UCDPA Project Report

**Twitter Stock price prediction using Time series analysis**

Date: 20 Dec 2021 Submitted by: Shoumik Goswami

## GitHub URL (insert URL here)

<https://github.com/shoumikgoswami/UCDPA_shoumikgoswami/tree/main/Stock%20price%20prediction>

## Abstract (Short overview of the entire project and features)

This project has been done as a part of project submission for UCD Specialist Certificate in Data Analytics. The objective of this project is to fetch stock prices data using stock python API and use a simple time series model based on Facebook's Prophet model to predict the stock prices for next 1 year. The project uses Yahoo finance API to fetch prices of the stock since inception and extensive EDA to understand how the prices have changed over the period.

## Introduction (Explain why you chose this project use case)

Twitter had an interesting run after it went public in 2013, where the prices dipped below the IPO price and stagnated for a couple of years. The prices revived after 2018, peaking in 2021 and heading towards a decline again in 2022. This looks like an interesting use case to be able to predict how the prices will behave in the coming years as looking at the trends of this stock, it may head towards another round of decline and stagnation in the coming years.

The project also uses an API based approach of fetching live data for data analysis. This has been done specifically to meet the course requirements of fetching data using multiple techniques.

## Dataset (Provide a description of your dataset and source. Also justify why you chose this source)

The data has been fetching using stock python API on Yahoo finance – yfinance. Link to the library - <https://pypi.org/project/yfinance/>

This API is simple and easy to use, allowing users to fetch data of any stock since inception. This API is a stock python API and as a result is free and does not require any private API keys to access the data.

## Implementation Process (Describe your entire process in detail)

The analysis has been performed in the below steps –

**Analysis pipeline - the OSEMN approach**

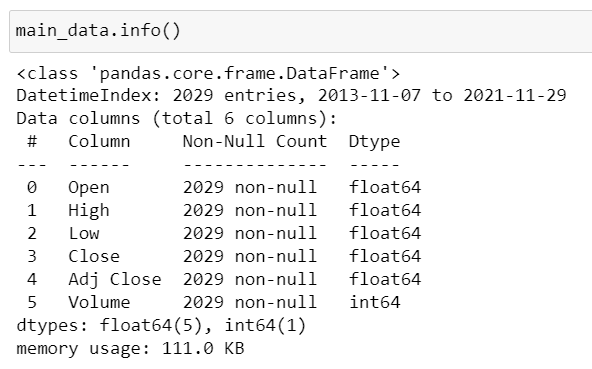
* Obtain the data  
  The data was for Twitter stock exchange was loaded from 1st Jan 2010 to 30 Nov 2021. The stock went public in 2013 so the prices are available since 2013 itself till end of Nov 2021. The data provides daily prices such as Open, Close, High, Low, Adjusted Close and Volume. The prices are only available for trading days.
* Scrubbing / Cleaning the data  
  Once the data is loaded, we check the information of the dataset has 2029 entries and 6 columns. All the data elements are numerical datasets and the index is a datetime index. There are no null values in the dataset so no further cleaning or imputation is required.
* Exploring / Visualizing our data  
  To understand what is in the dataset, we check the descriptive statistics. Then we plot different charts to understand the price trends, volume trends and month-wise variation in the data for the observation period. We try to observe any seasonality in the data as this is key to creating a good model. We also create 50 day and 100 day moving average as a technical indicator to understand the price movements and forecast the prices trend for the future.
* Modeling the data  
  We will be using Facebook’s prophet algorithm to create a timeseries model for this project. The model requires the input to be of a specific format so we transform the dataset. We convert the index to a column so that we have the days in a single column, named **ds.** Then we pick the adjusted closing price and name it to **y.** We then load the prophet model and fit it on the dataset. We set the seasonality to false as we did not observe any seasonality in the dataset. Once the model is created, we use it to forecast the prices for the next 365 days or 1 year. We plot the trends and forecast to see the changes.
* iNterpreting the results  
  While the model does a good job of fitting the existing trends, it does not do well predicting the prices at a future date. This will require further hyper parameter tuning to improve the predictions. (This can be covered as a separate project).

**Environment set-up and loading dependencies**

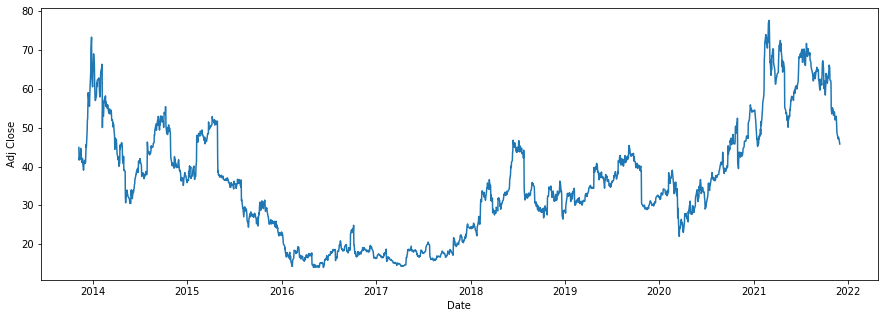
Jupyter notebook is used to do the analysis and Github is used to version the changes.

## Results (Include the charts and describe them)

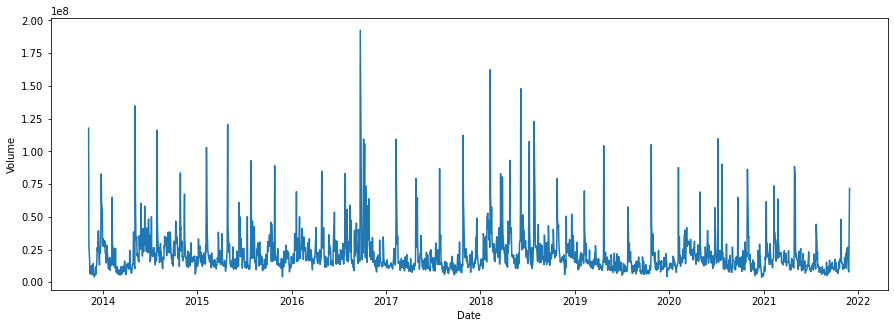
Below are the results based on the analysis



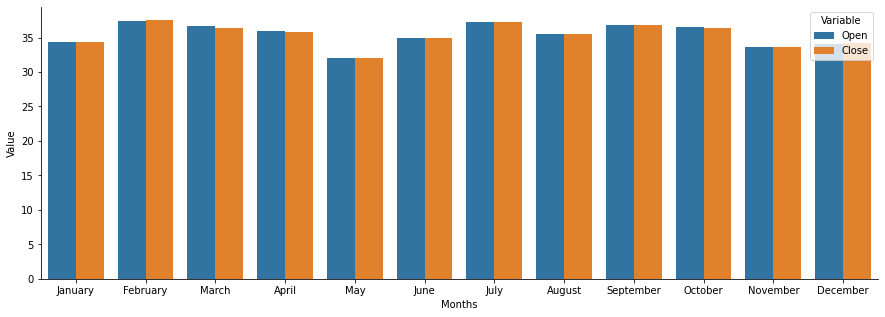
There are no null values in the dataset. Additionally, we have 2029 records in the data - one record per day of Twitter stock price.



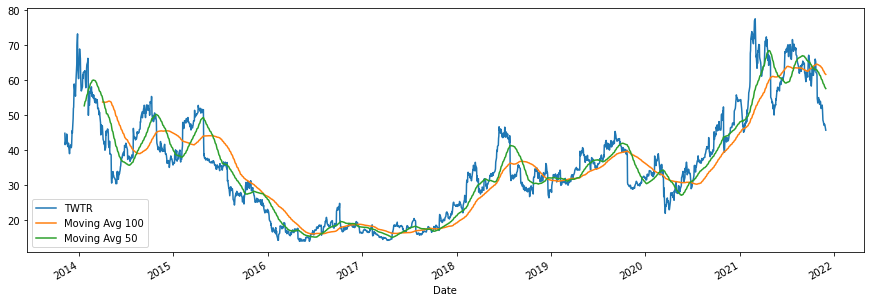
Based on the trend of adjusted closing prices, Twitter's stock price went up exponentially after its IPO but soon it had to go through a market correcting for almost 3 years. After 2018, the stock price started gaining positive momentum until 2021 where it seems to trend downwards.



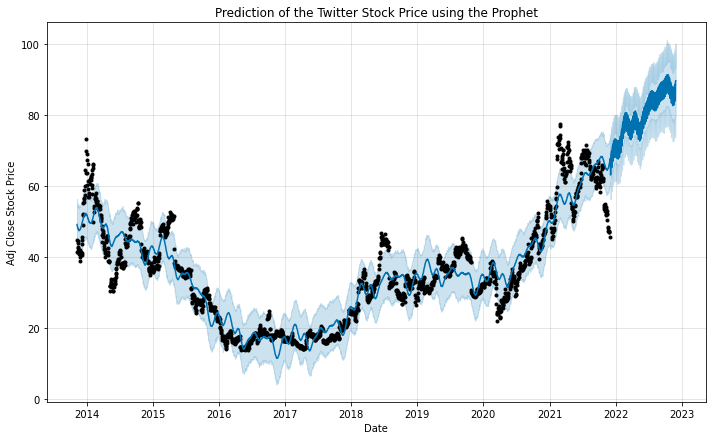
Looking at the volume, the highest volumes were traded during 2017-2018 when the stock price was at its lowest. The volume has declined since 2018.



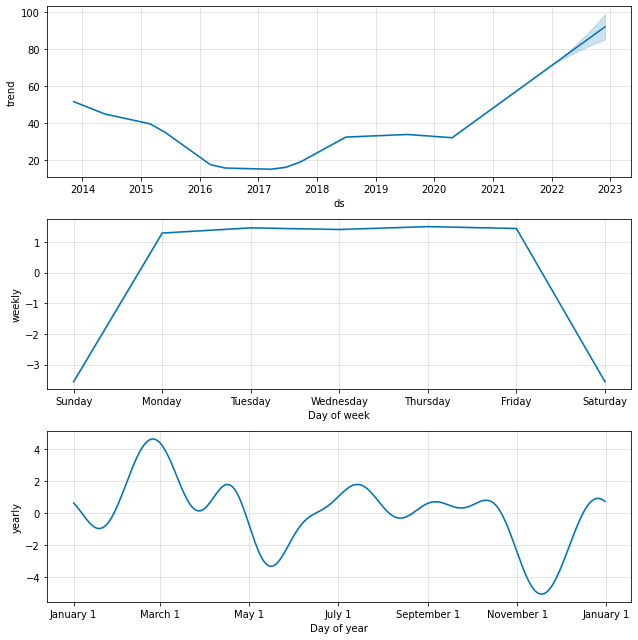
Looking at the month wise plots, there is no visible seasonality in the data. The prices remain nearly constant over the period.



Looking at the 50 day and 100 day moving average technical indicator, it seems the prices of the stock may trend downwards in the future.



While the model does a good job of fitting the existing trends, it does not do well predicting the prices at a future date. This will require further hyper parameter tuning to improve the predictions.



Based on this, there is no specific trend that can be observed in the data. The prices are expected to increase in the coming years as per the model but the future prices will only be accurate once the model is tuned. Additionally, no such trends are observed during the days of the week as the prices tend to remain constant.

## Insights (Point out at least 5 insights in bullet points)

* Based on the trend of adjusted closing prices, Twitter's stock price went up exponentially after its IPO but soon it had to go through a market correcting for almost 3 years. After 2018, the stock price started gaining positive momentum until 2021 where it seems to trend downwards.
* Looking at the volume, the highest volumes were traded during 2017-2018 when the stock price was at its lowest. The volume has declined since 2018.
* Looking at the month wise plots, there is no visible seasonality in the data. The prices remain nearly constant over the period.
* Looking at the 50 day and 100 day moving average technical indicator, it seems the prices of the stock may trend downwards in the future.
* The prices are expected to increase in the coming years as per the model but the future prices will only be accurate once the model is tuned. Additionally, no such trends are observed during the days of the week as the prices tend to remain constant.

## References (Include any references if required)

* <https://facebook.github.io/prophet/>
* <https://analyticsindiamag.com/hands-on-guide-to-using-yfinance-api-in-python/>
* <https://towardsdatascience.com/making-a-trade-call-using-simple-moving-average-sma-crossover-strategy-python-implementation-29963326da7a>