STOCK PRICE PREDICTION

TEAM MEMBERS

MANIKANDAN.B

KARTHICK.M

MATHAN.S

SHIBIN.G

PRASANNA.KM

PHASE 4 SUBMISSION DOCUMENT

PHASE 4: DEVELOPMENT PART 2



INTRODUCTION

In this notebook I will show you how to write a python program that predicts the price of stocks using a machine learning technique called Long Short-Term Memory (LSTM). This program is really simple and I doubt any major profit will be made from this program, but it’s slightly better than guessing! Remember the stock price can be affected by many different things.

NECESSARY STEPS TO FOLLOW:

MAKING TECH LIST

Tech\_list = ['AAPL', 'GOOG', 'MSFT', 'AMZN']

Company\_list = [AAPL,GOOG,MSFT,AMZN]

company\_name = ['AAPL', 'GOOG', 'MSFT', 'AMZN']

for company, comp\_name **in** zip(Company\_list,company\_name):

company["company\_name"] = comp\_name

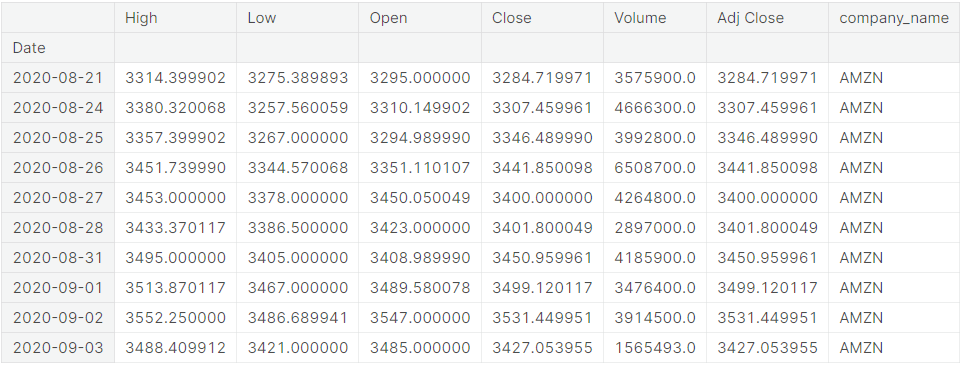
LOAD THE DATASET:

To load a dataset for credit card fraud detection, We can use the Pandas library in Python. Here's how we can load a dataset from a CSV file, which is a common data format:

A COMPANY LIST

df = pd.concat(Company\_list,axis=0)

df.tail(10)



AAPL.info()

Output:

<class 'pandas.core.frame.DataFrame'>

DatetimeIndex: 255 entries, 2019-09-03 to 2020-09-03

Data columns (total 7 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 High 255 non-null float64

1 Low 255 non-null float64

2 Open 255 non-null float64

3 Close 255 non-null float64

4 Volume 255 non-null float64

5 Adj Close 255 non-null float64

6 company\_name 255 non-null object

dtypes: float64(6), object(1)

memory usage: 15.9+ KB

Data processing :

Data processing occurs when data is collected and translated into usable information. Usually performed by a data scientist or team of data scientists, it is important for data processing to be done correctly as not to negatively affect the end product, or data output.

Basic Summary Statistics:

STATISTICAL ANALYSIS AND DATA VISUALIZATION :

Use Pandas to obtain summary statistics of the dataset, which can give a quick overview of the data, including counts, means, standard deviations, and percentiles.

Program:

plt.figure(figsize=(12, 8))

plt.subplots\_adjust(top=1.25, bottom=1.2)

for i, company **in** enumerate(Company\_list, 1):

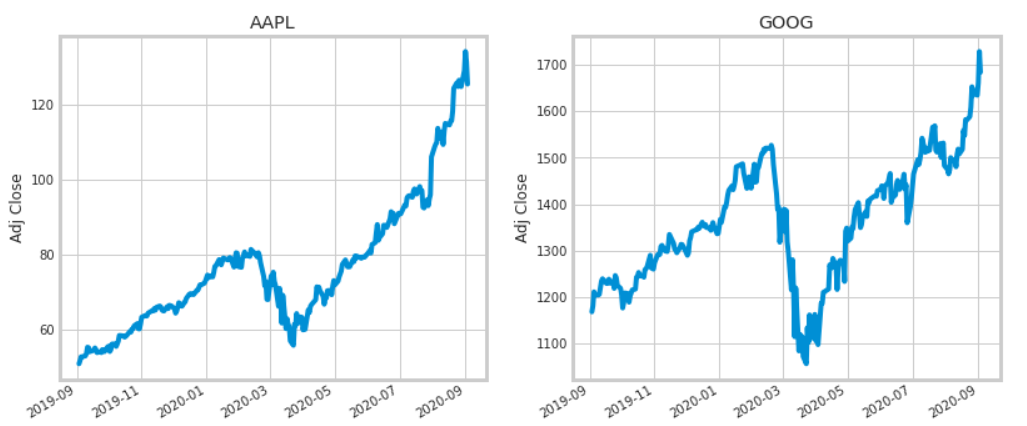
plt.subplot(2, 2,i)

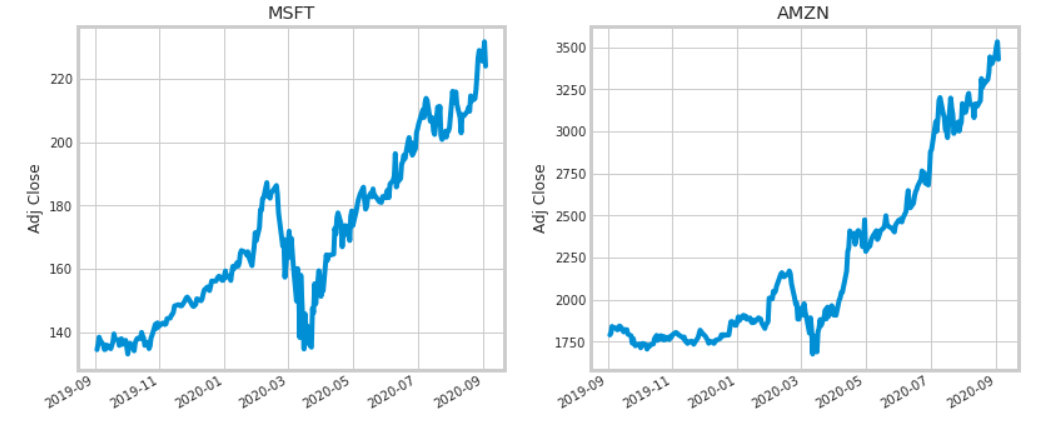
company['Adj Close'].plot()

plt.ylabel('Adj Close')

plt.xlabel(None)

plt.title(f"**{**Tech\_list[i - 1]**}**")





MOVING AVERAGE OF THE VARIUOS STOCKS

There are three important moving averages that can be applied to your charts that will help you trade better. They are the 10 moving average, the 20 moving average and the 50 moving average. The 20 moving average (10MA) is the short-term outlook. The 50 moving average (20MA) is the medium term outlook. The 200 moving average (50MA) is the trend bias.

fig, axes = plt.subplots(nrows=2,ncols=2)

fig.set\_figheight(8)

fig.set\_figwidth(15)

AAPL[['Adj Close', 'MA for 10 days', 'MA for 20 days', 'MA for 50 days']].plot(ax=axes[0,0])

axes[0,0].set\_title('APPLE')

GOOG[['Adj Close', 'MA for 10 days', 'MA for 20 days', 'MA for 50 days']].plot(ax=axes[0,1])

axes[0,1].set\_title('GOOGLE')

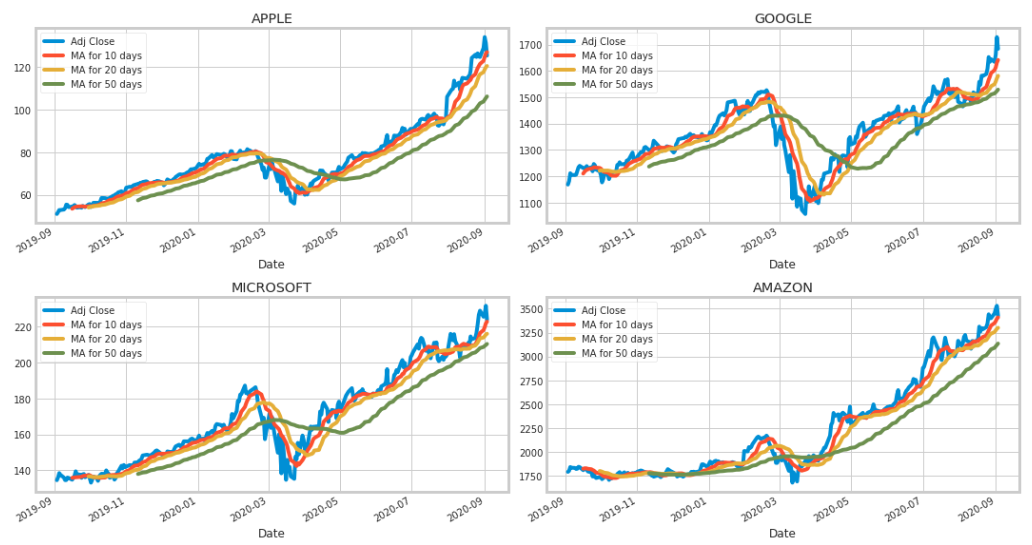
MSFT[['Adj Close', 'MA for 10 days', 'MA for 20 days', 'MA for 50 days']].plot(ax=axes[1,0])

axes[1,0].set\_title('MICROSOFT')

AMZN[['Adj Close', 'MA for 10 days', 'MA for 20 days', 'MA for 50 days']].plot(ax=axes[1,1])

axes[1,1].set\_title('AMAZON')

fig.tight\_layout()



**Different stocks closing prices**

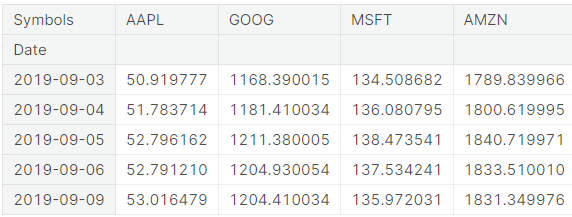
Even in the era of 24-hour trading, there is a closing price for any stock or other security, and it is the final price at which it trades during regular market hours on any given day. The closing price is considered the most accurate valuation of a stock or other security until trading resumes on the next trading day.

The closing price is the raw price, which is just the cash value of the last transacted price before the market closes. The adjusted closing price factors in anything that might affect the stock price after the market closes. A stock's price is typically affected by supply and demand of market participants.

Program:

closing\_df = web.DataReader(Tech\_list, 'yahoo', start, end)['AdjClose']

closing\_df.head()

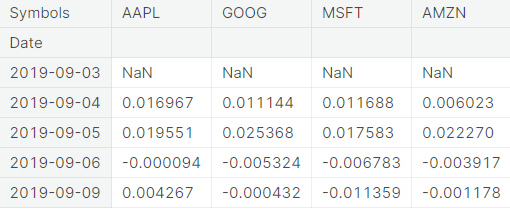


Data frame :

Program:

tech\_rets = closing\_df.pct\_change()

tech\_rets.head()

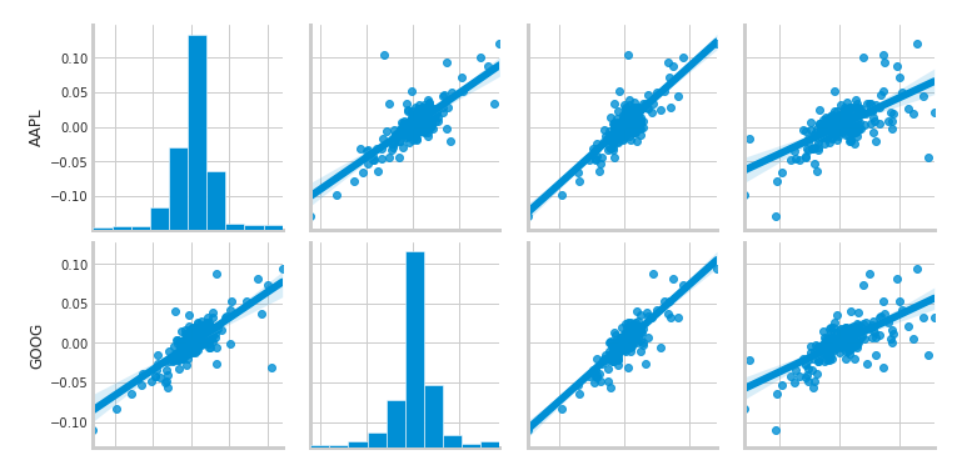


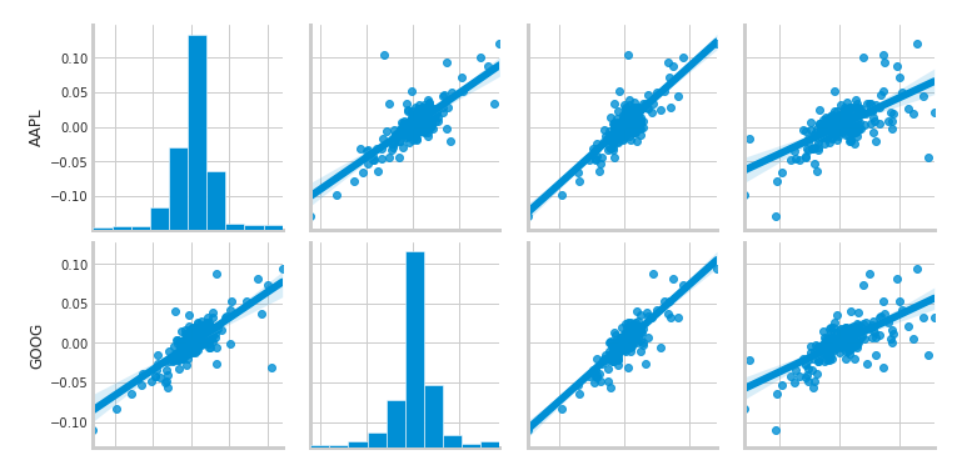
Dataframe for an automatic visual analysis

A DataFrame is essentially a two-dimensional, tabular data structure with rows and columns. Each column can represent a different variable or feature, and each row typically represents a single observation or data point. DataFrames are used to store and work with structured data, making them particularly well-suited for tasks like data cleaning, exploration, and analysis.

Program:

sns.pairplot(tech\_rets, kind='reg')



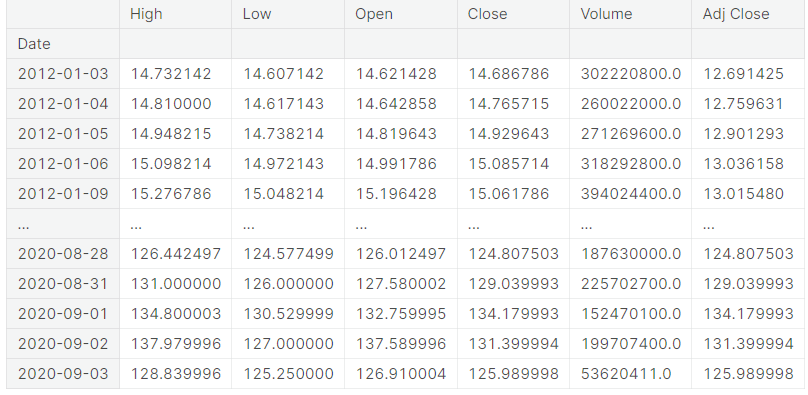


Building Models

**Predicting the closing price stock price of APPLE inc:**

**Program:**

df = web.DataReader('AAPL', data\_source='yahoo', start='2012-01-01', end=datetime.now())



df.shape

(2183, 6)

plt.figure(figsize=(16,8))

plt.title('Close Price History')

plt.plot(df['Close'])

plt.xlabel('Date', fontsize=18)

plt.ylabel('Close Price USD ($)', fontsize=18)

plt.show()

Now scale the data set to be values between 0 and 1 inclusive, I do this because it is generally good practice to scale your data before giving it to the neural network

Program:

train = data[:training\_data\_len]

valid = data[training\_data\_len:]

valid['Predictions'] = predictions

plt.figure(figsize=(16,8))

plt.title('Model')

plt.xlabel('Date', fontsize=18)

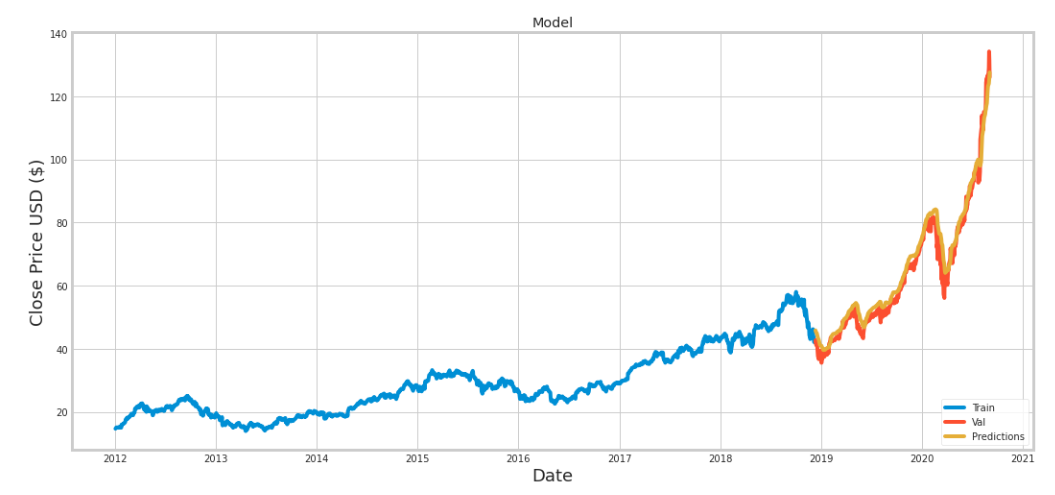
plt.ylabel('Close Price USD ($)', fontsize=18)

plt.plot(train['Close'])

plt.plot(valid[['Close', 'Predictions']])

plt.legend(['Train', 'Val', 'Predictions'], loc='lower right')

plt.show()



Outout

print(valid)

Close Predictions

Date

2018-12-12 42.275002 45.852188

2018-12-13 42.737499 45.508003

2018-12-14 41.369999 45.266945

2018-12-17 40.985001 44.958637

2018-12-18 41.517502 44.606354

... ... ...

2020-08-28 124.807503 123.218452

2020-08-31 129.039993 123.972687

2020-09-01 134.179993 124.918785

2020-09-02 131.399994 126.340744

2020-09-03 125.989998 127.540886

CONCLUSION

Summarize your findings and their implications. Did your model provide valuable insights, or is further research needed? Be cautious about making bold claims, as stock price prediction is inherently uncertain. In conclusion, stock price prediction is a complex task that involves various models and data sources. It's important to be transparent about your methodology, results, and limitations. While predictive models can offer insights, it's crucial to remember that stock markets are influenced by numerous unpredictable factors, and investing decisions should not be based solely on automated predictions.

In conclusion, stock price prediction is a challenging and complex endeavor that involves the analysis of a wide range of financial data and the application of diverse modeling techniques. Our study utilized historical stock prices, trading volumes, and economic indicators, and employed machine learning models to predict future stock prices. While our model showed promise and performed well in training, it is essential to acknowledge its limitations and the inherent unpredictability of financial markets.

Stock prices are influenced by an array of dynamic and often unexpected factors, making precise predictions a formidable task. Furthermore, our analysis assumes that historical patterns will repeat, which may not always hold true. As such, any stock price prediction should be approached with caution and not be the sole basis for investment decisions. Our work underscores the need for continuous research and the exploration of additional data sources, model improvements, and the consideration of unforeseen market events. In the ever-evolving landscape of finance, accurate and reliable stock price predictions remain an ongoing challenge.

This study has several limitations that can provide new directions for future studies. First, we only collected social media text data from one platform. Although we collected as much data as possible from large companies, the investors of other platforms may present different emotions and one website is less representative. We will try to collect more financial social media documents from different platforms in the future. Second, only one stock is selected for prediction in our study.