Project Deliverable and Report

Background

My previous topic was Text Mining in Business. The first thing to explain is the background and importance of predicting stock prices. The stock market is a complex system influenced by many factors including economic indicators, political events and social trends. Investors and financial analysts rely on accurate forecasts of stock prices to make informed decisions about buying and selling stocks. However, accurately predicting stock prices is a challenging task due to the complexity of the market and the large amount of data available. With the rapid development of technology and the availability of large amounts of financial data, data mining techniques are becoming more and more popular in predicting stock prices. Data mining involves the process of analyzing large data sets to identify patterns and relationships that can be used to predict future events. In the context of the stock market, data mining can help investors and financial analysts make informed decisions about buying and selling stocks. Data mining techniques can be applied to a wide range of financial data, including stock prices, financial statements, news articles, and social media data. By analyzing these data sets, we can identify patterns and relationships between stock prices and various factors such as revenue, earnings per share, and market trends. These insights can be used to make informed decisions about buying and selling stocks, managing portfolio risk and maximizing returns. I'm tracking stock prices in a project. Furthermore, the use of machine learning in data mining makes it possible to identify complex patterns and relationships in data that may not be apparent using traditional statistical methods. In summary, due to the availability of large amounts of financial data and advances in machine learning algorithms, data mining techniques are gaining popularity in predicting stock prices. By applying data mining techniques to financial data, investors and financial analysts can gain valuable insights into market trends, stock behavior, and investor sentiment, allowing them to make informed decisions. My project focuses on forecasting stock prices based on time series analysis. Analyzing and predicting stock prices is an important task in the financial industry. Time series analysis involves analyzing and modeling the time-dependent behavior of data over time. In the context of the stock market, time series analysis can help investors and financial analysts make informed decisions about buying and selling stocks, managing portfolio risk, and maximizing returns. One of the main benefits of time series analysis is its ability to capture trends and patterns in data over time. By analyzing historical stock price data, financial analysts can identify patterns in stock price behavior and predict future trends. This information can be used to make informed decisions about buying and selling stocks at the right time, maximizing returns and minimizing risk.

Problem statement and data description

The problem I solved in the project was to predict the Facebook stock price in 21-22 by analyzing the Facebook stock price before 2021. The dataset I am using is the historical Facebook stock price dataset downloaded from DataWorld, which includes Facebook stock price data for the past 12 to 20 years. The data set consists of 7 columns, date, open, high, low, close, adj_close and volume, as shown in the figure. The Date column represents the date the stock price was recorded, while the Open, High, Low, and Close columns represent the open, high, low, and close prices of Facebook stock on that particular date. The adj_close column represents

the adjusted closing price, which takes into account any dividends or other financial factors that may have affected the stock price. Finally, the "Volume" column indicates the number of shares traded on that particular date. This dataset can be used in time series analysis to predict future stock prices and gain insights into market trends.

Methods

- **Worked**
  I am using an ARIMA (Autoregressive Integrated Moving Average) model to predict the weekly stock price of Facebook. First, preprocess the data by setting the date column as an index, removing unnecessary columns, resampling the data at a weekly frequency, and removing any missing values. Next, train an ARIMA model using data up to the end of 2020. The order of the model is (0, 10, 2), which means it has zero-order autoregressive, ten-difference and second difference-order moving average. Then, I use the trained model to predict the stock price for each week from January 3, 2021, to December 25, 2022. And any missing values in the predictions are filled using linear interpolation. The model is then evaluated by computing the root mean square error (RMSE) between the predicted values and the null array. RMSE is 67.68, I think this number is good for this model. Finally, it visualizes actual and predicted stock prices using a line graph, where actual prices are in blue and forecast prices are in orange.

- **Not worked**
  I also tried to use the "ExponentialSmoothing" function to predict the model, but the RMSE of the running result is as large as 300, so I think this method has a large error, so I didn't use it. In addition, I also tried the decision tree and linear regression methods and got a lot of RMSE, so these methods are not applicable.

Finding

After running the code, I get the result as shown in the figure: firstly, the blue line represents the data ARIMA provided by me is to model this part of the data, and the orange line is the prediction result. And according to the code result RMSE is 67.68, I think this error is acceptable, indicating that the prediction results of the model have certain reliability. It can be found that the beginning of the orange line 2021-01-03 is much lower than the end of the blue line, which means that the stock price will decrease in 21 to 22 years. We also see that the orange line is a horizontal line rather than a zig-zag line, this is because the data is not very seasonal and this happens when the model finds it difficult to predict the future, so it simply takes the previous values The mean and forecast for the future thus yields a straight line. All in all, it is predicted based on the average value of the stock price before 21 years, and the stock price will be greatly reduced from 21 to 22 years, so it is recommended that those who buy Facebook stock sell the stock to reduce losses.

Facebook Stock Price Forecast
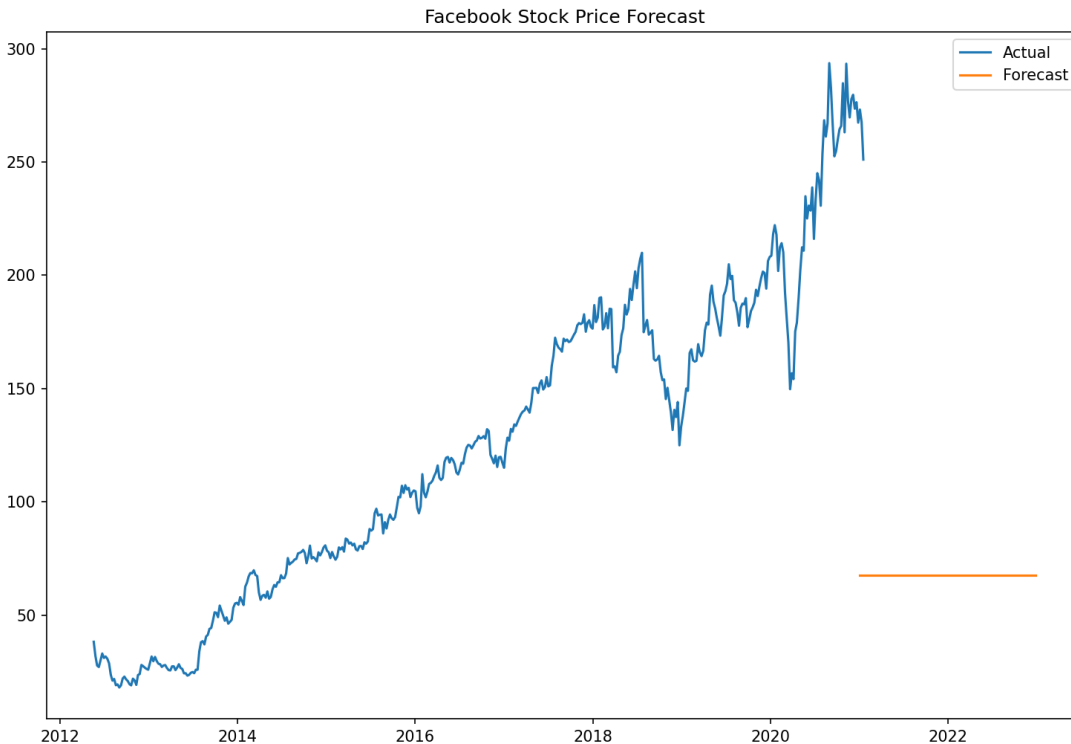
Conclusion and potential future work

Based on the above analysis, the stock price of Facebook will drop sharply from the 21st to the 22nd. Investors are advised to sell Facebook stock. In addition, for Facebook, possible future work could be used to study the reasons for future declines in Facebook's stock price in order to prevent a large decline in the stock price. In addition, the limitations of ARIMA can also be seen from the analysis process. The clarity of an ARIMA analysis depends on the seasonality of the data. If the data is not seasonal or the frequency of the data is difficult to find, then the ARIMA forecast is based on the average of previous data and cannot predict the exact time. Therefore, potential future work might try to address this issue by ignoring seasonality in order to develop the ARIMA algorithm to be more effective.