**Google Stock Price Prediction Project**

Hareesh Vijayanandan Nair

MSCC 696: Data Science Practicum II

John Koenig

May 1, 2022

**Problem or Situation**

Stock is a major investment option. Most of the retirement fund is also modelled in such a way that it is also used as a stock investment. This scenario gives more challenges to normal middleclass people on their future investment. People are trying to protect their hard-earned money without much loss. Modelling and prediction of stock price is getting more attention. This project is to research on how to predict Google stock price as a model.

**Research Question**

How to understand and predict the stock price impact during the volatile market where national and international affairs are bringing impacts. This prediction model should help the users to avoid any loss on stock investment and maximize the profit.

**Data**

Google Stock data set for this project is obtained from yahoo finance stock data.

|  |  |
| --- | --- |
| **Field** | **Description** |
| Date | When the transaction is processed. |
| Open | The open price when the stock exchange is opened. |
| Close | The close price when the stock exchange is closed. |
| High | High price of stock on this day. |
| Low | Low price on this day. |
| Volume | The total volume of the stock transacted on this day. |
| Adj Close | The adjusted closing price amends a stock close price. |

**Proposed methodology**

Long short-term memory (LSTM) with Recurrent neural Network (RNN) is used for this prediction. This model proved to be a good candidate for time series data prediction.

**Results**

Exploratory Data Analysis (EDA) and Data Visualization plots are used to ensure the Data validation and stability. Multiple plots are generated to ensure the validation result. Augmented Dickey Fuller test (ADF) and Kwiatkowski-Phillips-Schmidt-Shin Test (KPSS) is conducted to ensure the data is reliable for a Time series prediction. Google Stock Price Predication is conducted using LSTM and RNN.

**Google Stock Price Action from 2010 to 2022 - Open Stock**

Chart, line chart

Description automatically generated

**Google Stock Price Action from 2010 to 2022 – Volume of stock traded in millions**

Chart, line chart

Description automatically generated

**Augmented Dickey Fuller Test (ADF) and KPSS Result**

Both of the tests p value is less than .05 which proves it is a timeseries data set.

|  |  |
| --- | --- |
| Chart  Description automatically generated | Text  Description automatically generated with medium confidence |

**Prediction Diagram and LSTM output**

|  |  |
| --- | --- |
|  | Chart, line chart  Description automatically generated |

**Conclusions**

* From RNN and LSTM method the plot looks almost similar to the validation results.
* This model can be further improved and can be used to predict the Google stock price results.
* On Further tuning with other larger data sets this model can be reused for other stock price predictions also.

**References**

* Data - <https://finance.yahoo.com/>
* Design - <https://learning.oreilly.com/library/view/python-for-finance/9781492024323/ch15.html#ts_clustering>
* Code References - <https://www.kaggle.com/search?q=google+stock+price+in%3Anotebooks>