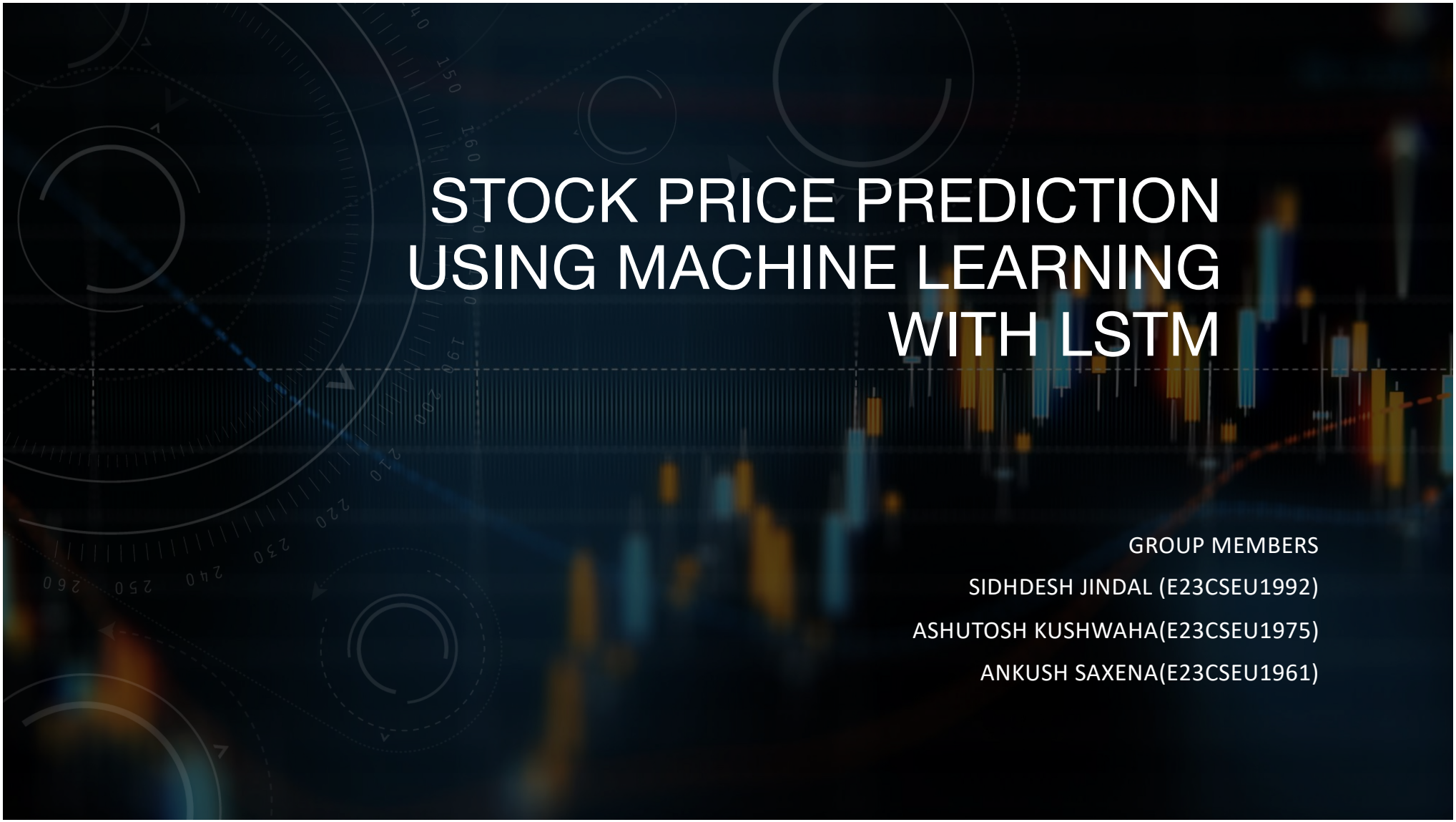


STOCK PRICE PREDICTION USING MACHINE LEARNING WITH LSTM

GROUP MEMBERS

SIDHDESH JINDAL (E23CSEU1992)

ASHUTOSH KUSHWAHA(E23CSEU1975)

ANKUSH SAXENA(E23CSEU1961)

PROJECT OVERVIEW

This project involves building a stock price prediction model for leading tech companies using historical data, with an emphasis on visualizing key trends and patterns in stock prices and volumes. By leveraging machine learning—specifically, a Long Short-Term Memory (LSTM) neural network—the project aims to predict future prices for each stock and analyse factors affecting price fluctuations. The four companies selected for this study are Apple (AAPL), Google (GOOG), Microsoft (MSFT), and Amazon (AMZN).

1.DATA COLLECTION AND PREPARATION

Using Yahoo Finance, we gather historical data on:

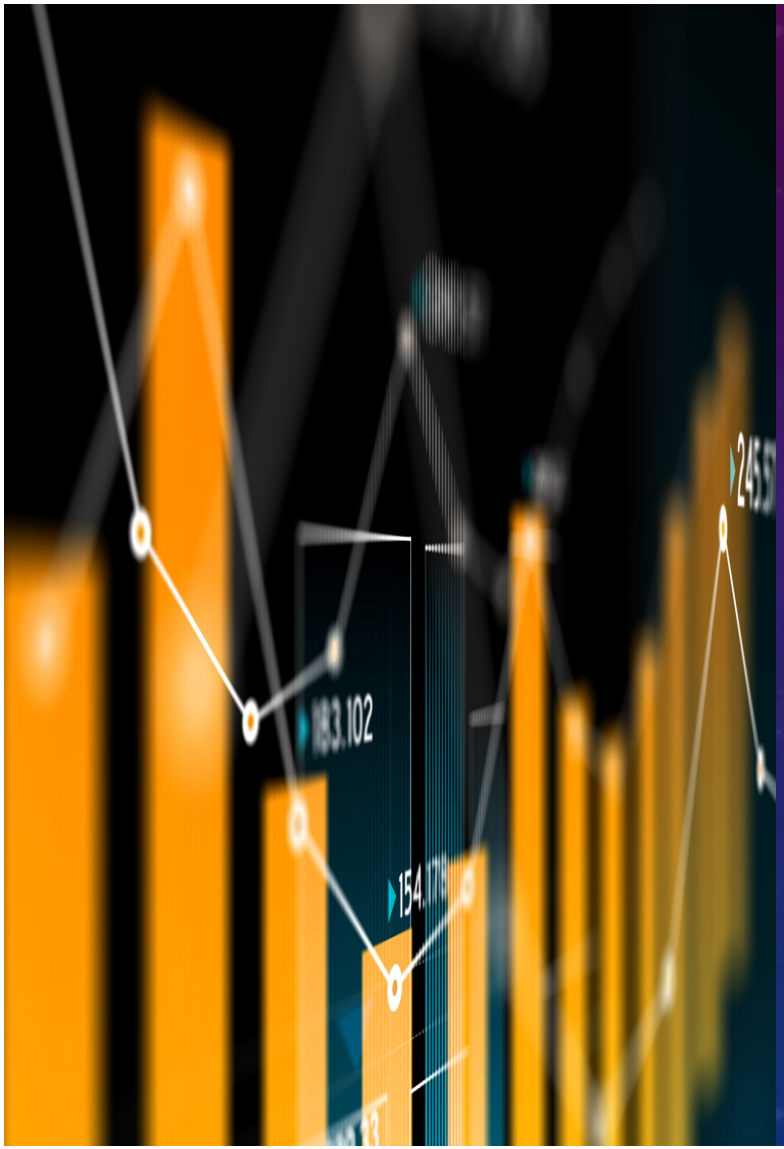
- Adjusted Closing Price: The adjusted closing price accounts for corporate actions like stock splits and dividends.
- Volume: The total number of shares traded daily.
- After loading and inspecting the data, we clean and concatenate it into a single DataFrame. This consolidated data serves as the basis for both the LSTM model and the stock price analysis.

2. STOCK PRICE ANALYSIS

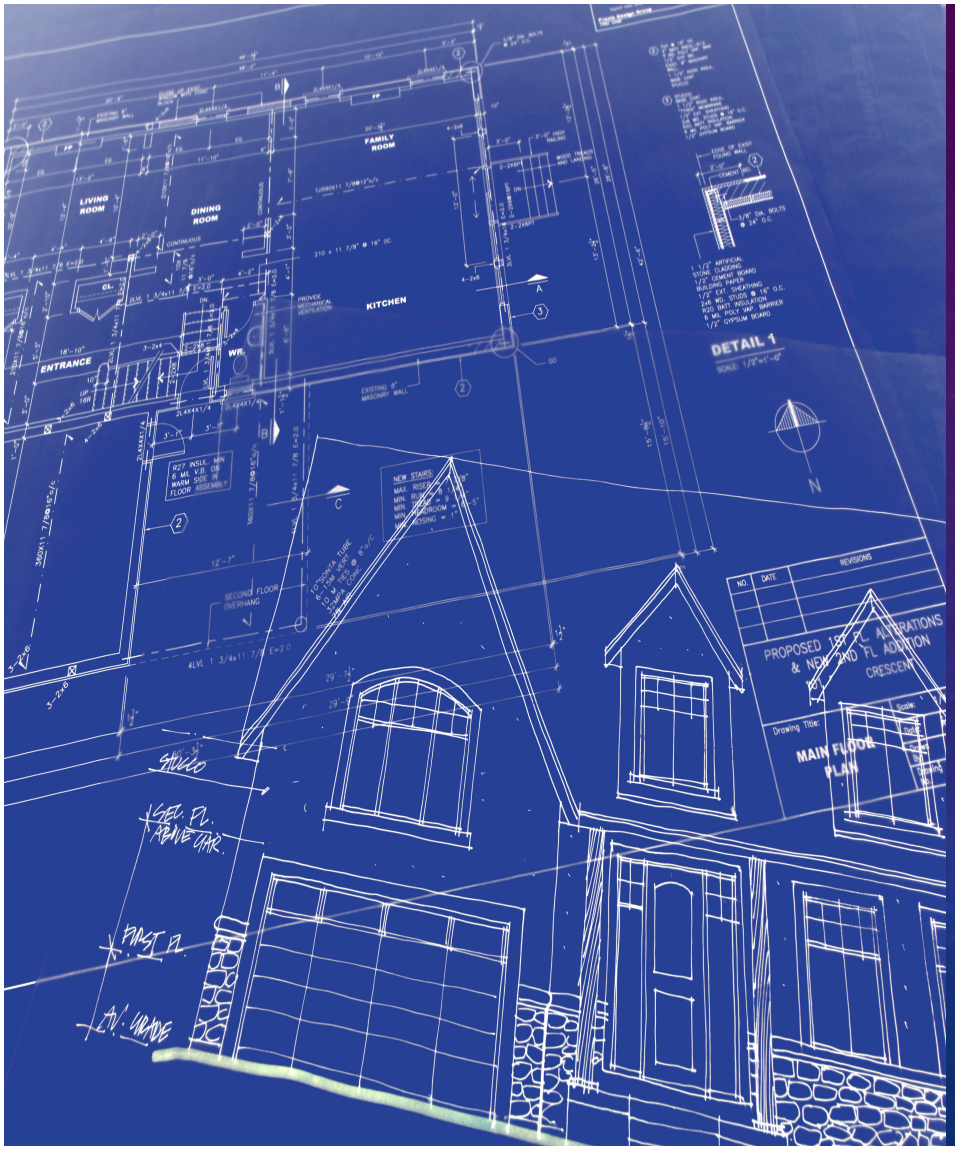
This section covers various key financial metrics that help us better understand stock trends and risk:

2.1 Historical Price Trends
In this section, we plot the Adjusted Closing Prices of each stock over time, allowing us to visually assess price fluctuations and identify trends in each company's stock performance.

2.2 Trading volumes
We plot trading volume to gauge investor activity. High trading volume often reflects significant news, earnings announcements, or other impactful events.

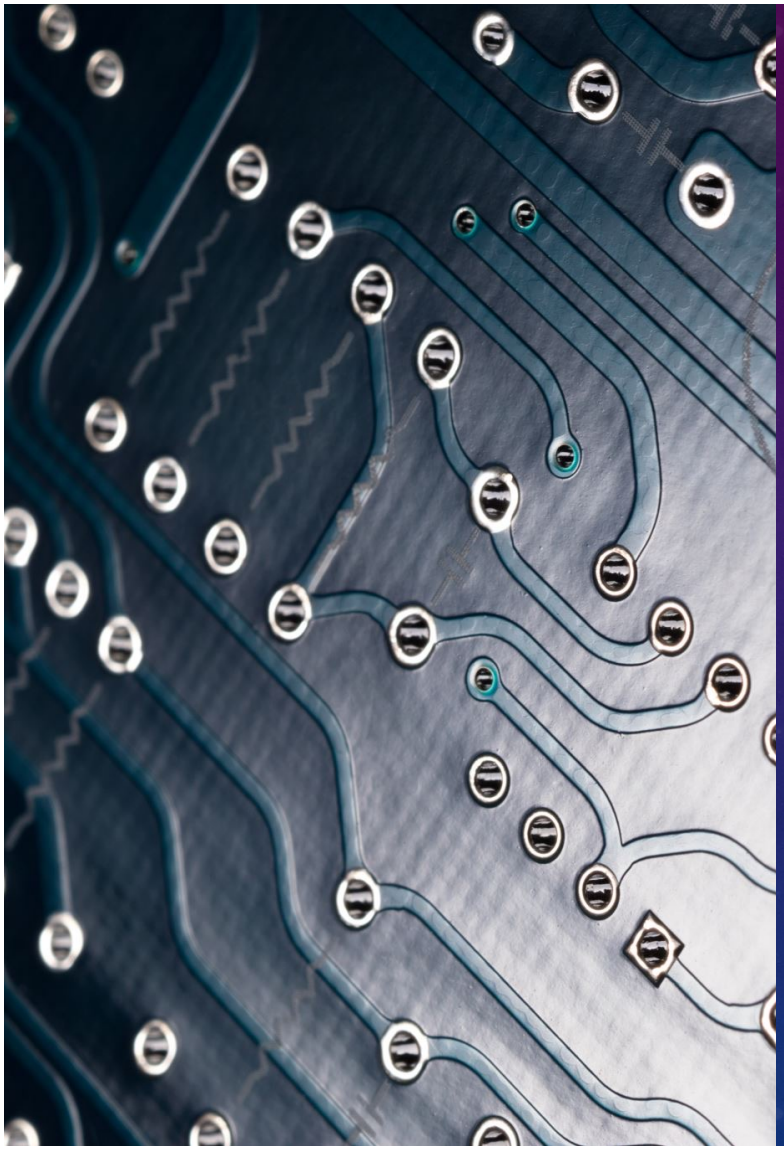


- 2.3 Moving Averages- **Short-Term Moving Average (e.g., 10-day):Used to gauge recent price trends, reflecting short-term momentum in the stock's price.
- Medium-Term Moving Average (e.g., 20-day) and Long-Term Moving Average (e.g., 50-day):** These indicators help identify trends and potential shifts. For instance, when a short-term moving average crosses above a long-term average, it might signal a bullish trend (uptrend) in the stock price.
- 2.4 Daily Returns:
- The daily return calculation indicates the percentage change in the adjusted closing price from one day to the next. Analysing these returns helps evaluate volatility, where high fluctuations signify higher risk and opportunity.
- 2.5 Distribution of Daily Returns:
- Histograms show how returns are distributed, with most values near zero, indicating that small daily price changes are common. Extreme values represent unusually large price changes, often linked to impactful events.
- 2.6 Risk vs. Expected Return:
- - **Expected Return:** The average of daily returns over the selected period.
- - **Risk (Standard Deviation):** This quantifies price volatility. A higher standard deviation suggests more price fluctuation and, therefore, more investment risk.
- The scatter plot of risk vs. expected return allows investors to assess whether a stock's potential return compensates for its level of risk.



METHODOLOGY

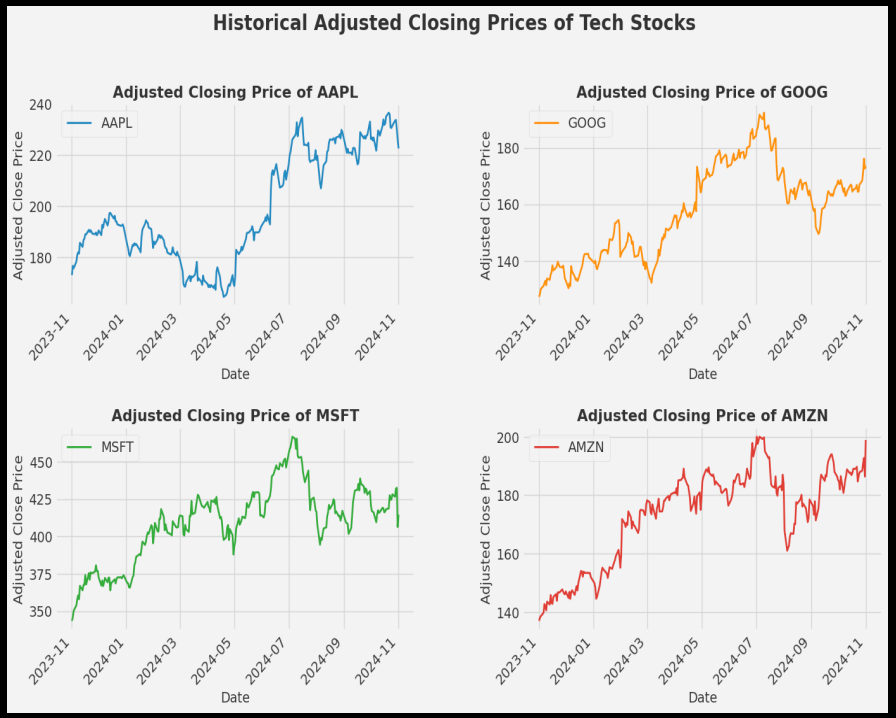
- **Data Cleaning and Normalization**: Ensures the data is free of inconsistencies and standardized for better modelling.
- **AI Model - LSTM**: A deep learning model chosen for its ability to handle sequential data, retaining long-term dependencies essential for time series analysis.
- **Tools and Libraries**: Python is used for scripting, TensorFlow/Keras for deep learning implementation, and Matplotlib for data visualization.

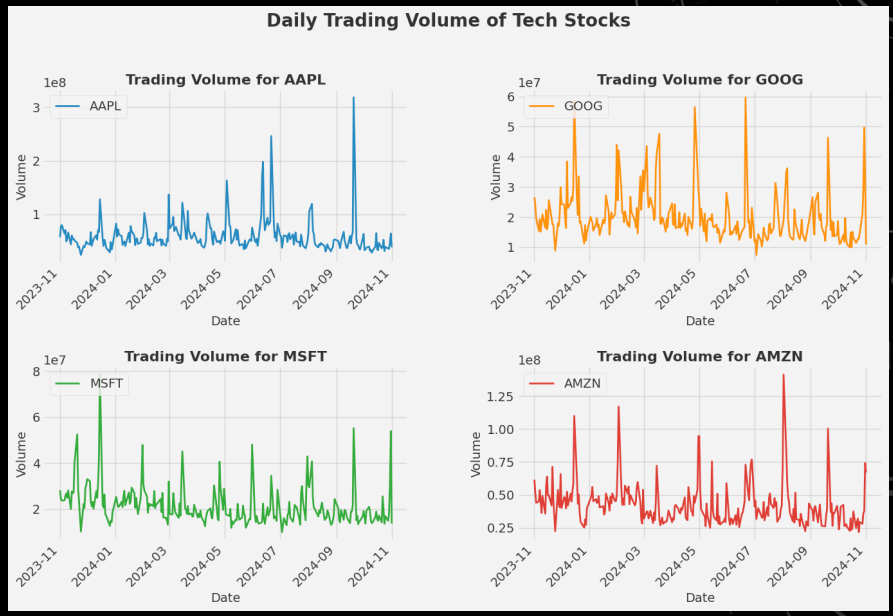


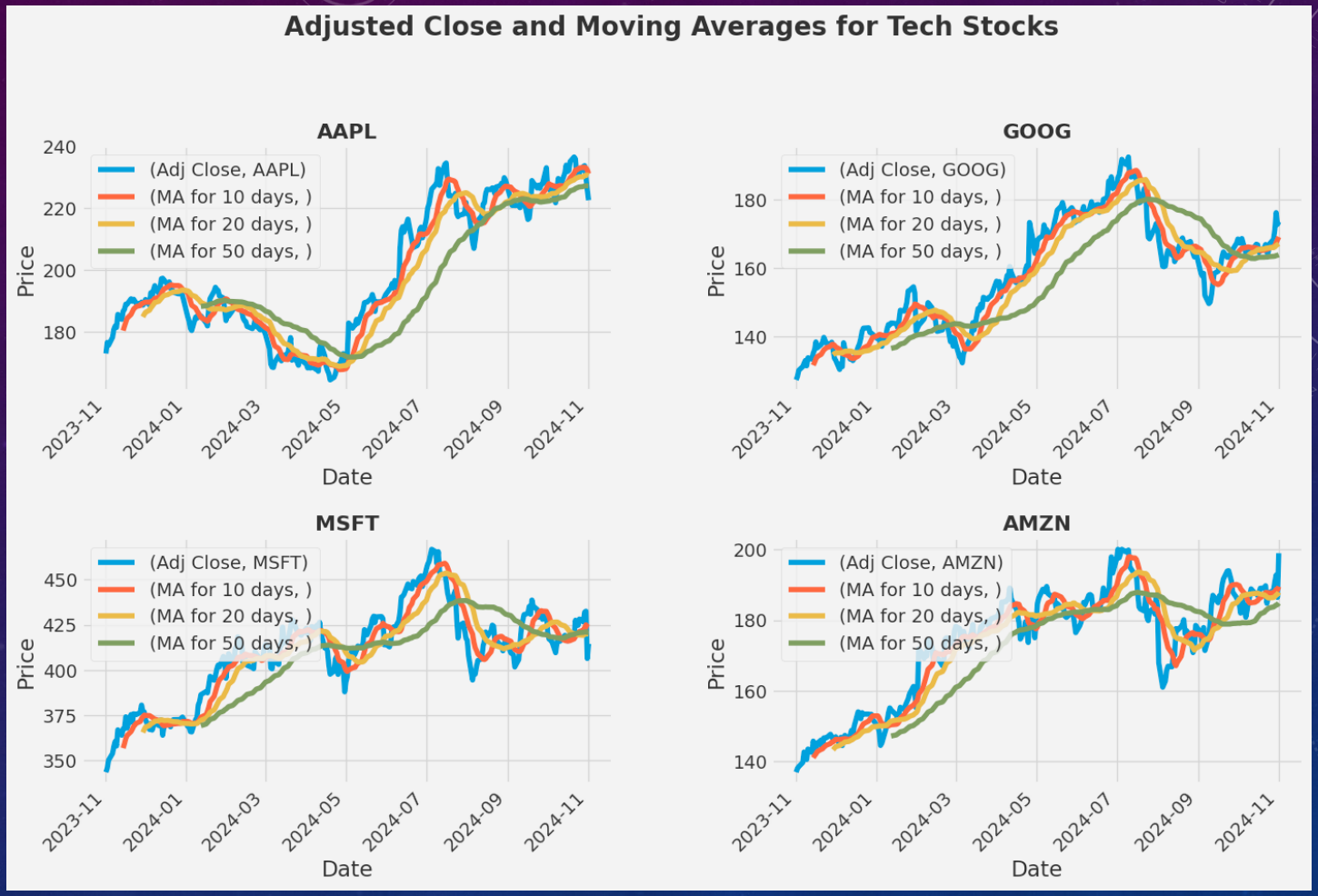
LSTM- LONG SHORT-TERM MEMORY

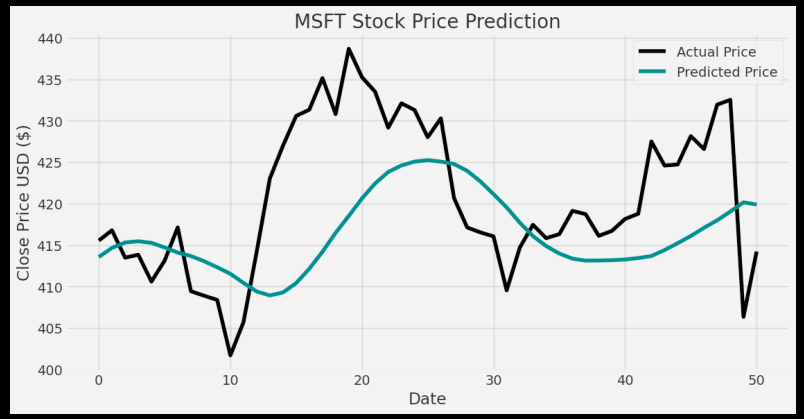
is a type of recurrent neural network (RNN) designed to handle sequential data like time series. It overcomes the limitations of standard RNNs, such as the vanishing gradient problem, by using memory cells and gates (input, forget, and output gates) to retain long-term dependencies. This makes LSTMs ideal for tasks like stock price prediction, natural language processing, and speech recognition.

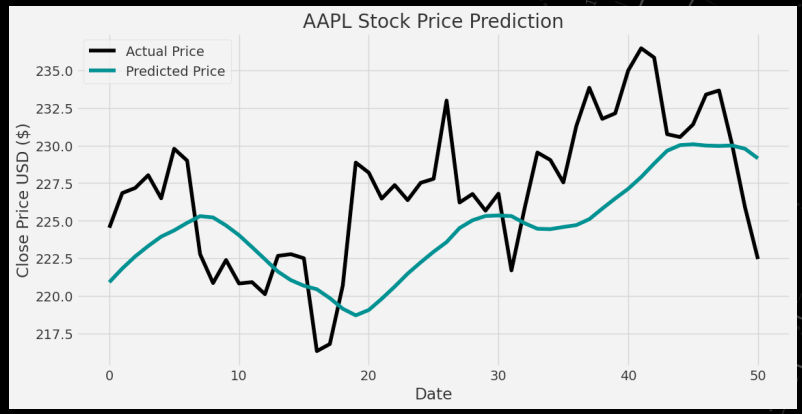
It aims to provide a short-term memory for RNN that can last thousands of timesteps (thus "*long* short-term memory"). The name is made in analogy with long-term memory and short-term memory and their relationship, studied by cognitive psychologists since the early 20th century.

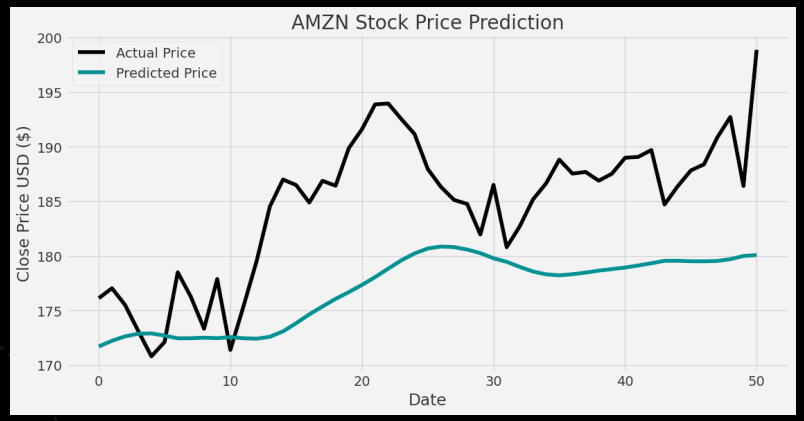

Historical Adjusted Closing Prices of Tech Stocks
Adjusted Closing Price of AAPL
AAPL
Adjusted Close Price
Date
Adjusted Closing Price of GOOG
GOOG
Adjusted Close Price
Date
Adjusted Closing Price of MSFT
MSFT
Adjusted Close Price
Date
Adjusted Closing Price of AMZN
AMZN
Adjusted Close Price
Date

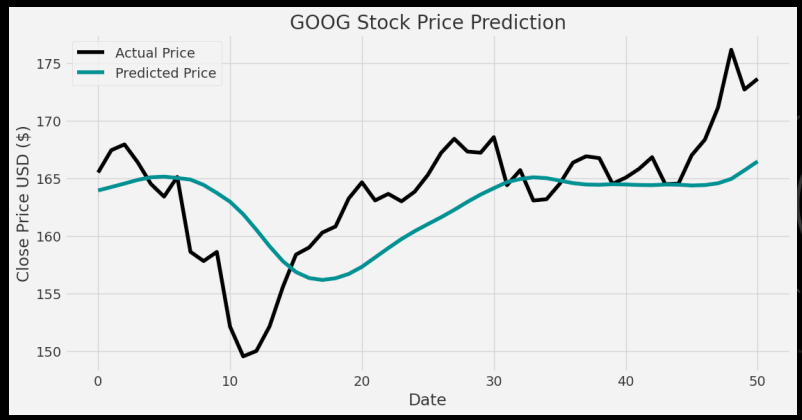

Daily Trading Volume of Tech Stocks
Trading Volume for AAPL
1e8
AAPL
Volume
Date
Trading Volume for GOOG
1e7
GOOG
Volume
Date
Trading Volume for MSFT
1e7
MSFT
Volume
Date
Trading Volume for AMZN
1e8
AMZN
Volume
Date


Adjusted Close and Moving Averages for Tech Stocks
AAPL
(Adj Close, AAPL)
(MA for 10 days,)
(MA for 20 days,)
(MA for 50 days,)
Price
Date
GOOG
(Adj Close, GOOG)
(MA for 10 days,)
(MA for 20 days,)
(MA for 50 days,)
Price
Date
MSFT
(Adj Close, MSFT)
(MA for 10 days,)
(MA for 20 days,)
(MA for 50 days,)
Price
Date
AMZN
(Adj Close, AMZN)
(MA for 10 days,)
(MA for 20 days,)
(MA for 50 days,)
Price
Date


MSFT Stock Price Prediction
Actual Price
Predicted Price
Close Price USD ($)
Date
440
435
430
425
420
415
410
405
400
0
10
20
30
40
50


AAPL Stock Price Prediction
Actual Price
Predicted Price
Close Price USD ($)
Date
235.0
232.5
230.0
227.5
225.0
222.5
220.0
217.5
0
10
20
30
40
50


AMZN Stock Price Prediction
Actual Price
Predicted Price
Close Price USD ($)
Date
200
195
190
185
180
175
170
0
10
20
30
40
50


GOOG Stock Price Prediction
Actual Price
Predicted Price
Close Price USD ($)
Date
175
170
165
160
155
150
0
10
20
30
40
50

CONCLUSION

- The project demonstrates that LSTM effectively predicts stock market trends by capturing long-term dependencies in sequential data. Despite its success, external factors like market volatility can affect accuracy. Key takeaways include the importance of data preprocessing and the advantages of deep learning for time series analysis. While powerful, LSTMs are not foolproof, and incorporating external factors can enhance predictions.