Project Overview :

This project involves the analysis and modeling of stock data using machine learning. The project is divided into two main files: one for data collection and the other for building and evaluating machine learning models.

1. Data Collection (Yfinance Data Retrieval)

The data collection process involves obtaining stock data for 495 stocks using the Yfinance library. To reduce the hassle of running the data retrieval code each time, the data has been pre-downloaded and saved as a CSV file. Additionally, a detailed PDF document is provided, containing the code and outputs of each cell in the data collection file.

Data Retrieval Code File: data_collection.ipynb
Data CSV File: stock_data.csv
Code and Output PDF: data_collection_code_outputs.pdf

Link to access the data_collection.ipynb :
https://colab.research.google.com/drive/1m2JcD7tcylagTfT57wyX_1uhHjNU52zN?usp=sharing

2. Machine Learning Models (LR - Linear Regression)

The machine learning part focuses on building and evaluating models using the stock data obtained. The code is organized in a Colab notebook named LR.ipynb. To run this notebook successfully, follow the instructions below:

Download the Dataset CSV File:
Use the provided stock_data.csv file obtained from the data collection process.
Upload Dataset to Google Drive:
Upload the stock_data.csv file to your Google Drive.
Mount the LR Colab File to Google Drive:
Open the LR.ipynb file using Google Colab and mount it to your Google Drive to access the dataset.

Update Dataset Path in LR Code:
In the Colab notebook, locate the line where the dataset is loaded (pd.read_csv('path_to_dataset.csv')) and update it with the correct path to the stock_data.csv file in your Google Drive.
Run Each Cell:
Execute each cell in the notebook to perform the necessary data preprocessing, model training, and evaluation.
Code and Output PDF:
Refer to the accompanying PDF document (LR_code_outputs.pdf) for a detailed overview of the code and outputs in each cell of the LR Colab notebook.
Additional Notes:
Make sure to have the required libraries and dependencies installed before running the code.

Link to access the LR.ipynb:
https://colab.research.google.com/drive/1wfviB3yYCeWiqTsQVuGeZ4rkzWvY8AjS?usp=sharing