Stock Price Forecasting: Using LSTM

Sanskriti Sharma (21100BTCSAII09438)
Harshit Rajpal (20100BTCSAII07162)
Vivek Handa(20100BTCSAII07198)

Introduction

Welcome to the presentation on Stock Price Forecasting: Using LSTM .Time Series forecasting & modelling plays an important role in data analysis. Time series analysis is a specialized branch of statistics used extensively in fields such as Econometrics & Operation Research. Time Series is being widely used in analytics & data science. Stock prices are volatile in nature and price depends on various factors. The main aim of this project is to predict stock prices using Long Short Term Memory (LSTM) with Least mean squares (LMS) Algorithm. By leveraging LSTM, we can enhance the accuracy of predictions and enable more informed decision-making in the financial market.

Introduction

LSTM, a type of recurrent neural network, is designed to capture long-term dependencies and patterns in data, making it particularly effective for time series forecasting such as stock prices.Its ability to remember important information over longer periods of time allows it to make more accurate predictions compared to traditional forecasting models. Additionally, LSTM can handle non-linear relationships and adapt to changing market conditions, making it a powerful tool for investors and financial analysts. With the use of LSTM, investors can make more informed decisions and optimize their portfolio strategies based on reliable and accurate predictions of stock prices. By incorporating

Literature Survey

1) Study on the prediction of stock price based on the associated network model of LSTM

The prediction methods can be roughly divided into two categories, statistical methods and artificial intelligence methods. Statistical methods include logistic regression model, ARCH model, etc. Artificial intelligence methods include multi-layer perceptron, convolutional neural network, naive Bayes network, back propagation network, single-layer LSTM, support vector machine, recurrent neural network, etc. They used Long short-term memory network (LSTM). Long short-term memory network: Long short-term memory network (LSTM) is a particular form of recurrent neural network (RNN).

Working of LSTM:

LSTM is a special network structure with three "gate" structures. Three gates are placed in an LSTM unit, called input gate, forgetting gate and output gate. While information enters the LSTM's network, it can be selected by rules. Only the information conforms to the algorithm will be left, and the information that does not conform will be forgotten through the forgetting gate.

The experimental data in this paper are the actual historical data downloaded from the Internet. Three data sets were used in the experiments. It is needed to find an optimization algorithm that requires less resources and has faster convergence speed.

Literature Survey

2) An innovative neural network approach for stock market prediction

An innovative neural network approach for stock market prediction Xiongwen Pang1 Yanqiang Zhoul Pan Wang1 Weiwei Lin2 · Victor Chang3

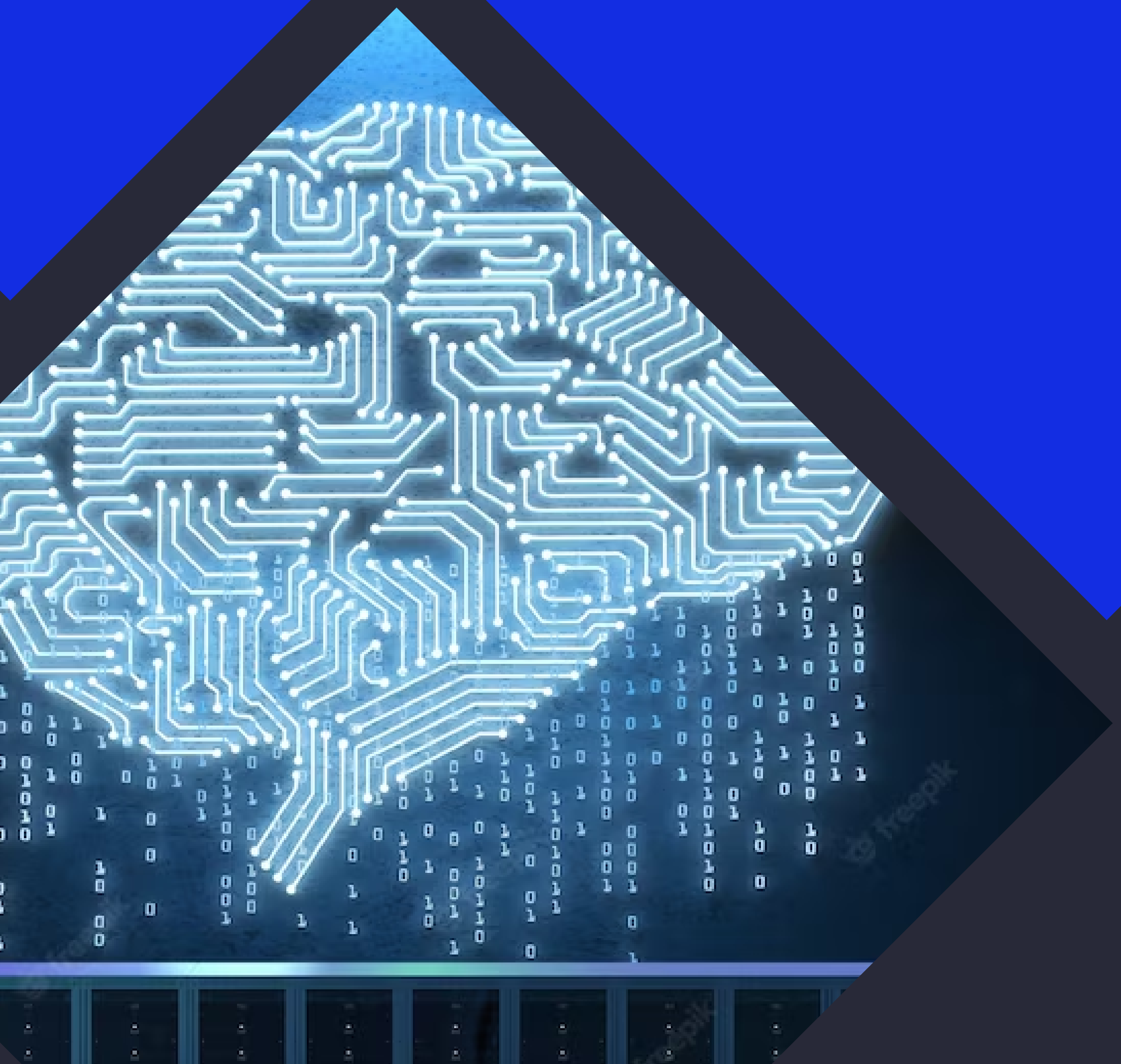
Used Long Short-term Memory (LSTM) with embedded layer and the LSTM neural network with automatic encoder.

LSTM is used instead of RNN to avoid exploding and vanishing gradients.

In this project python is used to train the model, MATLAB is used to reduce dimensions of the input. MySQL is used as a dataset to store and retrieve data.

The historical stock data table contains the information of opening price, the highest price, lowest price, closing price, transaction date, volume and so on.

The accuracy of this LSTM model used in this project is 57%.

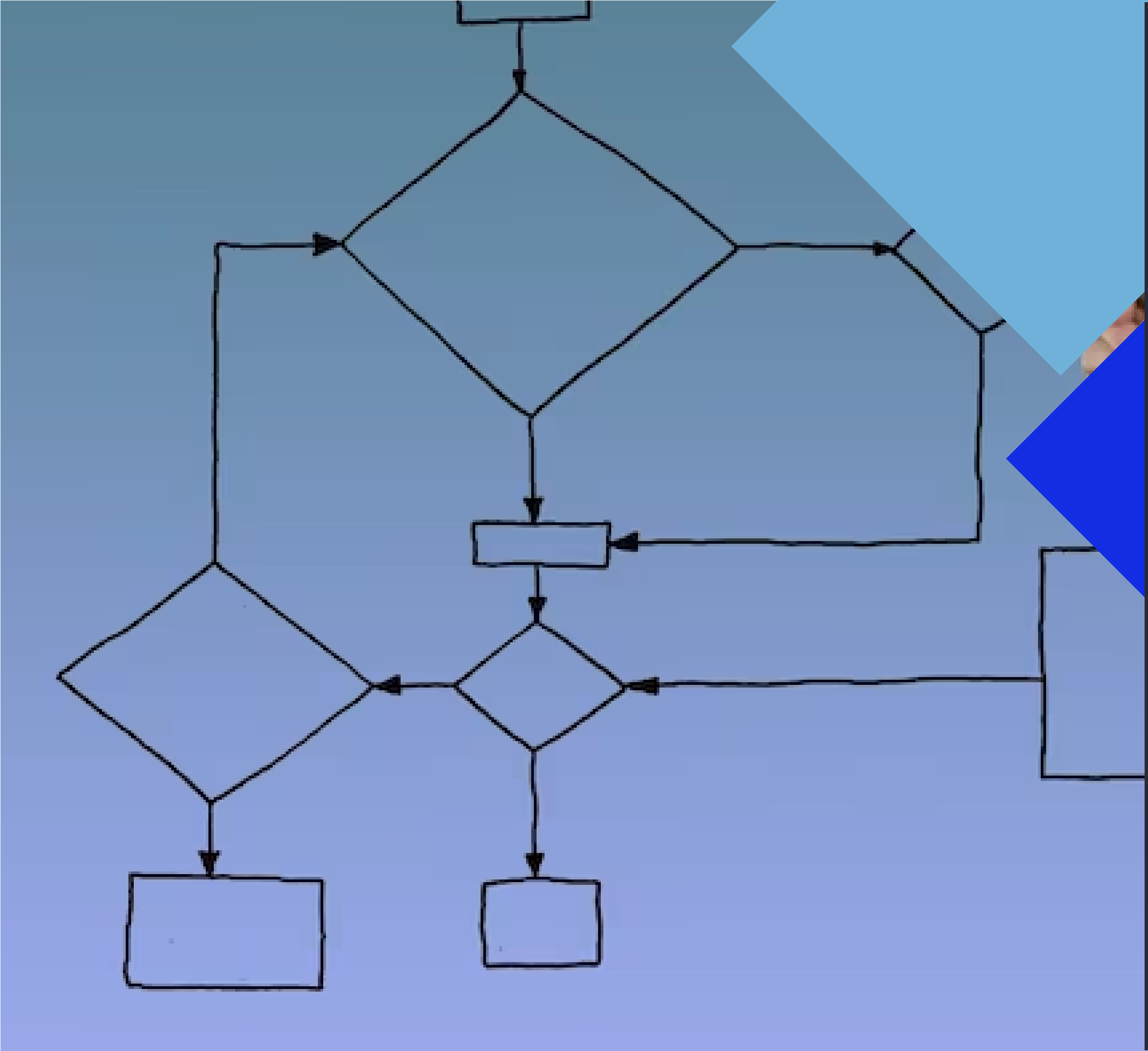


Understanding LSTM

LSTM is a type of recurrent neural network (RNN) that excels in capturing long-term dependencies in sequential data. It overcomes the limitations of traditional models by utilizing specialized memory cells and gates to handle information flow. LSTM's ability to retain important patterns in time series data makes it a valuable tool for stock price forecasting. In addition to its applications in stock price forecasting, LSTM can also be used for natural language processing tasks such as sentiment analysis, text generation, and machine translation. Its versatility and effectiveness in handling sequential data make it a popular choice among researchers and practitioners in the field of deep learning.

LSTM Model Architecture

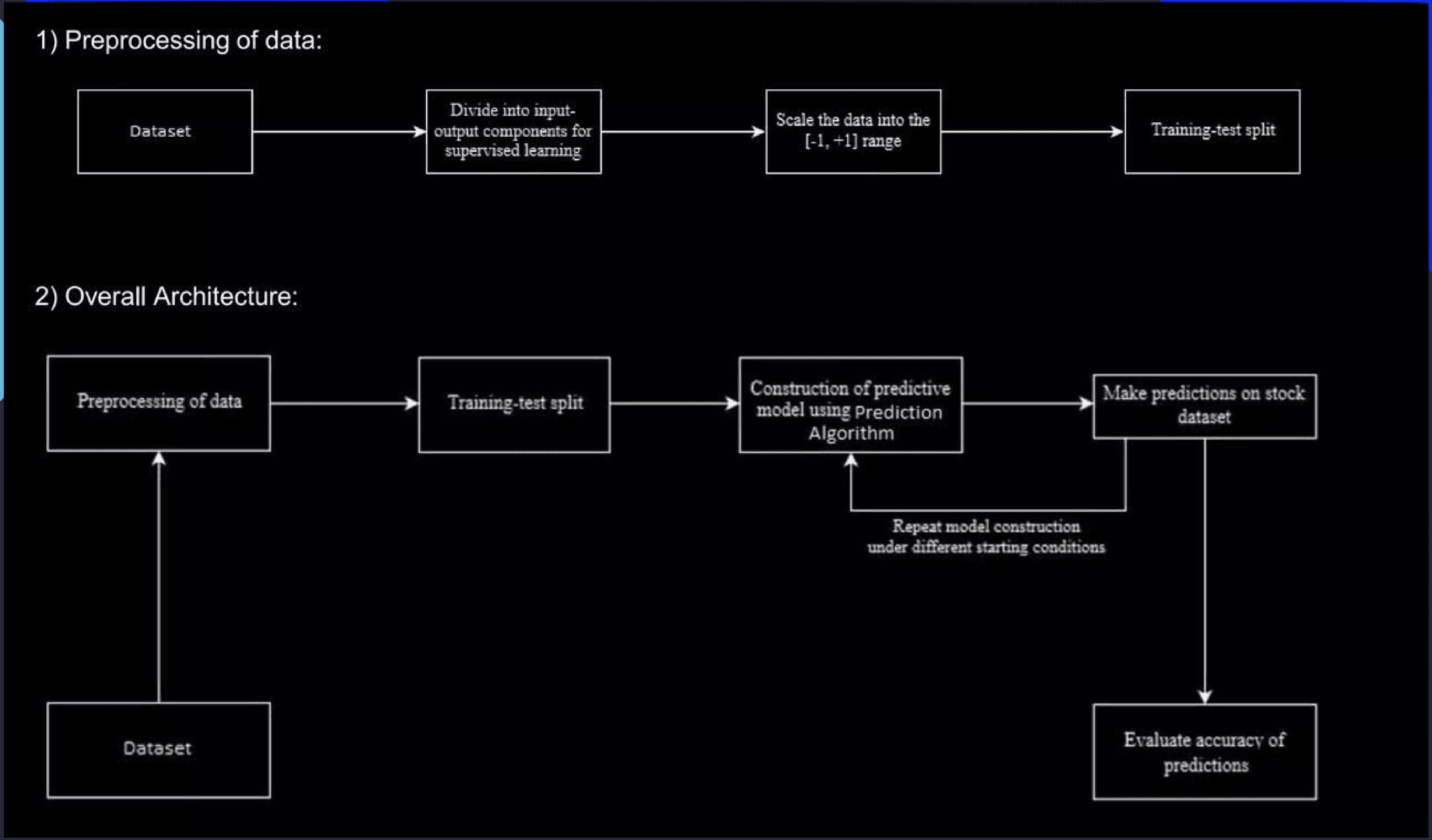
The LSTM model consists of stacked LSTM layers, which learn hierarchical representations of the input data. These layers are followed by one or more dense layers for prediction. The model's hyperparameters, such as the number of LSTM units and the learning rate, need to be carefully tuned to achieve optimal performance. To further enhance the model's performance, additional techniques such as dropout regularization and batch normalization can be applied. Furthermore, the input data can be preprocessed using techniques like feature scaling or one-hot encoding to improve the model's ability to learn from the data. Additionally, different activation functions can be explored for the LSTM units, such as sigmoid or tanh, to capture different types of patterns in the data. Overall, the LSTM model offers a powerful tool for sequential data analysis



Data Preprocessing

Preparing data for LSTM involves normalizing the stock price values to a common scale, splitting the dataset into training and testing sets, and constructing input sequences. Feature engineering techniques such as lagging and windowing can be applied to capture relevant patterns and improve forecast accuracy. In addition, it is important to preprocess the data by removing outliers and handling missing values. Exploratory data analysis can also be performed to gain insights into the underlying patterns and relationships in the data. Furthermore, feature selection methods such as correlation analysis and forward/backward selection can be used to identify the most relevant variables for the LSTM model. Finally, model evaluation techniques such as cross-validation and performance metrics like mean squared error can be employed to assess the accuracy and robustness of the LSTM model.

System Architecture

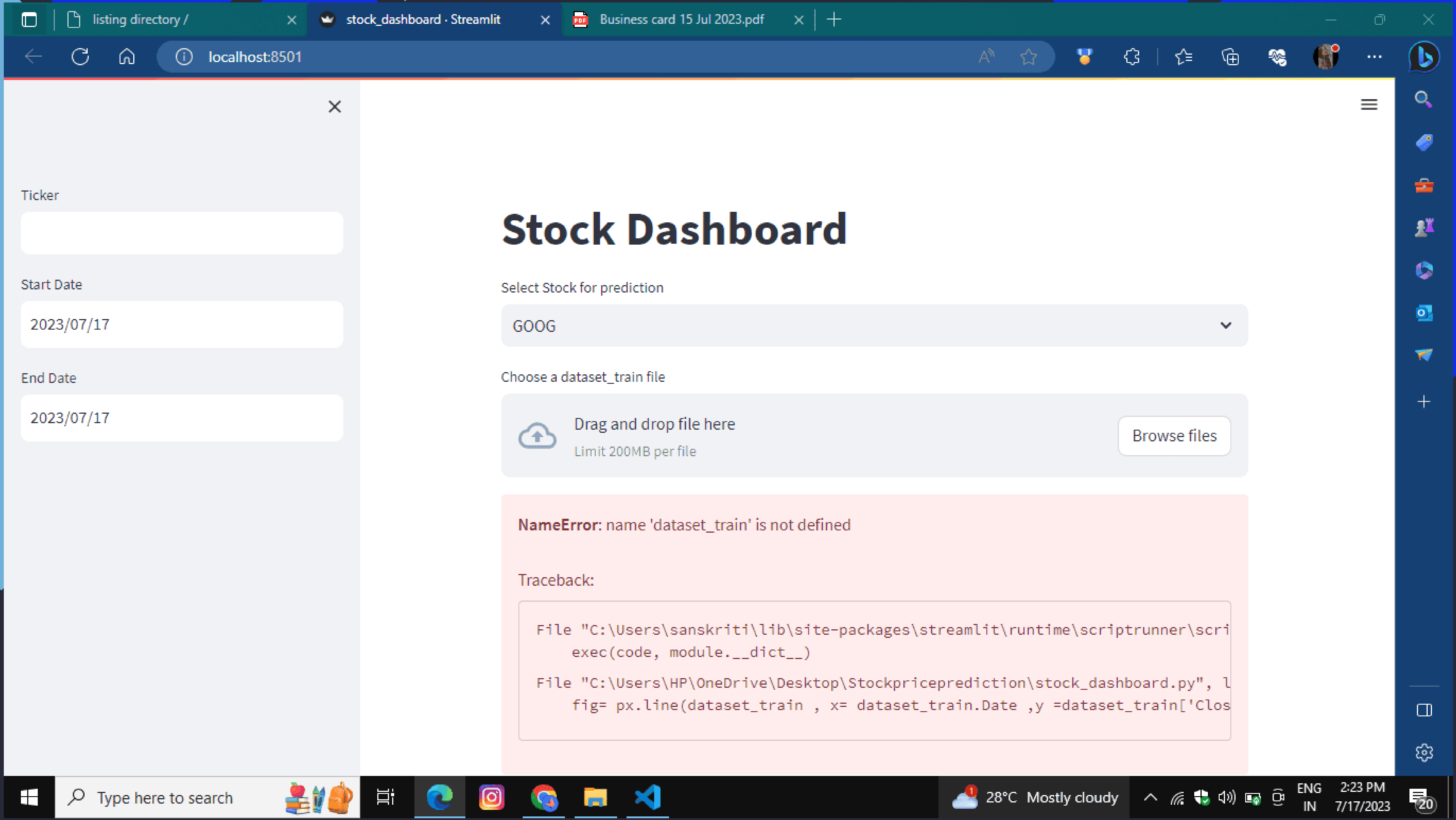


Importing Datasets

https://github.com/Sanskriti2227/Revolutionizing-Stock-Price-Forecasting-Using-LSTM-as-a-Game-Changer-

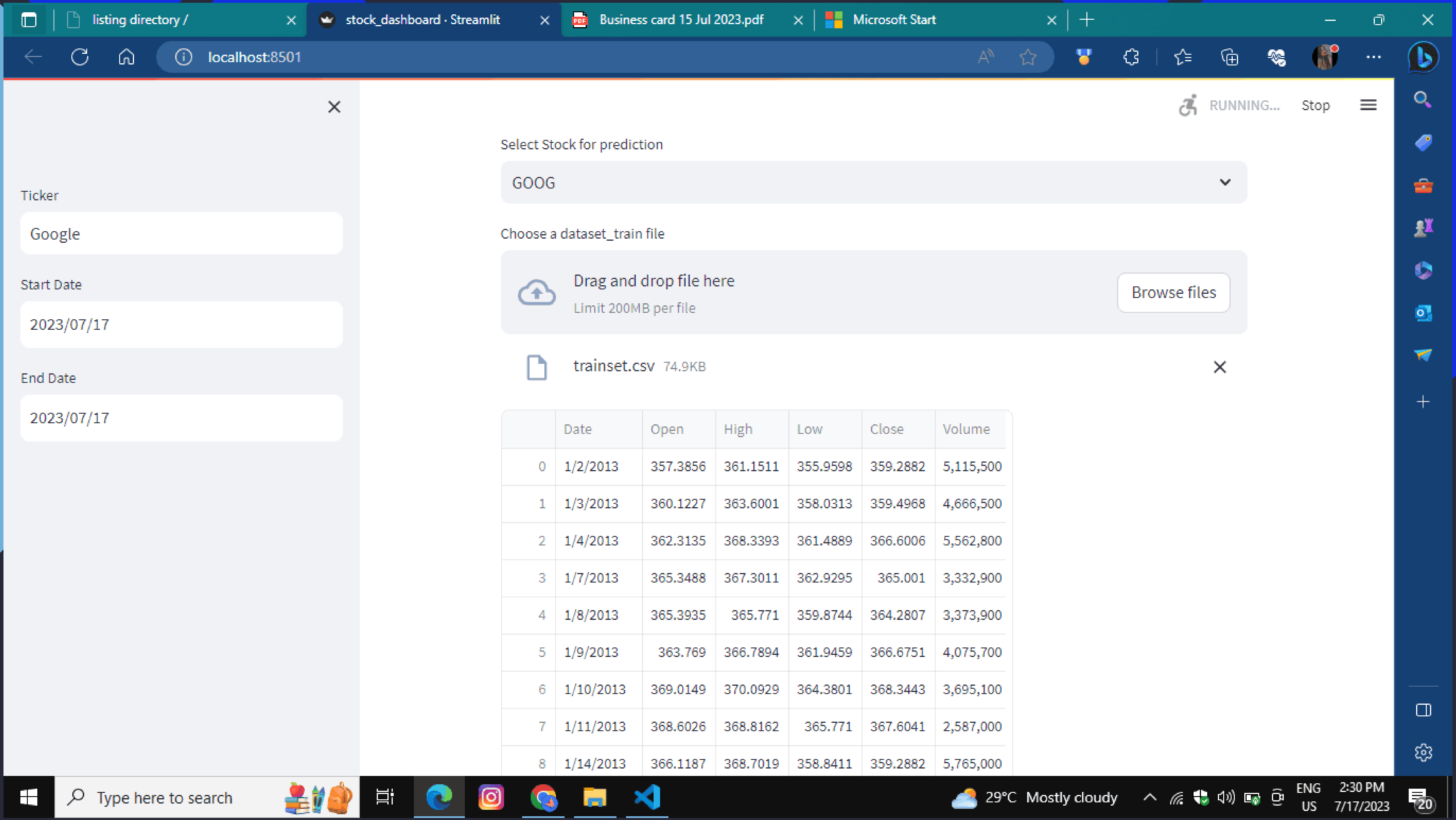
pd.read_csv('./Stock-Price-Prediction/datasets/tw_spydata_raw.csv')

Home Page



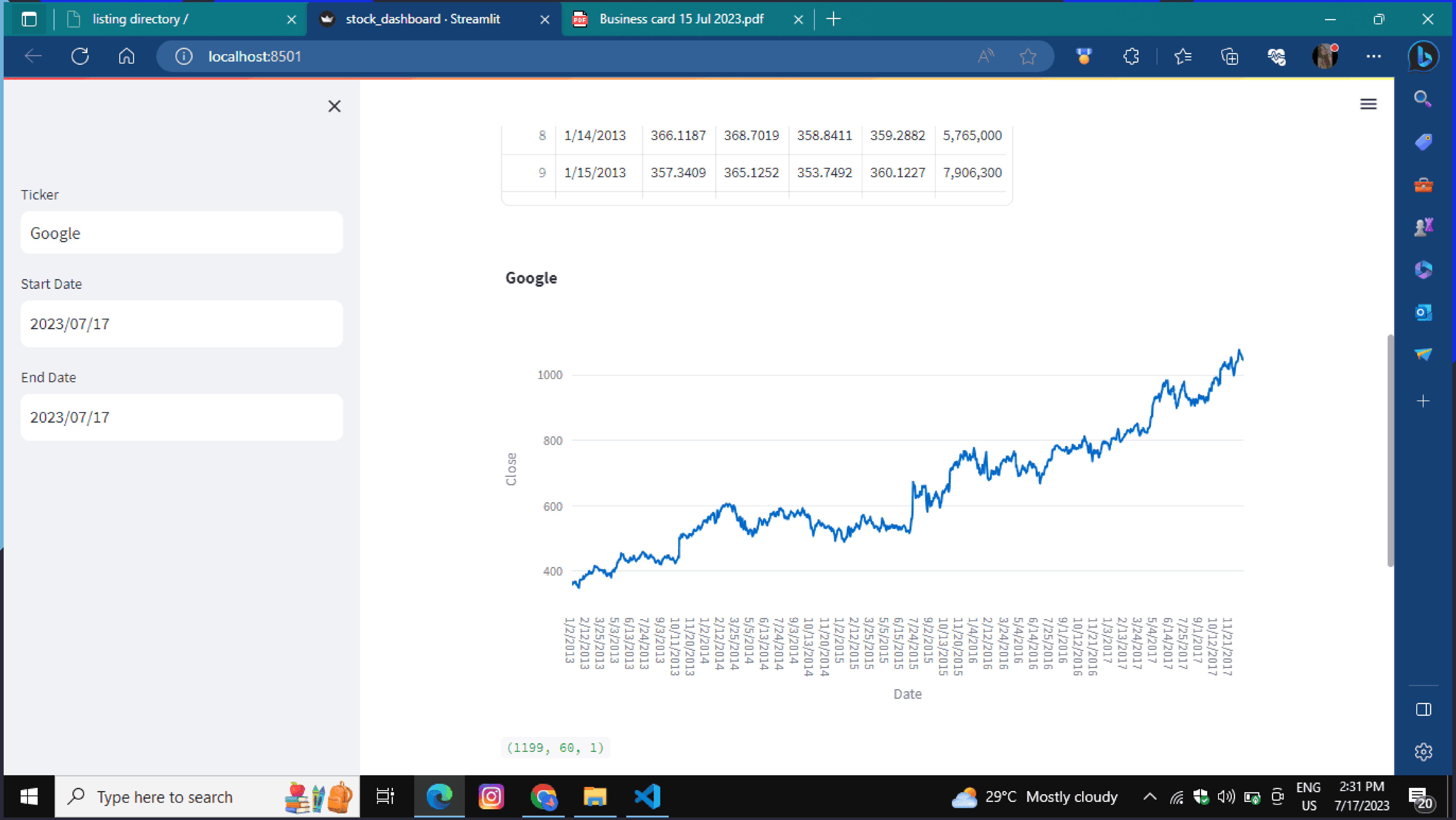
listing directory /
stock_dashboard · Streamlit
Business card 15 Jul 2023.pdf
localhost:8501
Ticker
Start Date
2023/07/17
End Date
2023/07/17
Stock Dashboard
Select Stock for prediction
GOOG
Choose a dataset_train file
Drag and drop file here
Limit 200MB per file
Browse files
NameError: name 'dataset_train' is not defined
Traceback:
File "C:\Users\sanskriti\lib\site-packages\streamlit\runtime\scriptrunner\scri
exec(code, module.__dict__)
File "C:\Users\HP\OneDrive\Desktop\Stockpriceprediction\stock_dashboard.py", l
fig= px.line(dataset_train , x= dataset_train.Date ,y =dataset_train['Clos
Type here to search
28°C Mostly cloudy
ENG
IN
2:23 PM
7/17/2023

While Training

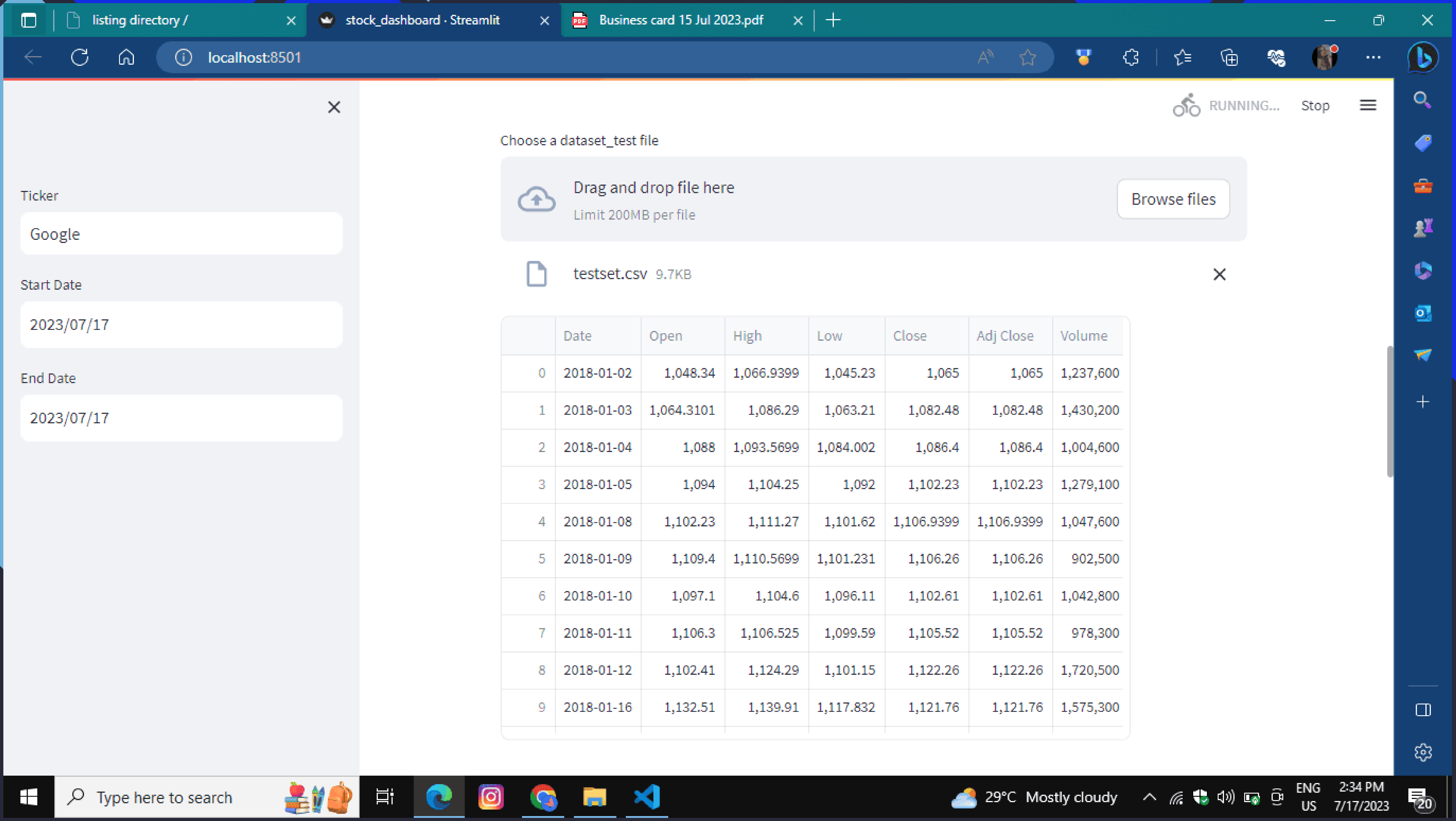


	Date	Open	High	Low	Close	Volume
0	1/2/2013	357.3856	361.1511	355.9598	359.2882	5,115,500
1	1/3/2013	360.1227	363.6001	358.0313	359.4968	4,666,500
2	1/4/2013	362.3135	368.3393	361.4889	366.6006	5,562,800
3	1/7/2013	365.3488	367.3011	362.9295	365.001	3,332,900
4	1/8/2013	365.3935	365.771	359.8744	364.2807	3,373,900
5	1/9/2013	363.769	366.7894	361.9459	366.6751	4,075,700
6	1/10/2013	369.0149	370.0929	364.3801	368.3443	3,695,100
7	1/11/2013	368.6026	368.8162	365.771	367.6041	2,587,000
8	1/14/2013	366.1187	368.7019	358.8411	359.2882	5,765,000

Training Completed

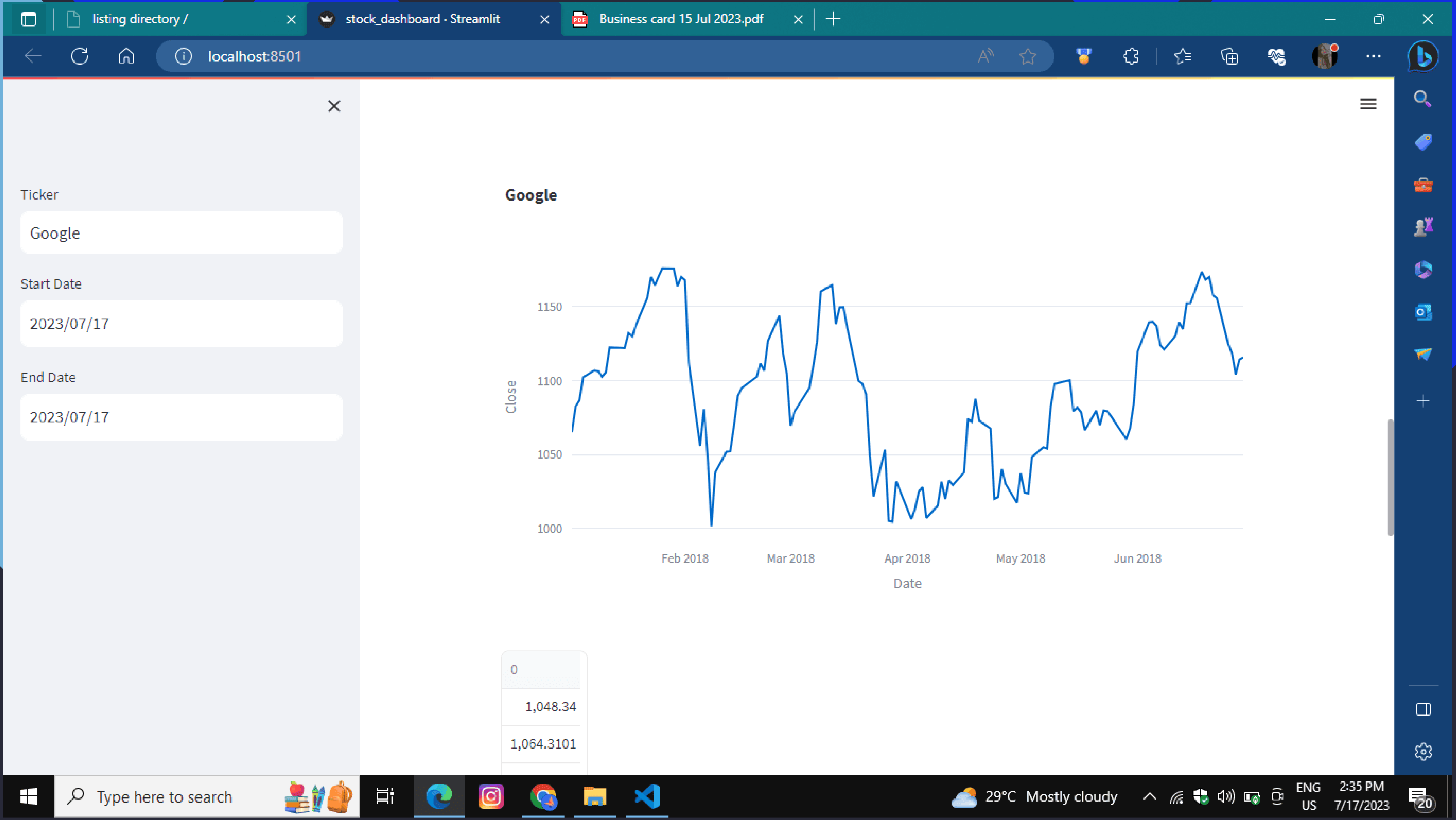


Testing Dataset -While Testing



	Date	Open	High	Low	Close	Adj Close	Volume
0	2018-01-02	1,048.34	1,066.9399	1,045.23	1,065	1,065	1,237,600
1	2018-01-03	1,064.3101	1,086.29	1,063.21	1,082.48	1,082.48	1,430,200
2	2018-01-04	1,088	1,093.5699	1,084.002	1,086.4	1,086.4	1,004,600
3	2018-01-05	1,094	1,104.25	1,092	1,102.23	1,102.23	1,279,100
4	2018-01-08	1,102.23	1,111.27	1,101.62	1,106.9399	1,106.9399	1,047,600
5	2018-01-09	1,109.4	1,110.5699	1,101.231	1,106.26	1,106.26	902,500
6	2018-01-10	1,097.1	1,104.6	1,096.11	1,102.61	1,102.61	1,042,800
7	2018-01-11	1,106.3	1,106.525	1,099.59	1,105.52	1,105.52	978,300
8	2018-01-12	1,102.41	1,124.29	1,101.15	1,122.26	1,122.26	1,720,500
9	2018-01-16	1,132.51	1,139.91	1,117.832	1,121.76	1,121.76	1,575,300

Testing Completed



Predictions

Ticker

Start Date

2023/07/17

End Date

2023/07/17

Price Movements

	Date	Open	High	Low	Close	Adj Close	Volume	% change
1	30/06/2010	1.7193	2.028	1.5533	1.5887	1.5887	257,806,500	-0.0025
2	01/07/2010	1.6667	1.728	1.3513	1.464	1.464	123,282,000	-0.0785
3	02/07/2010	1.5333	1.54	1.2473	1.28	1.28	77,097,000	-0.1257
4	06/07/2010	1.3333	1.3333	1.0553	1.074	1.074	103,003,500	-0.1609
5	07/07/2010	1.0933	1.1087	0.9987	1.0533	1.0533	103,825,500	-0.0192
6	08/07/2010	1.076	1.168	1.038	1.164	1.164	115,671,000	0.1051
7	09/07/2010	1.172	1.1933	1.1033	1.16	1.16	60,759,000	-0.0034
8	12/07/2010	1.1967	1.2047	1.1333	1.1367	1.1367	33,037,500	-0.0201
9	13/07/2010	1.1593	1.2427	1.1267	1.2093	1.2093	40,201,500	0.0639
10	14/07/2010	1.196	1.3433	1.184	1.3227	1.3227	62,928,000	0.0937

Annual Return is `51.23508600666341` %

Standard Deviation is `57.225213159926746` %

Risk Adj Return is `8953.236375629953`

Training and Evaluation

During training, the LSTM model learns to minimize the **mean squared error (MSE)** between the predicted and actual stock prices. The model's performance is evaluated using metrics like **root mean squared error (RMSE)** and **mean absolute error (MAE)**. Cross-validation techniques can enhance the robustness of the model's performance assessment.

Conclusion

In this project, we are predicting closing stock price of any given organization, we developed a web application for predicting close stock price using LMS and LSTM algorithms for prediction. We have applied datasets belonging to Google, Microsoft, Tesla Stocks and achieved above 95% accuracy for these datasets.

Thanks

Sanskriti Sharma
Harshit Rajpal
Vivek Handa