TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

People invest their savings in various forms of investments, most commonly Gold, Land, Bond, and Stocks. In the digital era every individual is using smartphone and many exchanges to buy and sell stocks but the prediction is a major task in the stock market.With the help of latest technologies, the value of stocks can be predicted with high accuracy

Stocks, also known as equity, are a security representing a holder's proportionate ownership of a corporation. Stockholders are therefore entitled to that portion of the corporation's assets and earnings. Companies issue stock in order to raise capital to finance future growth.

An AI is built to predict the trend of future stocks with the application of deep learning concept. The dataset is collected from yahoo finance using API. For better accuracy, the data is collected from the year 2012.Finally, by building a LSTM model, we put the data into it and predict the stock of particular ticker

CHAPTER 1
INTRODUCTION

A stock market, equity market, or share market is the aggregation of buyers and sellers of stocks (also called shares), which represent ownership claims on businesses; these may include securities listed on a public stock exchange, as well as stock that is only traded privately, such as shares of private companies which are sold to investors through equity crowdfunding platforms. Investment is usually made with an investment strategy in mind.

The total market capitalization of all publicly traded securities worldwide rose from US$2.5 trillion in 1980 to US$93.7 trillion at the end of 2020. As of 2016, there are 60 stock exchanges in the world. Of these, there are 16 exchanges with a market capitalization of $1 trillion or more, and they account for 87% of global market capitalization. Apart from the Australian Securities Exchange, these 16 exchanges are all in North America, Europe, or Asia. By country, the largest stock markets as of January 2022 are in the United States of America (about 59.9%), followed by Japan (about 6.2%) and United Kingdom (about 3.9%).

1.1 STOCK EXCHANGES

A stock exchange is an exchange (or bourse) where stockbrokers and traders can buy and sell shares (equity stock), bonds, and other securities. Many large companies have their stocks listed on a stock exchange. This makes the stock more liquid and thus more attractive to many investors. The exchange may also act as a guarantor of settlement. These and other stocks may also be traded "over the counter" (OTC), that is, through a dealer. Some large companies will have their stock listed on more than one exchange in different countries, so as to attract international investors.

Stock exchanges may also cover other types of securities, such as fixed-interest securities (bonds) or (less frequently) derivatives, which are more likely to be traded OTC. Trade in stock markets means the transfer (in exchange for money) of a stock or security from a seller to a buyer. This requires these two parties to agree on a price. Equities (stocks or shares) confer an ownership interest in a particular company.

Participants in the stock market range from small individual stock investors to larger investors, who can be based anywhere in the world, and may include banks, insurance companies, pension funds and hedge funds. Their buy or sell orders may be executed on their behalf by a stock exchange trader.

Some exchanges are physical locations where transactions are carried out on a trading floor, by a method known as open outcry. This method is used in some stock exchanges and commodities exchanges, and involves traders shouting bid and offer prices. The other type of stock exchange has a network of computers where trades are made electronically. An example of such an exchange is the NASDAQ.

A potential buyer bids a specific price for a stock, and a potential seller asks a specific price for the same stock. Buying or selling at the Market means you will accept any ask price or bid price for the stock. When the bid and ask prices match, a sale takes place, on a first-come, first-served basis if there are multiple bidders at a given price.

The purpose of a stock exchange is to facilitate the exchange of securities between buyers and sellers, thus providing a marketplace. The exchanges provide real-time trading information on the listed securities, facilitating price discovery. The New York Stock Exchange (NYSE) is a physical exchange, with a hybrid market for placing orders electronically from any location as well as on the trading floor. Orders executed on the trading floor enter by way of exchange members and flow down to a floor broker, who submits the order electronically to the floor trading post for the Designated market maker ("DMM") for that stock to trade the order. The DMM's job is to maintain a two-sided market, making orders to buy and sell the security when there are no other buyers or sellers. If a bid–ask spread exists, no trade immediately takes place – in this case, the DMM may use their own resources (money or stock) to close the difference. Once a trade has been made, the details are reported on the "tape" and sent back to the brokerage firm, which then notifies the investor who placed the order. Computers play an important role, especially for program trading.

The NASDAQ is an electronic exchange, where all of the trading is done over a computer network. The process is like the New York Stock Exchange. One or more NASDAQ market makers will always provide a bid and ask the price at which they will always purchase or sell 'their' stock.

The Paris Bourse, now part of Euronext, is an order-driven, electronic stock exchange. It was automated in the late 1980s. Prior to the 1980s, it consisted of an open outcry exchange. Stockbrokers met on the trading floor of the Palais Brongniart. In 1986, the CATS trading system was introduced, and the order matching system was fully automated.

People trading stock will prefer to trade on the most popular exchange since this gives the largest number of potential counter parties (buyers for a seller, sellers for a buyer) and probably the best price. However, there have always been alternatives such as brokers trying to bring parties together to trade outside the exchange. Some third markets that were popular are Instinet, and later Island and Archipelago (the latter two have since been acquired by Nasdaq and NYSE, respectively). One

advantage is that this avoids the commissions of the exchange. However, it also has problems such as adverse selection. Financial regulators have probed dark pools.

1.2 MAJOR PROBLEMS

As this stock market have been a important part of investment investors are facing many problems

- The stock graphs are unpredictable
- Beginners are facing difficulties to learn about the process of buying and selling stocks
- Has a huge risk of money loss

CHAPTER 2
DEFINE

2.1 INDIRECT VS. DIRECT INVESTMENT

Indirect investment involves owning shares indirectly, such as via a mutual fund or an exchange traded fund. Direct investment involves direct ownership of shares.

Direct ownership of stock by individuals rose slightly from 17.8% in 1992 to 17.9% in 2007, with the median value of these holdings rising from $14,778 to $17,000. Indirect participation in the form of retirement accounts rose from 39.3% in 1992 to 52.6% in 2007, with the median value of these accounts more than doubling from $22,000 to $45,000 in that time. Rydqvist, Spizman, and Strebulaev attribute the differential growth in direct and indirect holdings to differences in the way each are taxed in the United States. Investments in pension funds and 401ks, the two most common vehicles of indirect participation, are taxed only when funds are withdrawn from the accounts. Conversely, the money used to directly purchase stock is subject to taxation as are any dividends or capital gains they generate for the holder. In this way, the current tax code incentivizes individuals to invest indirectly.

2.2 PARTICIPATION BY INCOME AND WEALTH STRATA

Rates of participation and the value of holdings differ significantly across strata of income. In the bottom quintile of income, 5.5% of households directly own stock and 10.7% hold stocks indirectly in the form of retirement accounts. The top decile of income has a direct participation rate of 47.5% and an indirect participation rate in the form of retirement accounts of 89.6%. The median value of directly owned stock in the bottom quintile of income is $4,000 and is $78,600 in the top decile of income as of 2007. The median value of indirectly held stock in the form of retirement accounts for the same two groups in the same year is $6,300 and $214,800 respectively. Since the Great Recession of 2008 households in the bottom half of the income distribution have lessened their participation rate both directly and indirectly from 53.2% in 2007 to 48.8% in 2013, while over the same period households in the top decile of the income distribution slightly increased participation 91.7% to 92.1%. The mean value of direct and indirect holdings at the bottom half of the income distribution moved slightly downward from $53,800 in 2007 to $53,600 in 2013. In the top decile, mean value of all holdings fell from $982,000 to $969,300 in the same time. The mean value of all stock holdings across the entire income distribution is valued at $269,900 as of 2013.

2.3 PARTICIPATION BY RACE AND GENDER

The racial composition of stock market ownership shows households headed by whites are nearly four and six times as likely to directly own stocks than households headed by blacks and Hispanics respectively. As of 2011 the national rate of direct participation was 19.6%, for white households the participation rate was 24.5%, for black households it was 6.4% and for Hispanic households it was 4.3%. Indirect participation in the form of 401k ownership shows a similar pattern with a national participation rate of 42.1%, a rate of 46.4% for white households, 31.7% for black households, and 25.8% for Hispanic households. Households headed by married couples participated at rates above the national averages with 25.6% participating directly and 53.4% participating indirectly through a retirement account. 14.7% of households headed by men participated in the market directly and 33.4% owned stock through a retirement account. 12.6% of female-headed households directly owned stock and 28.7% owned stock indirectly.

2.4 DETERMINANTS AND POSSIBLE EXPLANATIONS OF STOCK MARKET PARTICIPATION

In a 2003 paper by Vissing-Jørgensen attempts to explain disproportionate rates of participation along wealth and income groups as a function of fixed costs associated with investing. Her research concludes that a fixed cost of $200 per year is sufficient to explain why nearly half of all U.S. households do not participate in the market. Participation rates have been shown to strongly correlate with education levels, promoting the hypothesis that information and transaction costs of market participation are better absorbed by more educated households. Behavioural economists Harrison Hong, Jeffrey Kubik and Jeremy Stein suggest that sociability and participation rates of communities have a statistically significant impact on an individual's decision to participate in the market. Their research indicates that social individuals living in states with higher-than-average participation rates are 5% more likely to participate than individuals that do not share those characteristics. This phenomenon also explained in cost terms. Knowledge of market functioning diffuses through communities and consequently lowers transaction costs associated with investing.

2.5 DATA SET WITH DESCRIPTION

Variables	Description
Date	Date for stock in each days
Open	Open price of the stock in that date
High	High price of the stock in that date
Low	Low price of the stock in that date
Close	Close price of the stock in that date
Volume	Volume of the stock in that date
Adj Close	Adjacent Close price of the stock in that date

	Date	Open	High	Low	Close	Volume	Adj Close
0	6/29/2010	19.000000	25.00	17.540001	23.889999	18766300	23.889999
1	6/30/2010	25.790001	30.42	23.299999	23.830000	17187100	23.830000
2	7/1/2010	25.000000	25.92	20.270000	21.959999	8218800	21.959999
3	7/2/2010	23.000000	23.10	18.709999	19.200001	5139800	19.200001
4	7/6/2010	20.000000	20.00	15.830000	16.110001	6866900	16.110001

CHAPTER 3
SYSTEM REVIEW

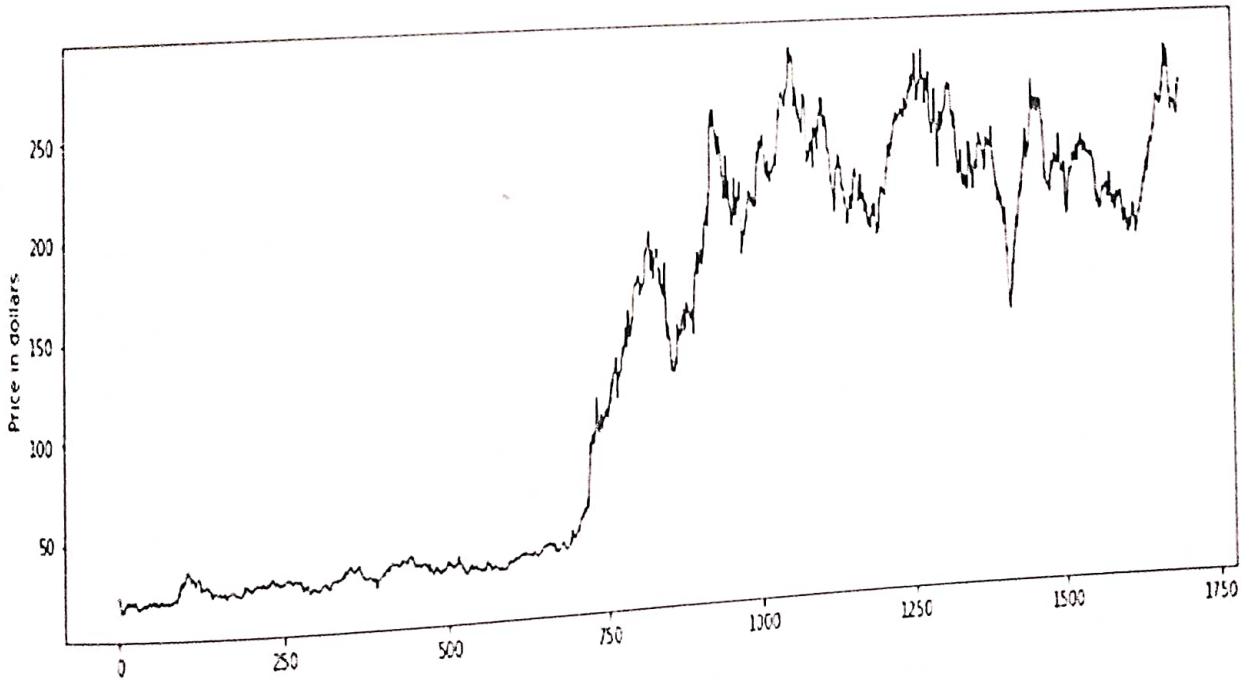
These problems can be resolved using a developing a app to predict the stock price An AI is built to predict the trend of future stocks with the application of deep learning concept.

The dataset is collected from yahoo finance using API. For better accuracy, the data is collected from the year 2012.

Finally, by building a LSTM model, we put the data into it and predict the stock of particular ticker.

3.1 EXPLORATORY DATA ANALYSIS

Exploratory Data Analysis (EDA) is an approach to analyze the data using visual techniques. It is used to discover trends, patterns, or to check assumptions with the help of statistical summary and graphical representations.



3.1 EXPLORATORY DATA ANALYSIS

3.2 SOFTWARE & LIBRARIES

By using python programming language and its libraries software is developed

SOFTWARE / LIBRARIES	USAGE
Python3 software	Programming language
Pandas	To work with data in data frames and data exploration
NumPy	For scaling and reshaping data
Sklearn	Pre-processing, normalizing, scaling
Matplotlib	For data visualization
TensorFlow, Kera's	For LSTM model

CHAPTER 4

PROJECT IMPLEMENTATION

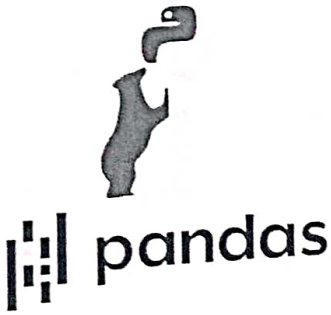
4.1 PYTHON3 SOFTWARE

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation

Python is a multi-paradigm programming language. Object-oriented programming and structured programming are fully supported, and many of their features support functional programming and aspect-oriented programming (including metaprogramming and metaobjects). Many other paradigms are supported via extensions, including design by contract and logic programming.

4.2 PANDAS

Pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. It is free software released under the three-clause BSD license. The name is derived from the term "panel data", an econometrics term for data sets that include observations over multiple time periods for the same individuals. Its name is a play on the phrase "Python data analysis" itself. Wes McKinney started building what would become pandas at AQR Capital while he was a researcher there from 2007 to 2010.



4.2 PANDAS SOFTWARE

4.3 NumPy

NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays. The ancestor of NumPy, Numeric, was originally created by Jim Hugunin with contributions from several other developers. In 2005, Travis Oliphant created NumPy by incorporating features of the competing Numarray into Numeric, with extensive modifications. NumPy is open-source software and has many contributors. NumPy is a NumFOCUS fiscally sponsored project.

4.4 SK LEARN

Scikit-learn (formerly **scikits.learn** and also known as **sklearn**) is a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support-vector machines, random forests, gradient boosting, *k*means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy. Scikit-learn is a NumFOCUS fiscally sponsored project.

4.5 MATPLOTLIB

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK. There is also a procedural "pylab" interface based on a state machine (like OpenGL), designed to closely resemble that of MATLAB, though its use is discouraged.[3] SciPy makes use of Matplotlib.

4.6 TENSORFLOW KERAS

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

Kera's is an open-source software library that provides a Python interface for artificial neural networks. Kera's acts as an interface for the TensorFlow library.

4.7 ARCHITECTURAL DIAGRAM

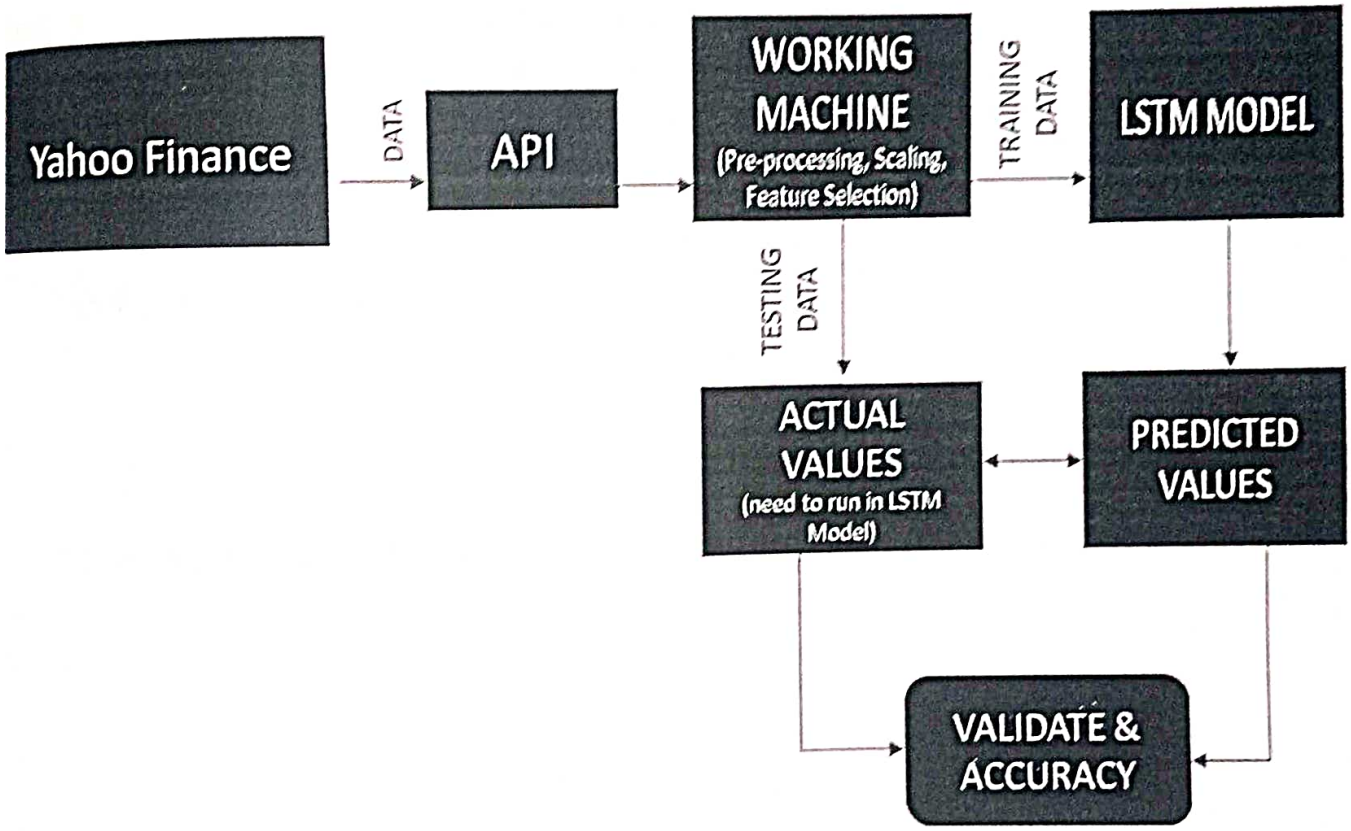


FIG 4.7 ARCHITECTURAL DIAGRAM

4.8 LSTM MODEL

- Unique version of RNN with 4 layers
- Remembering information for longer periods of time is their default behavior
- Made up of a memory cell, an input gate, an output gate and a forget gate

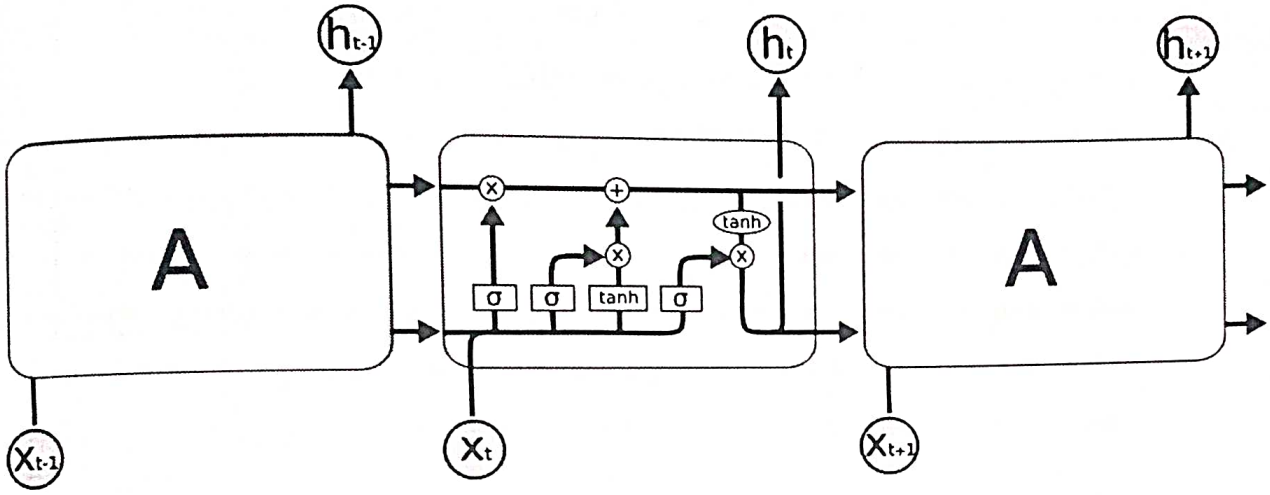


FIG 4.8 LSTM MODEL

4.9 PREDICTION GRAPH

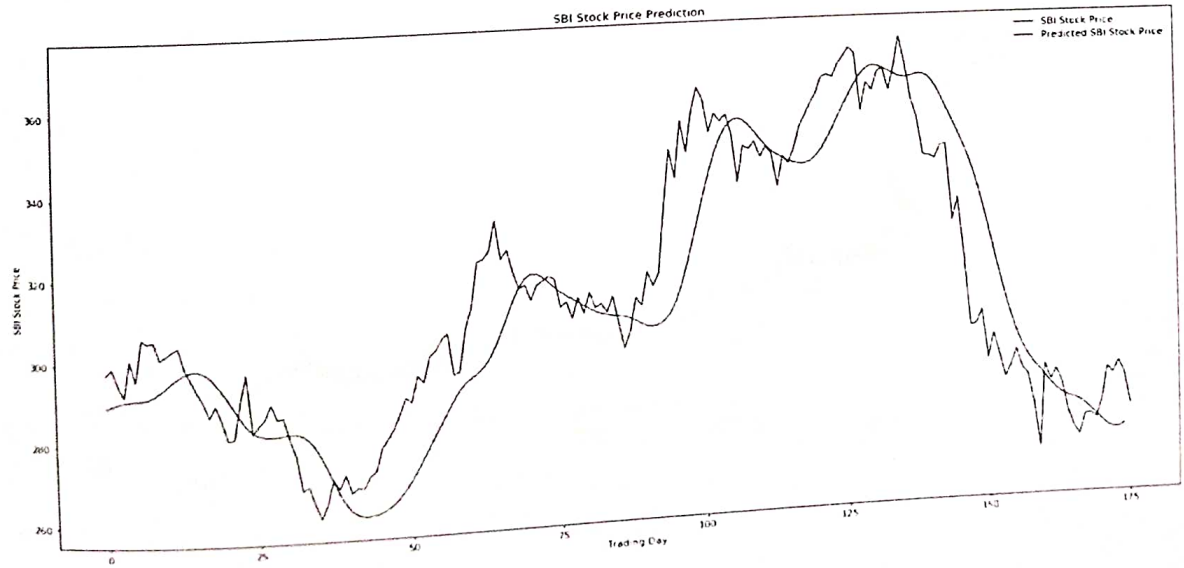
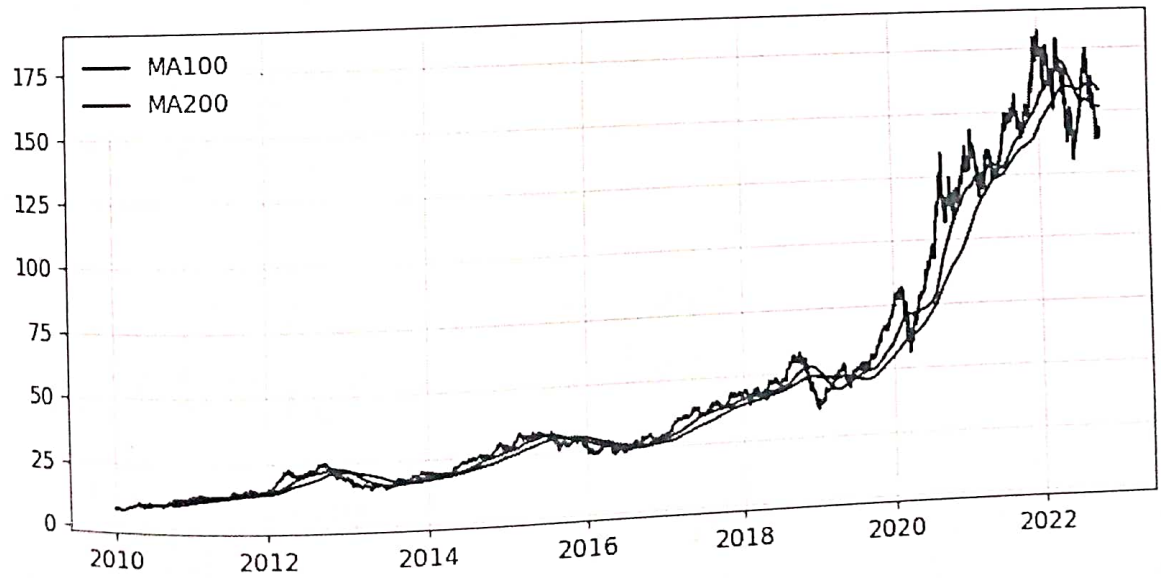


FIG 4.9 PREDICTION GRAPH

CHAPTER 5
TESTING

```
model.fit(X_train, y_train, validation_data=(X_test,y_test), epochs=50, batch_size=64, verbose=1)
```

```
Epoch 1/50
34/34 [==============================] - 19s 212ms/step - loss: 0.0018 - val_loss: 0.0053
Epoch 2/50
34/34 [==============================] - 6s 178ms/step - loss: 8.9337e-05 - val_loss: 0.0019
Epoch 3/50
34/34 [==============================] - 6s 175ms/step - loss: 5.9041e-05 - val_loss: 0.0040
Epoch 4/50
34/34 [==============================] - 6s 175ms/step - loss: 5.3890e-05 - val_loss: 0.0034
Epoch 5/50
34/34 [==============================] - 6s 185ms/step - loss: 5.1661e-05 - val_loss: 0.0061
Epoch 6/50
34/34 [==============================] - 6s 170ms/step - loss: 5.0241e-05 - val_loss: 0.0052
Epoch 7/50
34/34 [==============================] - 6s 173ms/step - loss: 5.0674e-05 - val_loss: 0.0081
Epoch 8/50
34/34 [==============================] - 6s 176ms/step - loss: 4.8786e-05 - val_loss: 0.0069
Epoch 9/50
34/34 [==============================] - 9s 267ms/step - loss: 4.8219e-05 - val_loss: 0.0069
Epoch 10/50
```



Set X_train, y_train and X_test, y_test values

```
def setValues(dataset, time_step):
    dataX, dataY = [], []
    for i in range(0, len(dataset)-time_step-1):
        a = dataset[i:(i+time_step), 0]
        dataX.append(a)
        dataY.append(dataset[i+time_step, 0])
    return np.array(dataX), np.array(dataY)

time_step = 100
X_train, y_train = setValues(train_data, time_step)
X_test, y_test = setValues(test_data, time_step)
```

```
print(len(X_train),X_train.shape)
print(len(y_train),y_train.shape)
```

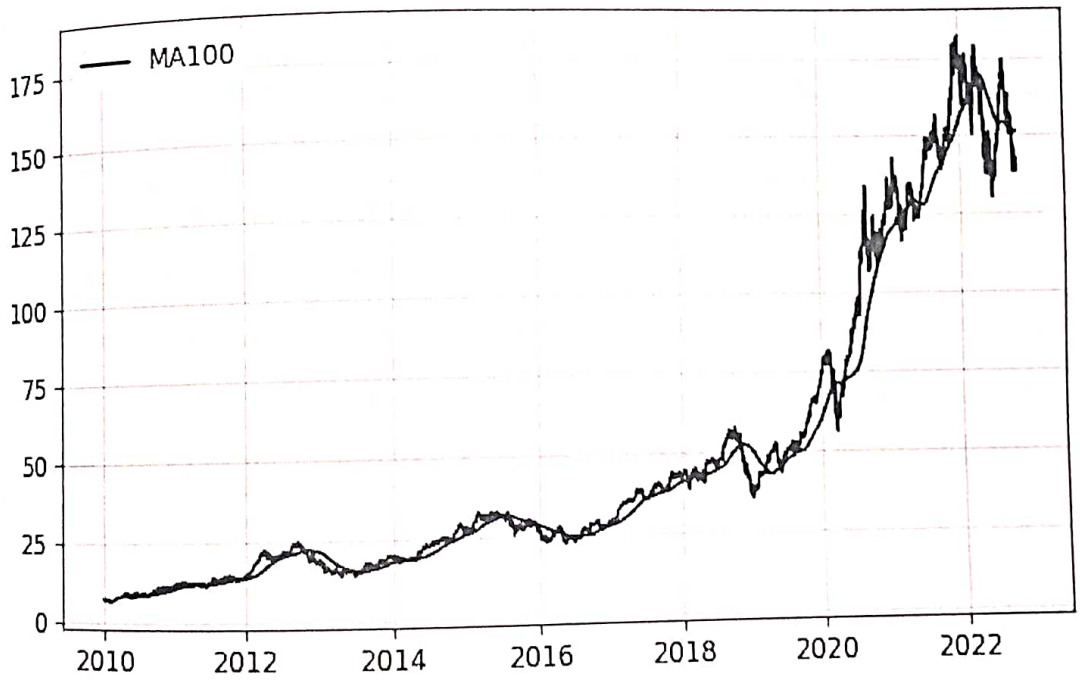
```
2153 (2153, 100)
2153 (2153,)
```

```
print(len(X_test),X_test.shape)
print(len(y_test),y_test.shape)
```

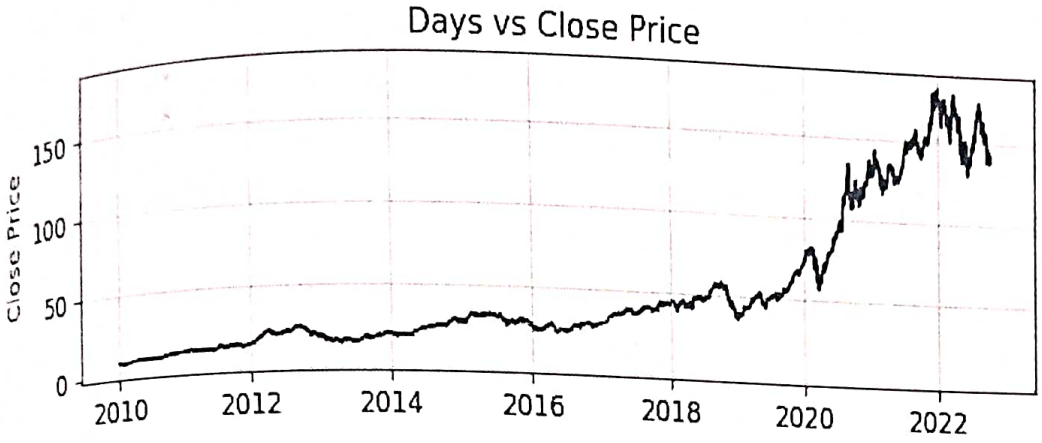
```
866 (866, 100)
866 (866,)
```

df.describe()

	High	Low	Open	Close	Volume	Adj Close
count	3221.000000	3221.000000	3221.000000	3221.000000	3.221000e+03	3221.000000
mean	50.348448	49.261857	49.798959	49.825841	2.590953e+08	48.025237
std	46.810203	45.658188	46.221046	46.256876	2.232075e+08	46.761350
min	7.000000	6.794643	6.870357	6.858929	4.100000e+07	5.856372
25%	19.016787	18.678572	18.885357	18.882143	1.039164e+08	16.556959
50%	29.602501	29.155001	29.379999	29.400000	1.694102e+08	27.182737
75%	55.570000	54.317501	54.990002	54.972500	3.525844e+08	53.254868
max	182.940002	179.119995	182.630005	182.009995	1.880998e+09	181.259933




```
from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler(feature_range=(0,1))
df1 = scaler.fit_transform(np.array(df1).reshape(-1,1))
```




```
df.head(5)
```

Date	High	Low	Open	Close	Volume	Adj Close
2009-12-31	7.619643	7.520000	7.611786	7.526071	352410800.0	6.426000
2010-01-04	7.660714	7.585000	7.622500	7.643214	493729600.0	6.526021
2010-01-05	7.699643	7.616071	7.664286	7.656429	601904800.0	6.537303
2010-01-06	7.686786	7.526786	7.656429	7.534643	552160000.0	6.433319
2010-01-07	7.571429	7.466071	7.562500	7.520714	477131200.0	6.421425

Model Creation

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
from tensorflow.keras.layers import LSTM

model = Sequential()

model.add(LSTM(50, return_sequences=True, input_shape=(100,1)))
model.add(LSTM(50, return_sequences=True))
model.add(LSTM(50))
model.add(Dense(1))

model.compile(loss='mean_squared_error', optimizer='adam')

model.summary()
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 100, 50)           10400

 lstm_1 (LSTM)               (None, 100, 50)           20200

 lstm_2 (LSTM)               (None, 50)                20200

 dense (Dense)               (None, 1)                 51

=================================================================
Total params: 50,851
Trainable params: 50,851
Non-trainable params: 0
```

CHAPTER 6
CONCLUSION

6.1 CONCLUSION

Thus, stock market price prediction is quite possible by using this application Though price prediction is a mathematical strategy this application would crack the strategy and make the prediction easier. This application needs improvement with the modules in near future

CHAPTER 7
FUTURE WORKS

7.1 FUTURE WORKS

The model can be well trained with more data for higher accuracy. Apart from LSTM, there are more effective algorithms which can be trained heavily but in highend machines

REFERENCE

- https://en.wikipedia.org/w/index.php?title=Stock_market&action=edit&section=21
- https://en.wikipedia.org/w/index.php?title=Stock_market&action=edit&section=2
- https://en.wikipedia.org/w/index.php?title=Stock_market&action=edit&section=4
- https://en.wikipedia.org/w/index.php?title=Stock_market&action=edit&section=6
- https://en.wikipedia.org/w/index.php?title=Stock_market&action=edit&section=7
- https://en.wikipedia.org/w/index.php?title=Stock_market&action=edit&section=26
- https://en.wikipedia.org/w/index.php?title=Stock_market&action=edit&section=36