Stock Price Prediction using Time Series Analysis

What is the business problem?

Stock markets are where individual and institutional investors come together to buy and sell shares in a public venue. These exchanges exist as electronic marketplaces. The supply and demand help to determine the price for each security or the levels at which stock market participants, investors and traders are willing to buy or sell. The stock market is very unpredictable. Any geopolitical change can impact the trend of stocks in the share market. Therefore, it is very difficult to do a reliable trend analysis. The most efficient way to solve this kind of issue is with the help of machine learning and deep learning.

One of the best tools used for trend analysis and future prediction is time series forecasting. It is used to predict future values based on previously observed values. In the present problem, we are planning to split the complete stock market in different sectors. Then, we will pick any one sector and apply time series analysis followed by suitable forecasting methods. Finally, we will deploy the entire process across all the sectors. We anticipate that this thorough study will give an overall overview of past stock price trends across the sectors which will help to forecast the future stock price.

Who are the intended stakeholders, and why is this problem relevant to them?

Analysts

Where are the datasets available from?

https://www.kaggle.com/borismarjanovic/price-volume-data-for-all-us-stocks-etfs

This dataset has the information on historical stock prices(last updated 10/11/2017). It contains Date, Open, High, Low, Close, Volume, OpenInt columns.

What data science approaches do you anticipate you will use to model the business problem as a data science problem? (*)

We are planning to use time series forecasting as predictive modelling. Time series data is recorded at regular time intervals, and the order of these data points is important. Here time is an independent variable and the output of a model would be stock price. At first, we are planning to wrangle the data by cleaning out the inconsistencies, imputing the missing values, uniforming the data type, filtering out the specific time period for analysing the stock price etc. This will make the overall data tidy and ready for the next step, i.e., data exploration. Next, we will analyse the data by visualisation and try to find patterns (trend, seasonality, noise) out of it.
In the next step, we will plot correlogram plots (acf and pacf) and try to find out the parameters (p, q) for autocorrelation (AR) and moving average (MA) process. Finally, we will run different ARIMA (p,d,q) models manually and also using pmdarima and prophet model. We will pick the best model out of various models to predict the future stock price of any company.

How do you anticipate that the intended clients will use the results of your CP2 to address the original business problem?

General research of the stock or share market is not enough to make the decision for a new investor. The common trend of the stock market is highly ambiguous for investment; so most of the people are not able to make decisions based on common trends. So, we believe that the thorough analysis using advanced models will help both existing and new investors to understand and make a decision to invest in the share market.