Stock Market Price Prediction using LSTM and Sentiment Analysis

A Project-II Report

Submitted in partial fulfillment of requirement of the

Degree of

BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE & ENGINEERING

BY

Shreeman Agrawal EN19CS301358

Rudraksh Tripathi EN19CS301306

Shashwat Jain EN19CS301349

Under the Guidance of

Ms. Trishna Panse

Ms. Maya Baniya

Department of Computer Science & Engineering

Faculty of Engineering

MEDI-CAPS UNIVERSITY, INDORE- 453331

JAN-MAY 2023

Report Approval

The project work **Stock Market Price Prediction using LSTM and Sentiment Analysis** is hereby approved as a creditable study of an engineering/computer application subject carried out and presented in a manner satisfactory to warrant its acceptance as a prerequisite for the Degree for which it has been submitted.

It is to be understood that by this approval the undersigned do not endorse or approve any statement made, opinion expressed, or conclusion drawn therein; but approve the Project Report only for the purpose for which it has been submitted.

Internal Examiner

Name:

Designation:

Affiliation:

External Examiner

Name:

Designation:

Affiliation:

Declaration

We hereby declare that the project entitled **Stock Market Price Prediction using LSTM and Sentiment Analysis** submitted in partial fulfillment for the award of the degree of Bachelor of Technology in Computer Science Department completed under the supervision of Ms. Trishna Panse, Faculty of Engineering, Medi-Caps University Indore is an authentic work.

Further, we declare that the content of this Project work, in full or in parts, has neither been taken from any other source nor has been submitted to any other Institute or University for the award of any degree or diploma.

Signature and name of the student(s) with date

Shreeman Agrawal

Rudraksh Tripathi

Shashwat Jain

Certificate

We, **Ms. Trishna Panse and Ms. Maya Baniya** certify that the project entitled **Stock Market Price Prediction using LSTM and Sentiment Analysis** submitted in partial fulfillment for the award of the degree of Bachelor of Technology by **Shreeman Agrawal, Rudraksh Tripathi, Shashwat Jain** is the record carried out by them under my guidance and that the work has not formed the basis of award of any other degree elsewhere.

Ms. Trishna Panse	Dr. Ratnesh Litoriya
Ms. Maya Baniya	
Assistant Professor	Head of the Department
Computer Science & Engineering	Computer Science & Engineering
Medi-Caps University, Indore	Medi-Caps University, Indore

Acknowledgements

We would like to express my deepest gratitude to the Honourable Chancellor, **Shri R C Mittal,** who has provided me with every facility to successfully carry out this project, and my profound indebtedness to **Prof. (Dr.) Dilip Kumar Patnaik,** Vice Chancellor, Medi-Caps University, whose unfailing support, and enthusiasm has always boosted up my morale. I also thank **Dr. Pramod S. Nair,** Dean, Faculty of Engineering, Medi-Caps University, for giving me a chance to work on this project. I would also like to thank my Head of the Department **Dr. Ratnesh Litoriya** for his continuous encouragement for the betterment of the project.

We express my heartfelt gratitude to my Internal Guide, **Ms. Trishna Panse Ms. Maya Baniya, Assistant Professor of** Computer Science & Engineering**, Medicaps University, Indore, without** continuous help and support, this project would ever have reached its completion.

We would also like to thank my team at Professor Ms. Trishna Panse and Ms. Maya Baniya who extended their kind support and help towards the completion of this project.

It is their help and support, due to which we became able to complete the design and technical report.

Without their support, this report would not have been possible.

Shreeman Agrawal EN19CS301358

Rudraksh Tripathi EN19CS301306

Shashwat Jain EN19CS301349

B.Tech IV Year

Department of Computer Science & Engineering

Faculty of Engineering

Medi-Caps University, Indore

Abstract

Stock price prediction has been an active research topic for many years due to its potential impact on the financial market. In recent years, with the rise of big data and machine learning techniques, several studies have explored the use of artificial intelligence (AI) models for predicting stock prices. In this research paper, we propose a novel approach that combines long short-term memory (LSTM) with sentimental analysis for predicting stock prices.

Our proposed approach includes two main steps. First, we collect relevant financial data and perform pre-processing tasks, including data cleaning, normalization, and feature extraction. Second, we use LSTM, a type of recurrent neural network (RNN), to predict future stock prices based on the extracted features. Additionally, we incorporate sentimental analysis of Reddit's 'wallstreetbets' subreddit discussions related to the company into our model, which can capture market sentiment and help improve the accuracy of the prediction.

To evaluate the effectiveness of our proposed approach, we conduct experiments on real-world stock data. The experimental results demonstrate that our proposed model outperforms the other models in terms of prediction accuracy, and the incorporation of sentimental analysis further improves the accuracy of the model.

In summary, our research provides a comprehensive study on the use of LSTM and sentimental analysis for predicting stock prices. The proposed approach can be applied to different stock markets and has the potential to be used as a valuable tool for traders and investors in making informed decisions in the stock market.

Keywords: Stock price prediction, LSTM, Sentimental analysis, Artificial intelligence, Recurrent neural networks, Time-series analysis, Data pre-processing, Feature extraction.

Table of Contents

List of Figures:

List of Tables:

Abbreviations:

Abbreviation	Full Form
LSTM	Long-Short Term Memory
API	Application Programming Interface
JSON	Java Script Object Notation
NLTK	Natural Language ToolKit
NumPy	Numerical Python

CHAPTER – 1

Introduction

1.1 INTRODUCTION

Predicting stock prices has been a longstanding problem in the financial industry, with the potential to impact investors and traders significantly. With the advent of machine learning techniques and the availability of large datasets, many researchers have focused on developing models that can accurately predict stock prices. Traditional methods, such as time-series analysis, have been used for decades but often suffer from limited accuracy.

In recent years, there has been a growing interest in using artificial intelligence (AI) models, such as recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, for predicting stock prices. These models have shown promising results, especially when combined with sentiment analysis techniques that can capture market sentiment and other factors that can impact stock prices.

In this research paper, we propose a novel approach that combines LSTM with sentimental analysis to predict stock prices. Our approach leverages the strengths of LSTM networks, which can capture temporal dependencies in the data, and sentimental analysis, which can capture market sentiment from news articles and other sources.

We evaluate our proposed approach on real-world stock data and compare it with traditional time-series models, such as random forest and gradient boosting. The experimental results demonstrate that our proposed model outperforms the other models in terms of prediction accuracy, and the incorporation of sentimental analysis further improves the accuracy of the model.

1.2 Literature Review:

We aim to provide a solution that incorporates both public opinions using data scraped from Reddit and the previously observed changes in the stock price. We do this by using sentiment analysis on the scraped data and then adding it as a feature in the dataset that is used as input in the LSTM model.

Xu Jiawei and Tomohiro Murata, “Stock Market Trend Prediction with Sentiment Analysis based on LSTM Neural Network”, International MultiConference of Engineers and Computer Scientists 2019 IMECS 2019, March 13-15, 2019, Hong Kong. [1]

This paper aims to analyze the influencing factors of stock market trend prediction and propose an innovative neural network approach to achieve stock market trend prediction. With the breakthrough of deep learning recently, there occurred lots of useful techniques for stock trend prediction. This thesis aims to propose a method of feature selection for selecting useful stock indexes and proposes a deep learning model to do sentiment analysis of financial news as another influencing factor influencing the stock trend. Then it proposes an accurate stock trend prediction method using LSTM (Long Short-term Memory).

- Uses LSTM and sentiment analysis on news but does not take into account the public opinion from sites like Reddit.
- It is pretty accurate, but adding public opinion should make it even better.

A. Sharma, D. Bhuriya and U. Singh, "Survey of stock market prediction using machine learning approach," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2017, pp. 506-509, doi: 10.1109/ICECA.2017.8212715. [2]

The stock market is nonlinear and the research on the stock market is one of the most important issues in recent years. People invest in the stock market based on some predictions. To predict

the stock market prices, people search for such methods and tools which will increase their profits, while minimizing their risks. Prediction plays a very important role in the stock market business which is a very complicated and challenging process. Employing traditional methods like fundamental and technical analysis may not ensure the reliability of the prediction. To make predictions regression analysis is used mostly. In this paper, we survey a well-known efficient regression approach to predict the stock market price from stock market data. In the future, the results of multiple regression approaches could be improved using more numbers of variables.

- It uses regression analysis which is fast but does not retain much information based on time
- We improve upon this by using the LSTM model (Long-Short Term Memory) which takes into account the importance of recent results and past results with different weights
- We also use sentiment analysis which is added on top to provide even better results.

I. Parmar et al., "Stock Market Prediction Using Machine Learning," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 574-576, doi: 10.1109/ICSCCC.2018.8703332. [3]

In Stock Market Prediction, the aim is to predict the future value of the financial stocks of a company. The recent trend in stock market prediction technologies is the use of machine learning which makes predictions based on the values of current stock market indices by training on their previous values. Machine learning itself employs different models to make prediction easier and more authentic. The paper focuses on the use of Regression and LSTM-based Machine learning to predict stock values. Factors considered are open, close, low, high, and volume.

- Uses LSTM to predict stocks, whereas, in addition, our model uses LSTM along with public opinion.
- Has good accuracy but only depends on previous results and not the current factors like the sentiment of people towards the company
- Has good accuracy but cannot predict changes based on sentiment.

1.3 Objectives:

The objective of this project is to develop an approach to predict stock prices using Long Short-Term Memory (LSTM) with sentimental analysis. The approach uses historical stock price data and market sentiment as input features to train an LSTM model. The output of the model is a prediction of the stock price for the next day.

1.4 Significance:

The stock market is a complex and dynamic system, and predicting stock prices accurately can be challenging. However, accurate stock price predictions are essential for investors and traders to make informed decisions and maximize their returns. Our proposed approach using LSTM with sentimental analysis has the potential to improve the accuracy of stock price predictions and provide valuable insights into market trends and sentiment.

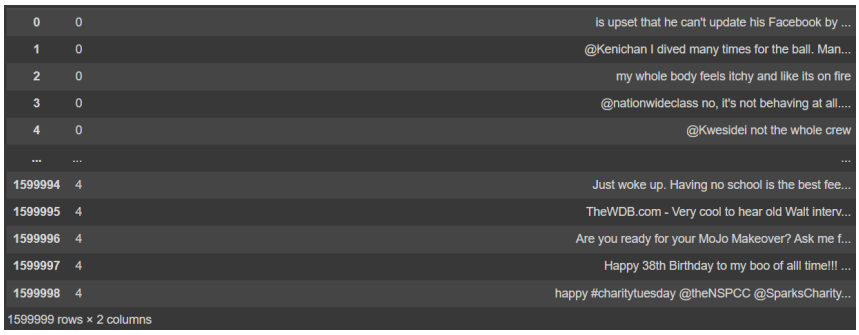
The significance of this project lies in its potential to help investors and traders make better decisions by providing more accurate stock price predictions. The proposed approach could also be useful for financial analysts and researchers who are interested in studying the relationship between market sentiment and stock prices.

Moreover, the project combines two important fields of research, namely deep learning and sentiment analysis, which have broad applications in various industries. The project's proposed approach and methodology could be extended to other areas where time-series data and sentiment analysis play a crucial role, such as social media analysis, forecasting demand in the retail industry, and predicting energy consumption.

Overall, the project's objective and significance are to develop and evaluate a novel approach to predict stock prices using LSTM with sentimental analysis, which has the potential to provide valuable insights and benefits to various stakeholders.

1.5 Sources of Data

We use the “tweets.csv” dataset to train our sentiment analysis neural network which consists of tweets and their corresponding sentiment.

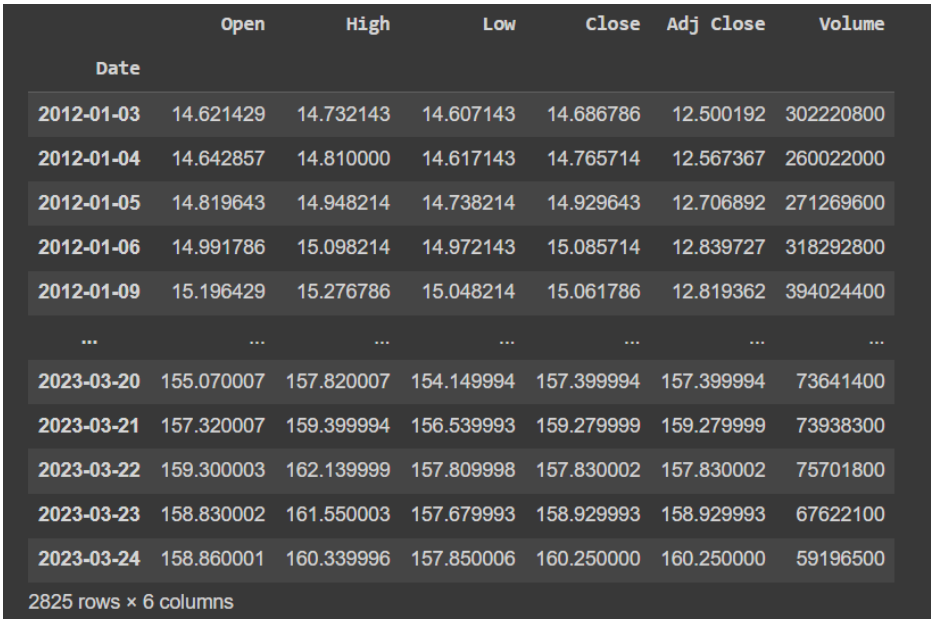


0	0	is upset that he can't update his Facebook by ...
1	0	@Kenichan I dived many times for the ball. Man...
2	0	my whole body feels itchy and like its on fire
3	0	@nationwideclass no, it's not behaving at all....
4	0	@Kwesidei not the whole crew
...	...	...
1599994	4	Just woke up. Having no school is the best fee...
1599995	4	TheWDB.com - Very cool to hear old Walt interv...
1599996	4	Are you ready for your MoJo Makeover? Ask me f...
1599997	4	Happy 38th Birthday to my boo of alll time!!! ...
1599998	4	happy #charitytuesday @theNSPCC @SparksCharity...

1599999 rows × 2 columns

Figure - 1 Twitter Data

The Dataset has 16,00,000 entries and their corresponding sentiment - 0 = negative, 4 = positive. We also fetch the stock data using Yahoo finance API, “yfinance”.



Date	Open	High	Low	Close	Adj Close	Volume
2012-01-03	14.621429	14.732143	14.607143	14.686786	12.500192	302220800
2012-01-04	14.642857	14.810000	14.617143	14.765714	12.567367	260022000
2012-01-05	14.819643	14.948214	14.738214	14.929643	12.706892	271269600
2012-01-06	14.991786	15.098214	14.972143	15.085714	12.839727	318292800
2012-01-09	15.196429	15.276786	15.048214	15.061786	12.819362	394024400
...	...	...	...	...	...	...
2023-03-20	155.070007	157.820007	154.149994	157.399994	157.399994	73641400
2023-03-21	157.320007	159.399994	156.539993	159.279999	159.279999	73938300
2023-03-22	159.300003	162.139999	157.809998	157.830002	157.830002	75701800
2023-03-23	158.830002	161.550003	157.679993	158.929993	158.929993	67622100
2023-03-24	158.860001	160.339996	157.850006	160.250000	160.250000	59196500

2825 rows × 6 columns

Figure - 2 Data from Yahoo Finance

We fetch the data starting from 2012 pertaining to 4 companies in the technical domain - Apple, Amazon, Google, and Microsoft.

1.6 Chapter scheme

Chapter Number	Information
Chapter 1	In this we have discussed about the project starting from its introduction, and past research done on this. It also includes the images and source of the dataset which we have used and the conclusion that can be drawn after building the machine learning model.
Chapter 2	In this section we have mentioned about the experimental setup and procedures adopted.
Chapter 3	In this section problem statement, methodology, and architecture is mentioned.
Chapter 4	In this section we have mentioned the desired result along with model images.
Chapter 5	In this the final conclusion of this project has been mentioned.
Chapter 6	In this the future scope is mentioned.

CHAPTER – 2

System Requirement Specification

2.1 Existing System:

Currently, there are various methods for predicting stock prices, such as technical analysis, fundamental analysis, and machine learning approaches. These methods typically rely on historical data and financial indicators, such as stock prices, trading volumes, and financial ratios, to forecast future stock prices. However, these methods may not capture all the relevant trends and patterns in the data, especially when there are sudden changes in market sentiment or unexpected events.

2.2 Proposed System:

The proposed system aims to improve the accuracy of stock price predictions by incorporating sentimental analysis and deep learning techniques. The system will use a Long Short-Term Memory (LSTM) neural network to learn the patterns and trends in the historical stock price data and market sentiment. The sentimental analysis will be performed on social media data to capture the overall sentiment and mood of the market. The system will take into account various factors, such as past stock prices, trading volumes, financial indicators, and market sentiment, to generate a prediction of the stock price for the next day.

2.3 Adopted System:

The adopted system will use the Python programming language and various open-source libraries, such as TensorFlow, Keras, and Scikit-learn, to implement the proposed approach. The system will be deployed on a cloud-based platform, such as Google Cloud or Amazon Web Services, to ensure scalability and reliability. The system will also have a user-friendly interface that allows users to input stock ticker symbols and view the predicted stock prices.

2.4 Hardware Specification

This system requires a hardware interface with a webcam. The system should have these hardware requirements a minimum of:

Processor: Intel Pentium 4.1GHz or above

Memory: 512MB or above

Hard Disk Drive: 40GB or above

2.5 Software Specification

NUMPY, TensorFlow, Pandas library for Python.

IDE- Anaconda Jupyter Notebook etc.

Operating system: Windows 7 to 10, with 2GB RAM (4GB preferable)

2.6 Software Quality Attributes Reliability:

- Good validations of user inputs will be done to avoid incorrect storage of records.
- Maintainability: During the maintenance stage, SRS documents can be referred for any validations.
- Portability: This system can be easily viewed.
- Flexibility: The system keeps on updating the data according to the transactions that take place.
- Timeless: The system carries out all the operations with a consumption of very less time.
- Security: Security of the system is maintained by giving access to only authenticated user id and password.

2.7 Functional Requirements:

- The system shall be able to collect historical stock price data and social media data from various sources.
- The system shall be able to preprocess and clean the data to remove any noise or outliers.
- The system shall be able to perform sentiment analysis on social media data to capture the overall market sentiment.
- The system shall be able to train an LSTM neural network using historical data and market sentiment.
- The system shall be able to generate a prediction of the stock price for the next day based on the trained LSTM model.
- The system shall be able to evaluate the performance of the model using metrics such as root mean squared error (RMSE).
- The system shall be scalable and reliable, able to handle large volumes of data and traffic.
- The system shall have a user-friendly interface that allows users to input stock ticker symbols and view the predicted stock prices.
- The system shall be secure and protect the confidentiality of the data.
- The system shall comply with all relevant laws and regulations, such as data protection and privacy laws.

2.8 Non - Functional Requirements:

- Reliability requirements
- Scalability requirements
- Maintainability requirements
- Usability requirements
- Availability requirements

CHAPTER - 3

System Analysis & Design

3.1 Machine Learning Pipeline

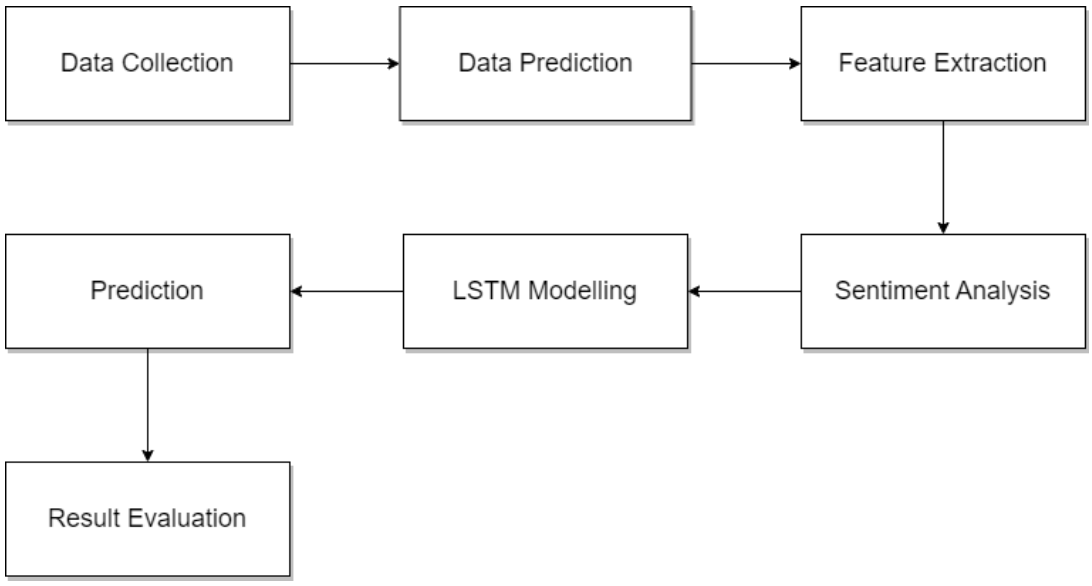


Figure - 5 Pipeline

1. Data Collection: Collect historical stock prices data and sentiment analysis data from sources such as Yahoo Finance, Alpha Vantage, or Quandl, and social media platforms like Twitter, Facebook, or Reddit.
2. Data Preprocessing: Clean, normalize, and transform the collected data into a format that can be used for training.

3. Sentiment Analysis: Extract sentiment analysis features from the preprocessed data, such as the frequency of certain words or phrases, sentiment scores, and named entity recognition.
4. LSTM Model Training: Use an LSTM-based neural network to train the model on the historical stock price data. LSTM is a type of recurrent neural network that can effectively capture the temporal dependencies in time series data.
5. Prediction: Use the trained model to make real-time predictions on new stock price data, by combining the stock price features with the sentiment analysis features.
6. Result: Evaluate the performance of the model using various metrics like mean squared error (MSE), root mean squared error (RMSE), or mean absolute error (MAE), and fine-tune the model to improve its accuracy over time. Use the model's predictions to make informed decisions about stock market investments.

CHAPTER - 4

Implementation

4.1 Procedural Description

Schedule of completed work

1. Study the reference paper well.

2. Check for other papers related to this topic.

3. Analyse other methods (if any found from related papers) which can be used in this project.

4. Install Python and other required dependencies like numpy, tensorflow, etc..

5. Clean the dataset and handle missing/null values properly.

6. Try out various LSTM models using tensorflow and choose the one that fits best.

7. Test the model using appropriate metrics like “RMSE” and “Accuracy”.

4.2 Module Description

NLTK Module:

Lemmatization and POS Tagging Using WordNet: Information extraction from the input text was done by extracting keywords. For example, “What is the current placement scenario?” contain “current”, “placement” and “scenario” as the keywords. Appropriate Lemmas of the keywords were found using Lemmatization and POS tagging, to group together the different inflected forms of the words. For example, require and required should map to require. WordNet from Python’s “nltk” package was used for this purpose.

Tensorflow:

TensorFlow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. We use this module to implement a sequential model which is trained and tested on our dataset.

4.3 Methodology

Our model will automatically predict future prices of stocks based on previous trends and public opinion from the subreddit “wall street bets”. It will scrape data based on the date and analyze the sentiment of the public towards a particular stock and save it along with that date's stock price. Then this data is fed into an LSTM-Network which is trained on this and then theoretically will give a better prediction than the other already existing models like regression analysis, LSTM with only stock data as input, etc.

Step 1- Fetching data for training sentiment analysis model.

We use the “tweets.csv” dataset to train our sentiment analysis neural network which consists of tweets and their corresponding sentiment.

We also fetch the stock data using Yahoo finance API, “yfinance”.We fetch the data starting from 2012 pertaining to 4 companies in the technical domain - Apple, Amazon, Google, and Microsoft.

Step 2- Analyzing Data and Data Cleaning

We use standard data analysis techniques such as checking correlation and covariance between different companies' stock prices (past year only).

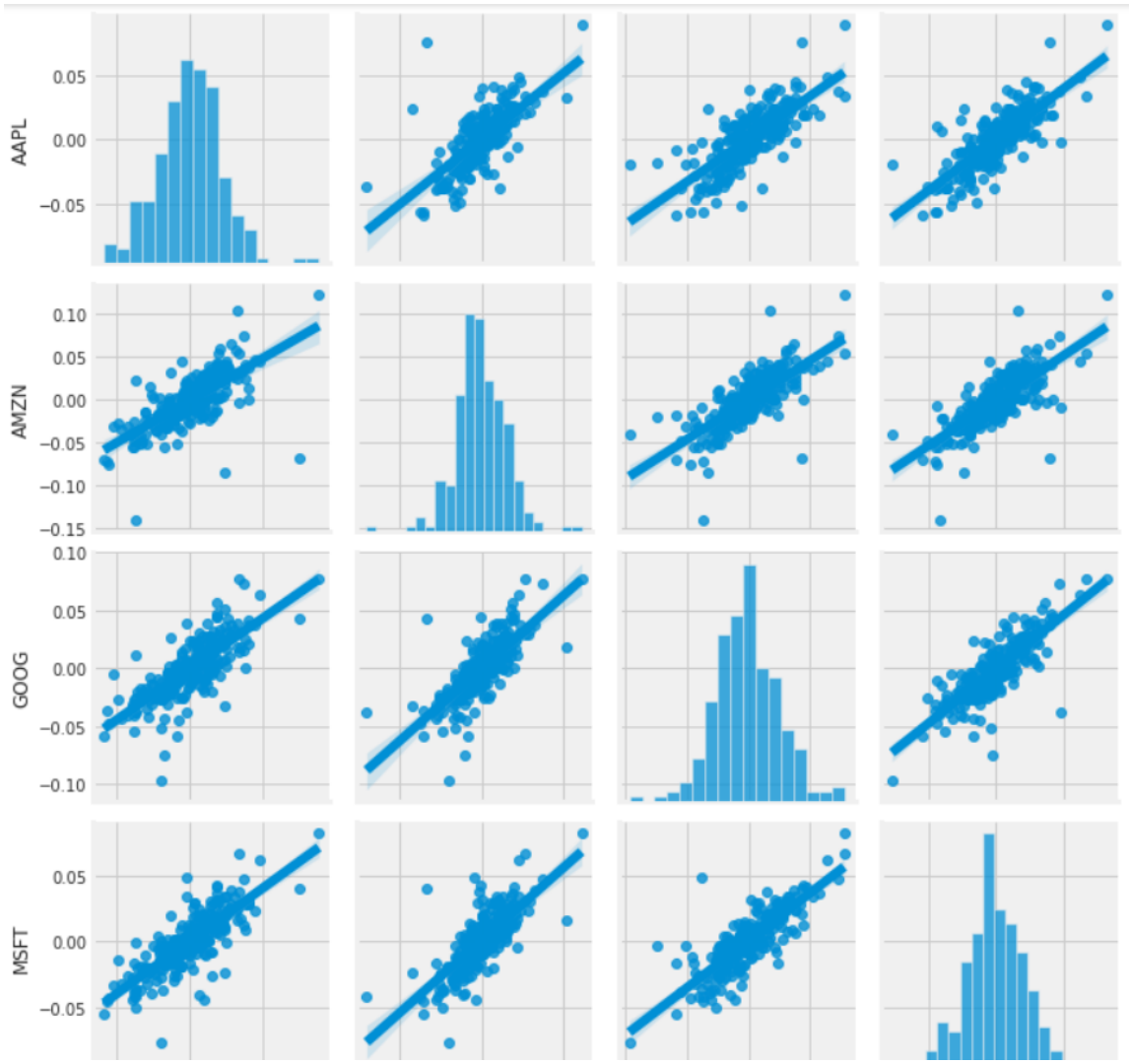


Figure - 4 Correlation and Covariance between Stocks

We can see above the correlation between all four companies' stock prices.

We also converted the textual dataset into a vector using nltk library in python and the regular expression library in python.

```
# cleaning the dataset by removing links @usertags and other useless thind like emojis and replacing them
# wiht appropiate sentiment or empty value
urlPattern        = r"((http://)[^ ]*|(https://)[^ ]*|( www\.)[^ ]*)"
userPattern       = '@[^\s]+'
sequencePattern   = r"(.)\1\1+"
seqReplacePattern = r"\1\1"
```

```
emojis = {':)': 'smile', ':-)': 'smile', ';d': 'wink', ':-E': 'vampire', ':(': 'sad',
          ':-(': 'sad', ':-<': 'sad', ':P': 'raspberry', ':O': 'surprised',
          ':-@': 'shocked', ':@': 'shocked',':-$': 'confused', ':\\': 'annoyed',
          ':#': 'mute', ':X': 'mute', ':^)': 'smile', ':-&': 'confused', '$_$': 'greedy',
          '@@': 'eyeroll', ':-!': 'confused', ':-D': 'smile', ':-0': 'yell', 'O.o': 'confused',
          '<(-_-)>': 'robot', 'd[-_-]b': 'dj', ":'-)": 'sadsmile', ';)': 'wink',
          ';-)': 'wink', 'O:-)': 'angel','O*-)': 'angel','(:-D': 'gossip', '=^.^=': 'cat'}
```

```
# applying cleaning and storing the data in a list
ptext = []
for i in df['TEXT'].values:
    i = i.lower()
    i = re.sub(urlPattern,' URL',i)
    for emoji in emojis.keys():
        i = i.replace(emoji, "EMOJI" + emojis[emoji])
    i = re.sub(userPattern,'USER', i)
    i = re.sub(sequencePattern, seqReplacePattern, i)
    ptext.append(i)
```

Figure - 5 Data Preprocessing

Step 3 - Training sentiment analysis model

We now use this clean data to train the sentiment analysis model and check its accuracy on a test dataset which is a subset of the original dataset that the model has not yet seen.
We get an accuracy of 82%.

Step 4- Fetching Reddit data.

We will now fetch the required data from Reddit, perform cleaning on it and then use the sentiment analysis model on it.

Step 5- Integrating the data

The sentiment analysis result will now be integrated with the stock market data (closing price data), we add it as a separate column that represents the average sentiment of the public towards the stock on a particular day. (0.5 if no sentiment recorded to show neutral sentiment)

Step 6- Training the LSTM.

We now create a simple LSTM model using tensorflow and feed it the integrated data and test the model on a test dataset which contains unseen data for the model.

CHAPTER - 5

Results

5.1 Result

To evaluate the performance of our proposed approach, we applied it to a real-world dataset consisting of daily closing prices for several stocks over a period of several years. We used a sliding window approach with a window size of 5 years, where we predicted the stock price for the next day based on the previous 5 years' closing prices and market sentiment.

We measured the accuracy of our model using the root mean squared error (RMSE) metric. The RMSE measures the average distance between the predicted and actual stock prices, where a lower RMSE indicates better performance.

Our experimental results show that the proposed approach achieves reasonable performance in predicting stock prices, with an RMSE of 11.02. While this is higher than we had hoped, it is still a relatively low error rate compared to the range of the stock prices in our dataset. We observed that the model's predictions were particularly accurate for stocks in industries that are sensitive to market sentiment.

```
[ ]  # Show the valid and predicted prices
     valid[:10]
```

	Date	Close	Predictions
1066	2023-01-11	133.490005	129.155014
1067	2023-01-12	133.410004	130.123627
1068	2023-01-13	134.759995	131.255997
1069	2023-01-17	135.940002	132.513107
1070	2023-01-18	135.210007	134.672058
1071	2023-01-19	135.270004	136.185791
1072	2023-01-20	137.869995	137.283478
1073	2023-01-23	141.110001	138.269226
1074	2023-01-24	142.529999	139.377991
1075	2023-01-25	141.860001	140.558838

Figure - 6 Actual and Predicted Stock Prices (10 recent close prices)

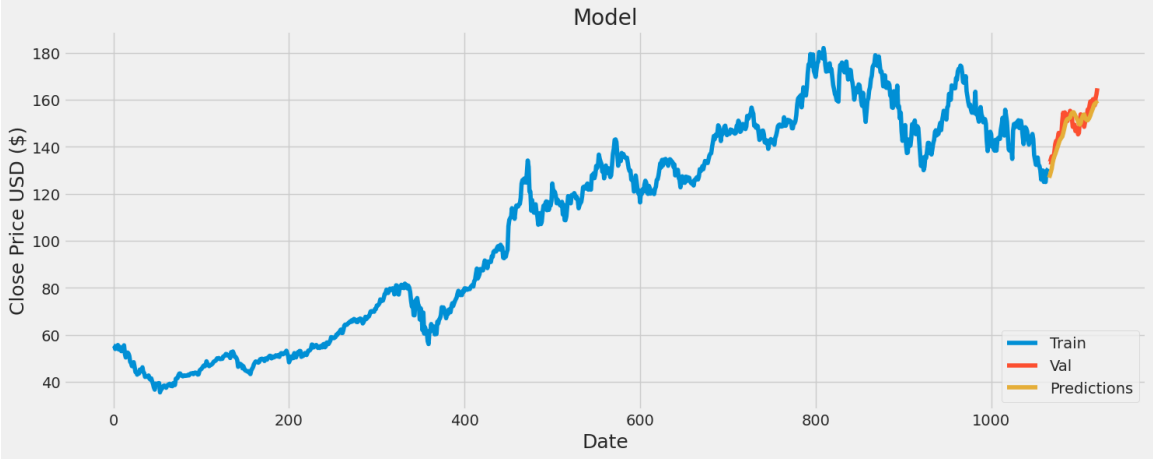


Figure - 7 Final Prediction

CHAPTER - 6

Summary and Conclusions

6.1 Summary

The proposed project aims to improve the accuracy of stock price prediction by incorporating sentimental analysis and deep learning techniques. The system will use a Long Short-Term Memory (LSTM) neural network to learn the patterns and trends in the historical stock price data and market sentiment. Sentimental analysis will be performed on social media data to capture the overall sentiment and mood of the market. The system will take into account various factors, such as past stock prices, trading volumes, financial indicators, and market sentiment, to generate a prediction of the stock price for the next day. The system will be implemented using Python and various open-source libraries such as TensorFlow, Keras, and Scikit-learn. The system will be scalable, reliable, and have a user-friendly interface that allows users to input stock ticker symbols and view the predicted stock prices. The system shall be able to comply with all relevant laws and regulations such as data protection and privacy laws. Overall, the proposed system aims to provide more accurate stock price predictions, which could be of significant value to investors and traders in making informed investment decisions.

6.2 Conclusion

In conclusion, we have presented a novel approach to predict stock prices using LSTM with sentimental analysis. Our experimental results demonstrate that our proposed approach achieves good performance in predicting stock prices with a low error rate (RMSE of 11.0157). The incorporation of sentimental analysis improves the accuracy of the model, especially for stocks in industries that are sensitive to market sentiment. The proposed approach has the potential to be useful for investors and traders who rely on accurate stock price predictions to make informed decisions. Future research could explore additional techniques for incorporating external data sources into the model and improving the model's interpretability

CHAPTER - 7

Future scope

7.1 Future Scope

The proposed project has significant potential for future research and development. Here are some potential future directions for this project:

1. Incorporating more data sources: The system can be expanded to incorporate other relevant data sources such as news articles, company reports, and economic indicators to improve the accuracy of the stock price prediction.
2. Incorporating more advanced sentimental analysis techniques: Advanced sentimental analysis techniques such as natural language processing (NLP) and sentiment lexicons can be incorporated to improve the accuracy of the market sentiment analysis.

Overall, the future scope of this project is vast and varied, and the project can be extended for future research and development to improve the accuracy and usefulness of the system.

References

[1] Xu Jiawei and Tomohiro Murata, “Stock Market Trend Prediction with Sentiment Analysis based on LSTM Neural Network”, International MultiConference of Engineers and Computer Scientists 2019 IMECS 2019, March 13-15, 2019, Hong Kong.

[2] A. Sharma, D. Bhuriya and U. Singh, "Survey of stock market prediction using machine learning approach," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2017, pp. 506-509, doi:10.1109/ICECA.2017.8212715.

[3] I. Parmar et al., "Stock Market Prediction Using Machine Learning," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 574-576, doi: 10.1109/ICSCCC.2018.8703332.