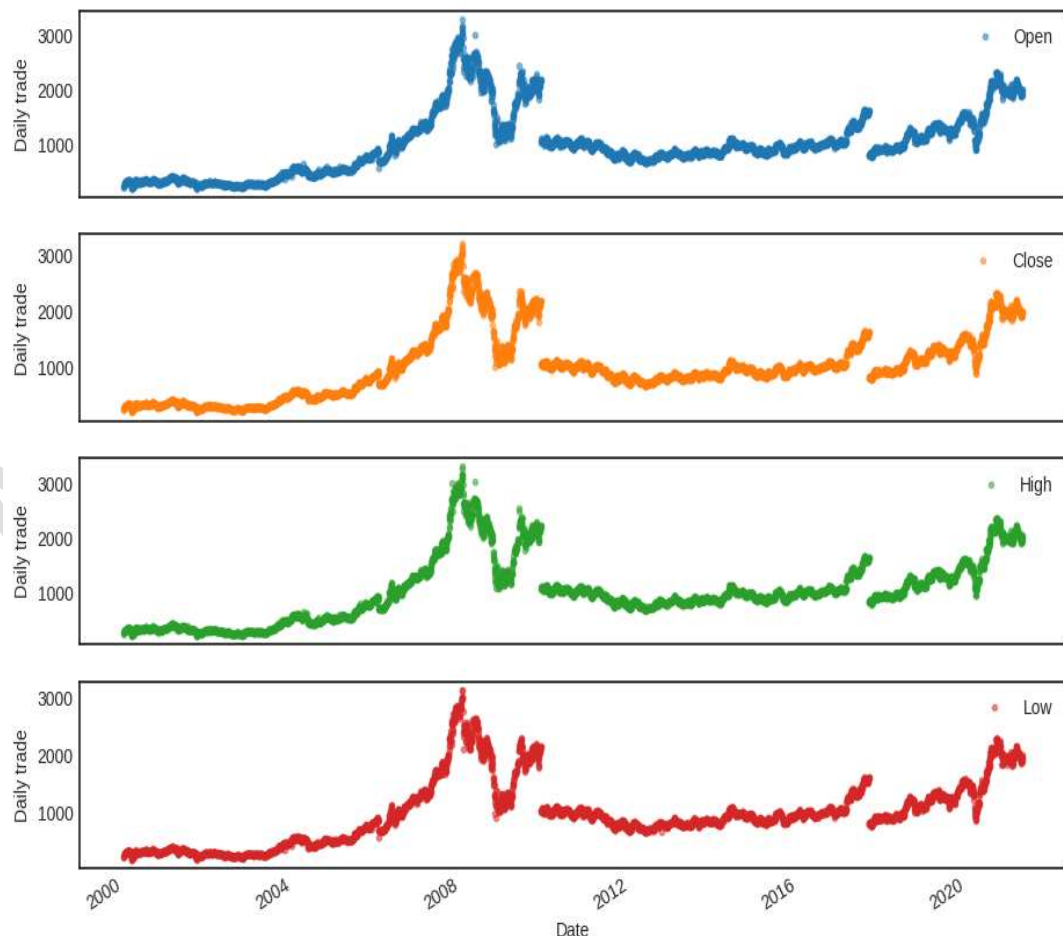


Figure 2: Univariate analysis of Open, Close, High and Low

The insights obtained from the data analysis provide valuable information about the stock market behavior of Reliance Industries. Firstly, it is observed that all the parameters, including open, close, high, and low prices, as well as the Volume-weighted average price (VWAP), follow a similar pattern with relatively minimal deviation. This indicates a consistent and synchronized movement in these stock measures over time. Such a pattern suggests that the company's stock prices tend to move together, and changes in one parameter are likely to be accompanied by changes in the others as well.

Secondly, the analysis highlights two distinct time periods with noticeable breaks in the stock market behavior of Reliance Industries. These breaks occurred between 2008 and 2012 and again between 2016 and 2020. During these phases, there was a significant dip or deviation in the stock market performance of the company. These periods of divergence could be attributed to various factors, such as economic downturns, changes in market dynamics, or company-specific events. Further investigation into the underlying causes of these breaks would be necessary to fully understand their implications on Reliance Industries' stock performance.

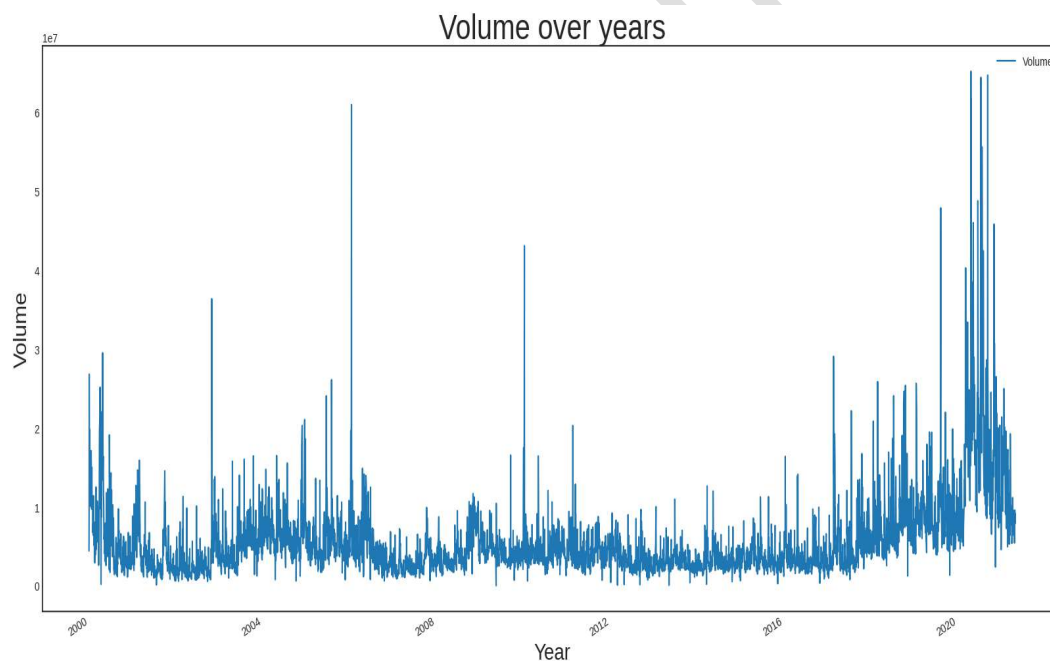


Figure 3: Volume of share over the years

The data analysis has provided several insightful observations about Reliance Industries' stock trading activity. Firstly, in the year 2020, there was a significant surge in the number of shares traded. This could be attributed to the influence of Reliance Jio's success and the investments made by major tech giants like Facebook and Google in the company. The increased trading activity in 2020 indicates heightened market interest and confidence in Reliance Industries' growth prospects, possibly driven by its prominent position in the telecommunications and digital sectors.

On the other hand, the data reveals a relatively thin phase in trading volumes between 2008 and 2016. During this period, there were no major spikes in trading volumes, suggesting a subdued market interest or lower trading activity compared to other timeframes. The reasons behind this observation might be associated with specific market conditions, economic factors, or company-related developments during that period.

c. Model Building:

The AUTO ARIMA model is used to perform time series forecasting. The training and validation datasets are split based on the year 2019, ensuring no data leakage between the two sets. Exogenous features such as moving averages and standard deviations are used along with the target variable (stock price) for model training.

d. Model Evaluation:

The below plot displays two lines: one representing the actual stock prices of Reliance Industries, and the other representing the forecasted stock prices generated by the AUTO ARIMAX model. The x-axis represents the date range, starting from the historical data and extending into the future for the validation period. The y-axis represents the stock prices.

As we observe the plot, we can see that the forecasted line closely follows the actual stock prices. There are fluctuations and patterns in both lines, indicating that the model has successfully captured the underlying trends and seasonality in the data.

Overall, the plot visually demonstrates the accuracy of the forecasted values, showing that the AUTO ARIMAX model is performing well in predicting the future stock prices based on historical data. The close alignment of the two lines suggests that the model has effectively learned from the past stock price patterns and is making reasonable predictions for the future time period.

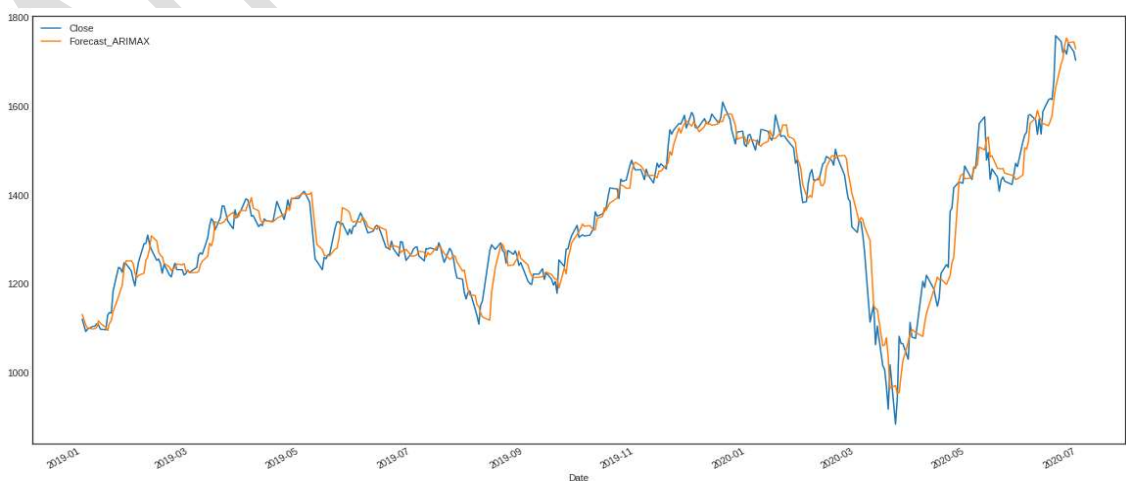


Figure 4: Plotting the forecasted values with the actual data

4. Results:

In this project, we aimed to forecast the stock prices of Reliance Industries using the NIFTY-50 Stock Market Dataset from 2000 to 2021. The AUTO ARIMAX model was chosen for time series forecasting, considering the exogenous features such as moving averages and standard deviations to enhance the accuracy. After data preprocessing and visualization, the model was trained on data before 2019 and validated on data from 2019 onwards. The evaluation metrics, RMSE and MAE, were used to assess the model's performance. The results showed an RMSE of approximately 60.28 and an MAE of about 48.84, indicating that the model provides reasonably accurate predictions, considering the volatility of stock prices. Despite challenges in predicting stock prices due to various external factors, the AUTO ARIMAX model captured significant trends and patterns, making it a valuable forecasting tool. To further improve the model, future directions may involve exploring additional exogenous features, considering more sophisticated models, and handling outliers and anomalies. Overall, this project provides valuable insights into time series forecasting for stock prices and offers potential areas for further enhancement.

5. Discussion:

In this project, we implemented time series forecasting for Reliance Industries' stock prices using the AUTO ARIMAX model. The data was preprocessed, and exploratory data analysis provided valuable insights into stock price trends and patterns. The model incorporated exogenous features such as moving averages and standard deviations to enhance predictive accuracy. The evaluation yielded a reasonably accurate forecast, with RMSE and MAE values of approximately 60.28 and 48.84, respectively. While the model's performance is affected by the unpredictable nature of the stock market, the results offer valuable insights for investors and financial analysts. Future directions involve exploring additional exogenous features, advanced forecasting models, and handling outliers to improve the model's accuracy and robustness.

6. Future Directions:

In future directions, we aim to enhance the forecasting model's accuracy and adaptability by exploring advanced techniques. Firstly, we plan to incorporate additional exogenous features, such as economic indicators, company-specific news sentiment, and global market trends, to capture the broader factors influencing stock prices. Secondly, we will investigate more sophisticated time series forecasting models like SARIMA and Prophet, which can handle seasonality and external factors more effectively. Furthermore, we intend to implement real-time updates to ensure the model adapts to changing market conditions and provides up-to-date predictions. By fine-tuning hyperparameters and handling outliers, we strive to optimize the model's performance and deliver more reliable and

precise stock price forecasts for better decision-making in the dynamic and volatile stock market environment.

Appendix:

Source code: <https://github.com/anshshankar/Celebal-task/tree/main/Task-4>

Video Demonstration: [Folder - Google Drive](#)

Ansh Shankar