Stock Price Predicting
DS 4002 Study Case by Linda Mu

Background:

In today's interconnected world, the allure of the stock market has captured the attention of many individuals, from seasoned investors to aspiring traders seeking to navigate the complexities of financial markets. With the rise of accessible trading platforms and a wealth of information at our fingertips, more people than ever are intrigued by the potential opportunities and risks that come with investing in stocks. However, within the expansive realm of financial data, accurately predicting stock prices continues to be an exceedingly challenging endeavor for most.

This is where the efficacy of data science, combined with advanced time series analysis methods such as ARIMA (AutoRegressive Integrated Moving Average), becomes evident. By leveraging historical stock price data and employing advanced statistical models, data scientists can uncover hidden patterns and trends that provide valuable insights into future market movements. From recognizing emerging market trends to refining investment strategies, the convergence of stock prediction and data science offers significant potential for practical use, influencing the choices made by investors and financial institutions alike as they strive for improved results in the ever-changing landscape of stock market.

Task:

You are a data scientist tasked with setting up a prediction model for stocks of interest. To accomplish this, you will utilize time series analysis with a model of your choice, based on historical data for several months or a year. Your objective is to develop a predictive model that can forecast future stock prices with a satisfied accuracy, enabling informed decision-making and potentially gaining insights into market trends. This project will require gathering and preprocessing relevant data, selecting appropriate modeling techniques, then train and test the performance of your model. Through this process, you will gain valuable experience in applying data science methodologies to real-world stock data, honing your analytical skills and enhancing your understanding of stock market dynamics.

You will be provided a code file along with 2 cleaned datasets: one for P&G (stable) stock prices and another for Minerva (volatile) stock prices. The provided code utilizes the ARIMA model in R. However, you are encouraged to explore other models and stocks of your interest. Our aim is to achieve a moderate level of accuracy, but there is no strict target for this. As long as you can justify your satisfaction with the results and avoid overfitting, you have the flexibility to experiment and refine your approach.

Reference:

- General information about price prediction:
  "The importance of price prediction | Blog Future Processing," www.future-processing.com, Mar. 07, 2023. https://www.future-processing.com/blog/the-importance-of-price-prediction/ [Accessed Mar 5.2024]
- General information about ARIMA model in R:
  "RPubs - Time series analysis stock market prediction using ARIMA Model in R," rpubs.com. https://rpubs.com/HassanOUKHOUYA/Time_series_analysis_stock_market_prediction_using_ARIMA_Model_in_R [Accessed Mar 5.2024]