Stock Price Prediction

Submitted in partial fulfillment of the requirements for the award of the Degree

of

Bachelor of Technology (B. Tech)

in

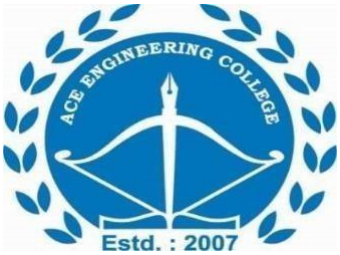
INFORMATION TECHNOLOGY

By

Doli SaiVishal 22AG1A1217

Under the esteemed Guidance of

Mrs. V .Vanaja
Assistant Professor



Department of Information Technology

ACE ENGINEERING COLLEGE

An Autonomous Institution

All the courses are Accredited by NBA and NAAC with A Grade

(Affiliated to Jawaharlal Nehru Technological University, Hyderabad, Telangana)

Ankushapur, Ghatkesar, Medchal, Hyderabad – 501 301

ACE

Engineering College

An Autonomous Institution

All the courses are Accredited by NBA and NAAC with A Grade

(Affiliated to Jawaharlal Nehru Technological University, Hyderabad, Telangana)

Website : www.aceec.ac.in E-mail: info@aceec.ac.in

CERTIFICATE

This is to certify that the Project work entitled "**Stock Price Pridiction"** is being submitted By **DoliSaiVishal(22AG1A1217)** in partial fulfilment for the award of Degree of **BACHELOR OF TECHNOLOGY in INFORMATION TECHNOLOGY** to the Jawaharlal Nehru Technological University, Hyderabad during the academic year 2023-2024 is a record of bona-fide work carried out by them under our guidance and supervision.

The results embodied in this report have not been submitted by the students

to any other university or institution for the award of any degree or diploma.

Internal Guide

Mrs.V.Vanaja

Assistant professor

Head of the Department

Prof. K. JAYA BHARATHI

Professor and Head

Dept. of IT

ACKNOWLEDGEMENT

We would like to express our gratitude to all the people behind the screen who have helped us to transform an idea into a real time application.

We would like to express our heart-felt gratitude to our parents without whom we would not have been privileged to achieve and fulfil our dreams.

A special thanks to our Secretary, **Prof. Y. V. GOPALA KRISHNA MURTHY** and Joint Secretary, **Mrs. M. PADMAVATHI** for having founded such an esteemed institution. We are also grateful to our beloved principal, **Dr. B. L. RAJU** for permitting us to carry out this project.

We profoundly thank **Prof. K. JAYA BHARATHI**, Head of the Department of Information Technology, who has been an excellent guide and a great source of inspiration to our work.

We are very thankful to our internal guide **Mrs.V.Vanaja**, who has been an excellent and given continuous support for the completion of our project work.

The satisfaction and euphoria that accompany the successful completion of the task would be great, but incomplete without the mention of the people who made it possible, whose guidance and encouragement crown all the efforts with success. In this context, we would like to thank all the other staff members, teaching and non- teaching, who have extended their timely help and eased our task.

Doli SaiVishal 22AG1A1217

DECLARATION

We hereby declare that the project entitled "**STOCK PRICE PRIDICTION**" submitted in partial fulfilment of the requirements for the award of degree of Bachelor of Technology in Information Technology. This dissertation is our original work and the project has not formed the basis for the award of any degree, associate ship, fellowship or any other similar titles and no part of it has been published or sent for the publication at the time of submission.

Doli SaiVishal 22AG1A1217

ABSTRACT

STOCK PRICE PRIDICTION

Our project aims to predict stock prices using machine learning algorithms. We gather historical stock data and employ various machine learning techniques to analyze and forecast future prices.Additionally, sentiment analysis of news articles and social media data is incorporated to enhance prediction accuracy. The performance of different models is evaluated using metrics like mean squared error and accuracy. Through this project, we seek to develop a reliable framework for predicting stock prices, aiding investors in making informed decisions.

INDEX

LIST OF FIGURES

List of Figures **Page No.**

LIST OF ABBREVATIONS

Abbreviation

DFD	Data Flow Diagram

CHAPTER 1
INTRODUCTION

1.1 MOTIVATION

Stock price prediction is a classic and important problem. With a successful model for stock prediction, we can gain insight about market behavior over time, spotting trends that would otherwise not have been noticed. With the increasingly computational power of the computer, machine learning will be an efficient method to solve this problem. However, the public stock dataset is too limited for many machine learning algorithms to work with, while asking for more features may cost thousands of dollars every day. our motivation is to design a public service incorporating historical data and users predictions to make a stronger model that will benefit everyone.

1.2 PROBLEM STATEMENT

The problem statement of stock price prediction is to develop a model or system that can accurately forecast the future prices of stocks or financial assets. The aim is to analyze historical price and volume data, as well as other relevant factors such as market trends, company news, economic indicators, and investor sentiment, to make predictions about the future movement of stock prices. The ultimate goal is to assist investors, traders, and financial institutions in making informed decisions about buying, selling, or holding stocks in order to maximize profits or minimize losses. The challenge lies in the complex and dynamic nature of financial markets, where stock prices are influenced by a multitude of factors, making accurate predictions a difficult task.

1.3 OBJECTIVE OF PROJECT

The objective of stock price prediction is to accurately forecast the future prices of stocks or financial assets. The main goals include:

1. Investment decision-making: The primary objective is to assist investors, traders, and financial institutions in making informed decisions about buying, selling, or holding stocks.

CHAPTER 2

LITERATURE SURVEY

A Recurrent Neural Network [RNN] approach in predicting daily stock prices an application to the Sri Lankan stock market in 2020 IEEE In national Conference on Industrial and Information Systems(ICHS).

• Author: A . Samarawickarma

• Algorithm: RNN

• Accuracy: 75%

21 A LSTM method for bitcoin price prediction in 2019 In national Conference on Electrical Engineering and CompuCHAPTER Science (ICECOS) presented by

• Author: S.H Othman

• Algorithm : LSTM

• Accuracy: 89%

31. A model for Stock price prediction using Reinforcement learning presented in 2021 IEEE In national Symposium on Industrial Electronics Proceedings.

• Author: Jae won lee

• Algorithm: Reinforcement Learning

• Accuracy: 85%

4]. A Timeseries Forecasting with RNN and LSTM predicting Nifty 50 stock prices presented by

• Author: Satyabrata Mishra

• Algorithm: RNN and LSTM

• Accuracy:80%

2.1 EXISTING SYSTEM

An Existing system for stock price prediction use traditional machine learning techniques such as RNN to make predictions based on historical price data due to their ability to process large volumes of data and identify complex .

Limitation of RNN:

RNN suffers from the vanishing gradient problem.

These algorithms, such as linear regression, support vector machines, random forests, and neural networks, learn from historical price data, financial indicators, and other relevant features generate

2.2 PROPOSED SYSTEM

Our proposed system combines multiple techniques to enhance the accuracy of predictions a multiple model algorithms LSTM with GRU may incorporate both fundamental analysis and machine learning algorithms to leverage the strengths of each approach.

By combining different methodologies, these proposed systems aim to improve the robustness and accuracy of stock price predictions.

CHAPTER 3

SOFTWARE REQUIREMENT ANALYSIS

FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are:

O ECONOMICAL FEASIBILITY

• TECHNICAL FEASIBILITY

• SOCIAL FEASIBILITY

3.1 ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources.

3.2 SOCIAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

3.3 FUNCTIONAL REQUIREMENTS : Functional Requirement defines a function of a software system and how the system must behave when presented with specific inputs or conditions. These may include calculations, data manipulation and processing and other specific functionality.

CHAPTER 4

SOFTWARE DESIGN

4.1 INTRODUCTION

Implementation is the stage of the project when the theoritical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.

The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve change over and evaluation of change over methods.

4.2 ARCHITECTURE

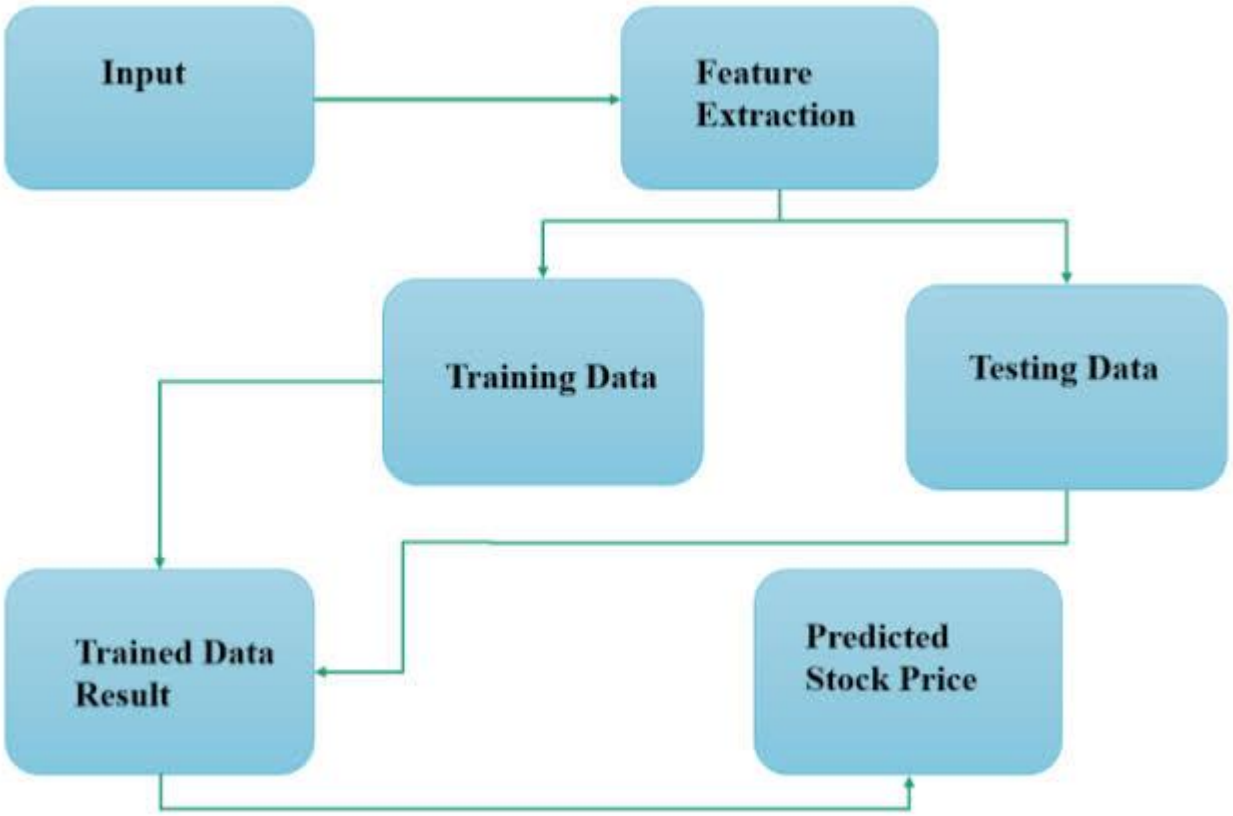


Fig-1

4.3 ACTIVITY DIAGRAM

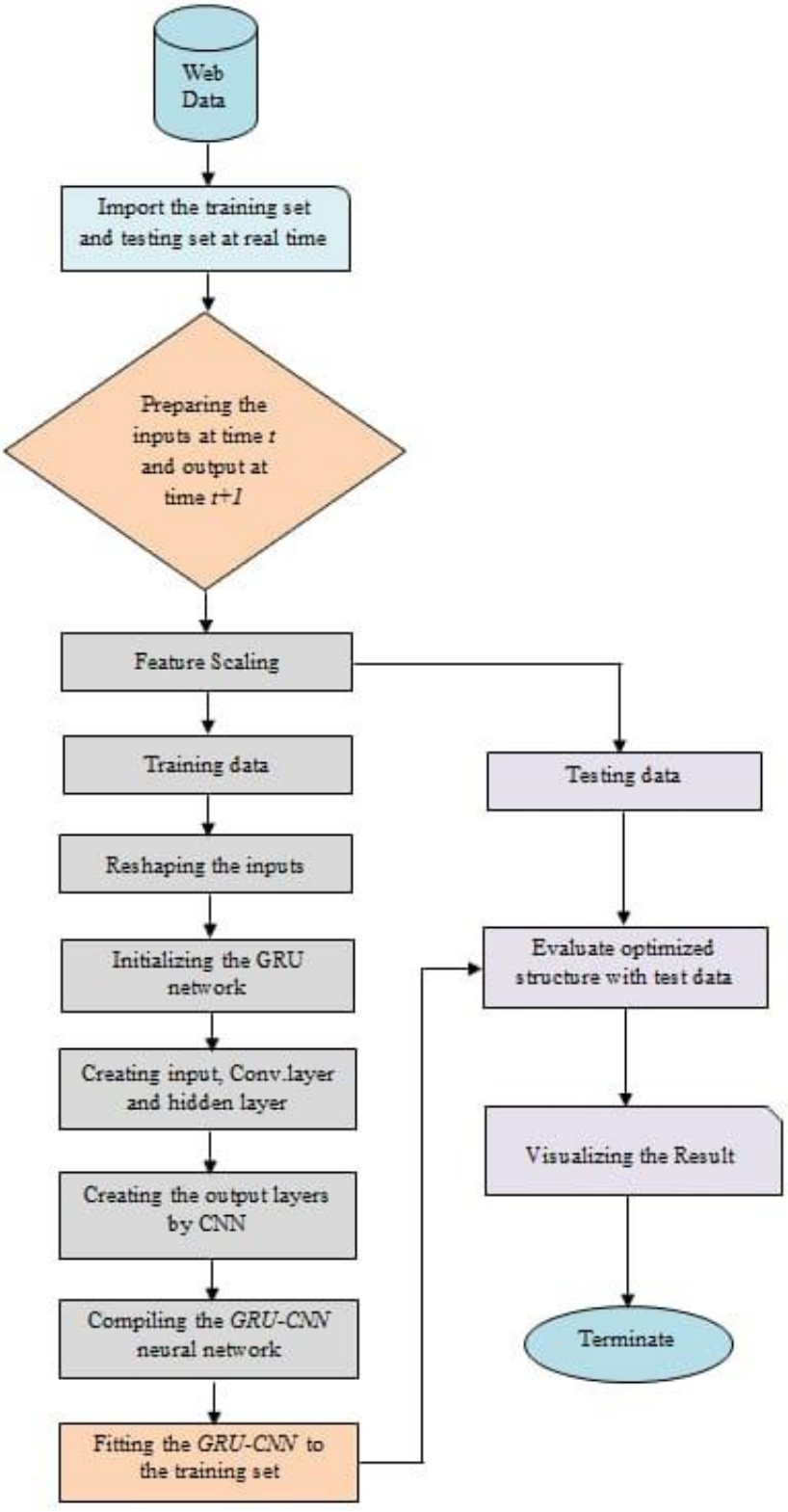
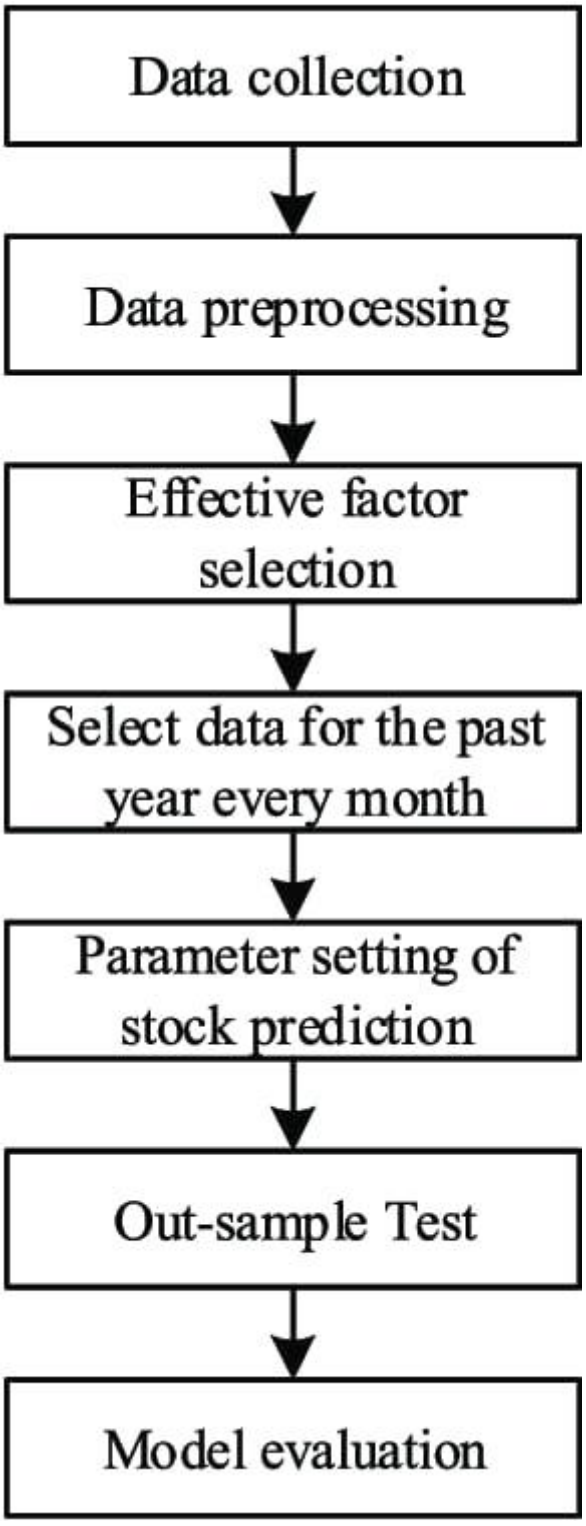


Fig-1

Data Flow Diagram:



CHAPTER 5

Software and Hardware requirements

5.1 HARDWARE REQUIREMENTS

Processor: 64-bit, four-core, 2.5 GHz minimum per core
RAM: 4 GB
Hard Disk: 250 GB
Input device: Standard Keyboard and Mouse.

5.2 SOFTWARE REQUIREMENTS

Operating System: Windows 10
Languages: Python
Frontend: Flask

CHAPTER 6
MODULES

There are Four modules in this projects:-

- DATA ACQUISITION
- DATA PREPROCESSING
- BUILDING THE MODEL
- PREDICTION

6.1 DATA ACQUISITION

The Stock price historical data were acquired from the yahoo finance website. The provided one-minute in val data which contain approximately 4.8 millions samples, including NaN values. Some columns in the data are open, high, low, close (OHLC) price, and the weighted price. All the timestamps are in UNIX time. The NaN values indicate that no trade or activity happened at that time.

6.2 DATA PREPROCESSING

Some pre-processing steps are performed to clean the stock price historical data, including feature selection, timestamp conversion, missing values removal, train-test split, and min-max scaling normalization.

As each dataset consists of many features, this work only utilizes three features for price prediction, namely timestamp, date, and closing price. Subsequently, timestamp conversion is carried out where the timestamp in UNIX is converted into the YY:MM:DD date format. The zeros and NaNs are filed out by dropping the associated rows. To avoid huge data losses and to provide more timely and detailed prediction, the samples are taken at one-min individuals. Due to the inconsistency of historical data and high sampling rates, the historical data of one week are used. The samples are further partitioned into six days for the training set and one day for the testing set. Apart from that, the features are subjected to min-max scaling normalization that transforms each feature into the range [0, 1].

6.3 BUILDING THE MODEL

Building the model with LSTM and GRU gives bet accuracy comparing with Other Models.

All models that are built so far do not allow for operating on sequence data.

LSTM (Long Short term- Memory) and GRU (Gated Recurrent Unit): These are types of recurrent neural network (RNN) architectures that can effectively model sequential data such as time series.

They have been widely used for stock price prediction due to their ability to capture long-term dependencies. LSTM and GRU networks can be implemented using deep learning frameworks such as TensorFlow or PyTorch.

An LSTM module (or cell) has 5 essential components which allows it to model both long term and short-term data. Hidden state (ht) - This is output state information calculated w.r.t. current input, previous hidden state and current cell input which you eventually use to predict the future stock market prices.

GRU is one of the variants of RNN. By introducing gating structure, it solves the problem that RNN is difficult to deal with long-distance information acquisition. Compared with LSTM, GRU is simplified and only update gate and reset gate are introduced. In GRU, the update (or input) gate decides how much input and previous output to be passed to the next cell and the reset gate is used to determine how much of the past information to forget. The current memory content ensures that only the relevant information needs to be passed to the next , which is determined by the weight W.

6.4 PREDICTION

The given input is preprocessed and the data regarding stock price is scraped based on yahoo finance. Then based on the input of stocks given it analyses from the dataset. With all this information, then defined different algorithm models are used for prediction. Final outcome is displayed through frontend. Result is displayed in the form of graphs of comparing last 15 days vs next 30 days plotting whole closing stock price prediction and final stock analysis chart.

Extension - LSTM ,SVR,RF, KNN,GRU,LSTM and GRU Machine Learning models are used with

proposed system one from which LSTM and GRU gives better accuracy and from that LSTM and GRU is used to build the model and save it from user prediction.

INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

• What data should be given as input?

• How the data should be arranged or coded?

• The dialog to guide the operating personnel in providing input. » Methods for preparing input validations and steps to follow when error occur.

OBJECTIVES

1. Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.

2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

3. When the data is entered it will check for its validity. Data can be entered with the help of screens.

OUTPUT DESIGN

A quality output is one, which mects the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.

Select methods for presenting information.

3. Create document, report, or other formats that contain information produced by the system. The output form of an information system should accomplish one or more of the following objectives.

• Convey information about past activities, current status or projections of the Future.

• Signal important events, opportunities, problems, or warnings.

• Trigger an act

CHAPTER 7

CODING TEMPLATES

```
import streamlit as st
import pandas as pd
import numpy as np
from keras.models import load_model
import matplotlib.pyplot as plt
import yfinance as yf

st.title("Stock Price Predictor App")

stock = st.text_input("Enter the Stock ID", "GOOG")

from datetime import datetime
end = datetime.now()
start = datetime(end.year-20,end.month,end.day)

google_data = yf.download(stock, start, end)

model = load_model("Latest_stock_price_model.keras")
st.subheader("Stock Data")
st.write(google_data)

splitting_len = int(len(google_data)*0.7)
x_test = pd.DataFrame(google_data.Close[splitting_len:])

def plot_graph(figsize, values, full_data, extra_data = 0, extra_dataset = None):
    fig = plt.figure(figsize=figsize)
```

```
    plt.plot(values,'Orange')
    plt.plot(full_data.Close, 'b')
    if extra_data:
        plt.plot(extra_dataset)
    return fig

st.subheader('Original Close Price and MA for 250 days')
google_data['MA_for_250_days'] = google_data.Close.rolling(250).mean()
st.pyplot(plot_graph((15,6), google_data['MA_for_250_days'],google_data,0))

st.subheader('Original Close Price and MA for 200 days')
google_data['MA_for_200_days'] = google_data.Close.rolling(200).mean()
st.pyplot(plot_graph((15,6), google_data['MA_for_200_days'],google_data,0))

st.subheader('Original Close Price and MA for 100 days')
google_data['MA_for_100_days'] = google_data.Close.rolling(100).mean()
st.pyplot(plot_graph((15,6), google_data['MA_for_100_days'],google_data,0))

st.subheader('Original Close Price and MA for 100 days and MA for 250 days')
st.pyplot(plot_graph((15,6),
google_data['MA_for_100_days'],google_data,1,google_data['MA_for_250_days'])
)

from sklearn.preprocessing import MinMaxScaler

scaler = MinMaxScaler(feature_range=(0,1))
scaled_data = scaler.fit_transform(x_test[['Close']])

x_data = []
y_data = []
```

```
for i in range(100,len(scaled_data)):
    x_data.append(scaled_data[i-100:i])
    y_data.append(scaled_data[i])

x_data, y_data = np.array(x_data), np.array(y_data)

predictions = model.predict(x_data)

inv_pre = scaler.inverse_transform(predictions)
inv_y_test = scaler.inverse_transform(y_data)

ploting_data = pd.DataFrame(
 {
  'original_test_data': inv_y_test.reshape(-1),
    'predictions': inv_pre.reshape(-1)
 } ,
    index = google_data.index[splitting_len+100:]
)
st.subheader("Original values vs Predicted values")
st.write(ploting_data)

st.subheader('Original Close Price vs Predicted Close price')
fig = plt.figure(figsize=(15,6))
plt.plot(pd.concat([google_data.Close[:splitting_len+100],ploting_data], axis=0))
plt.legend(["Data- not used", "Original Test data", "Predicted Test data"])
st.pyplot(fig)
```

CHAPTER 8

TESTING AND VALIDATION

7.1 TESTING STRATEGIES

7.1.1UNIT TESTING:

Unit testing, a testing technique using which individual modules are tested to determine if there are issues by the developer himself. it is concerned with functional correctness of the standalone modules. The main aim is to isolate each unit of the system to identify, analyze and fix the defects.

Unit Testing Techniques:

Black Box Testing - Using which the user interface, input and output are tested.

White Box Testing -Used to test each one of those functions behavior is tested.

7.1.2DATA FLOW TESTING:

Data flow testing is a family of testing strategies based on selecting paths through the program's control flow in order to explore sequence of events related to the status of Variables or data object. Dataflow Testing focuses on the points at which variables receive and the points at which these values are used.

7.1.3INTEGRATION TESTING:

Integration Testing done upon completion of unit testing, the units or modules are to be integrated which gives raise too integration testing. The purpose of integration testing is to verify the functional, performance, and reliability between the modules that are integrated.

Epoch 195/200

67/67 [==========]-1s 8ms/step-loss:0.0023-val_loss:0.0027

Epoch 196/200

67/67 [==========]-1s 8ms/step-loss:0.0027-val_loss:0.0013

Epoch 197/200

67/67 [==========] - Is 8ms/step - loss: 0.0025 - val _loss: 0.0015

Epoch 198/200

67/67 [==========] - 1s 8ms/step - loss: 0.0025 - val_loss: 0.0014

Epoch 199/200

67/67 [==========] - 1s 8ms/step - loss: 0.0025 - val _loss: 0.0014

Epoch 200/200

67/67 [==========] - Is 8ms/step - loss: 0.0025 - val loss: 0.002

11/11[==========] - 0s 2ms/step

5/5 [==========] - 0s 2ms/step

127.0.0.1 - - [22/Jun/2023 23:26:43] "POST /home HTTP/1.1" 200 -

127.0.0.1 - - [22/Jun/2023 23:26:43] "GET /static/app.js HTTP/1.1" 30

TEST CASE	IF AVAILABLE	IF UNAVAILABLE
User login	User login to be web page	There is no process
Enter the credentials	User can proceed next	There is no process

CHAPTER 9

OUTPUT SCREENS OF THE PROJECT

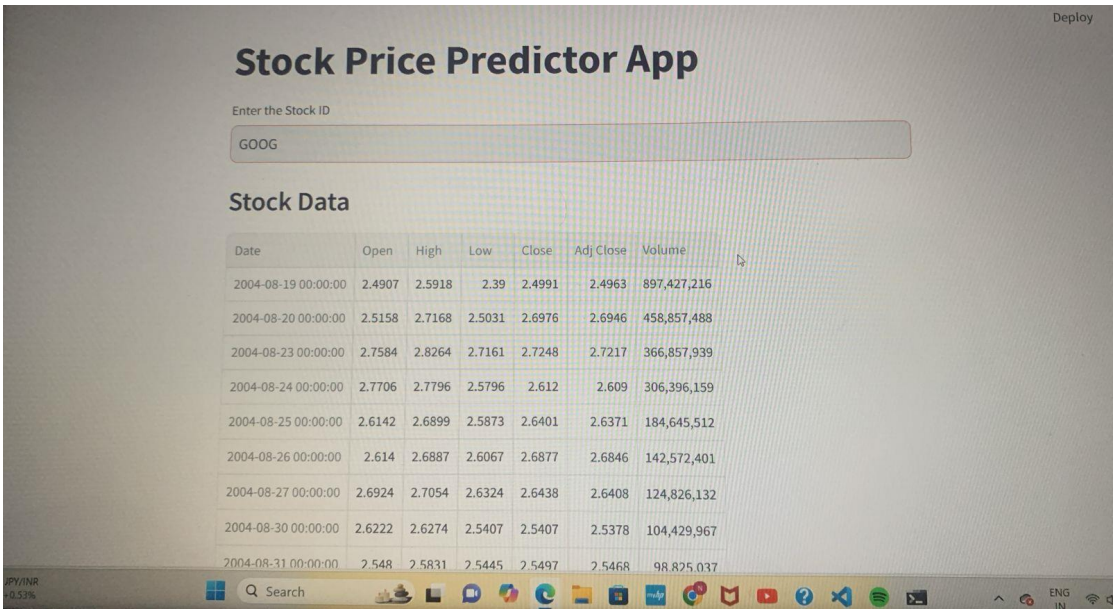
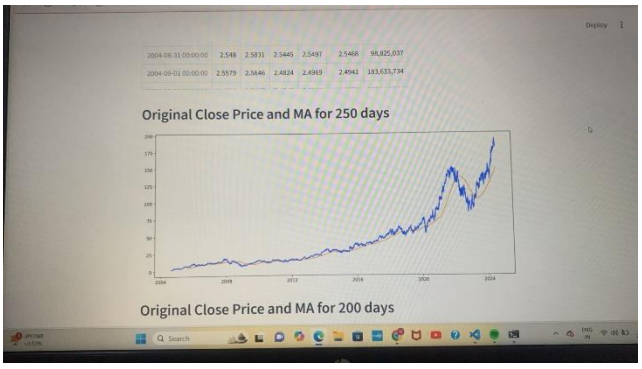
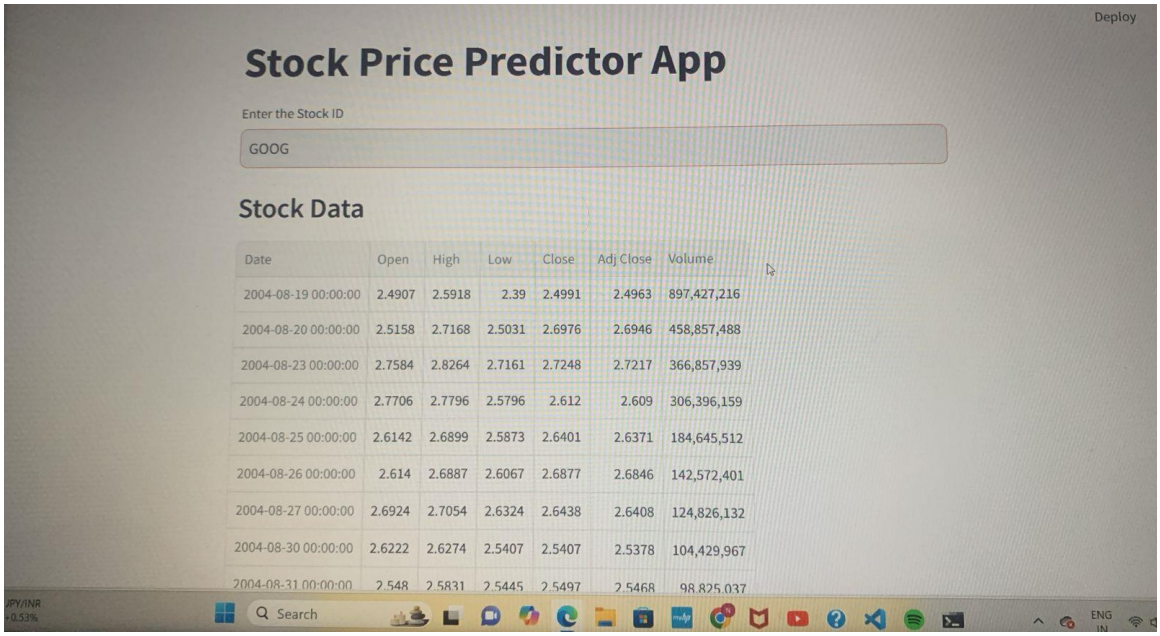


Fig -9.1




Deploy
Stock Price Predictor App
Enter the Stock ID
GOOG
Stock Data
Date Open High Low Close Adj Close Volume
2004-08-19 00:00:00 2.4907 2.5918 2.39 2.4991 2.4963 897,427,216
2004-08-20 00:00:00 2.5158 2.7168 2.5031 2.6976 2.6946 458,857,488
2004-08-23 00:00:00 2.7584 2.8264 2.7161 2.7248 2.7217 366,857,939
2004-08-24 00:00:00 2.7706 2.7796 2.5796 2.612 2.609 306,396,159
2004-08-25 00:00:00 2.6142 2.6899 2.5873 2.6401 2.6371 184,645,512
2004-08-26 00:00:00 2.614 2.6887 2.6067 2.6877 2.6846 142,572,401
2004-08-27 00:00:00 2.6924 2.7054 2.6324 2.6438 2.6408 124,826,132
2004-08-30 00:00:00 2.6222 2.6274 2.5407 2.5407 2.5378 104,429,967
2004-08-31 00:00:00 2.548 2.5831 2.5445 2.5497 2.5468 98,825,037
Search
ENG IN


Deploy
2004-08-31 00:00:00 2.548 2.5831 2.5445 2.5497 2.5468 98,825,037
2004-09-01 00:00:00 2.5579 2.5646 2.4824 2.4969 2.4941 183,633,734
Original Close Price and MA for 250 days
200
175
150
125
100
75
50
25
0
2004
2008
2012
2016
2020
2024
Original Close Price and MA for 200 days
JPY/INR +0.53%
Search
ENG IN

CONCLUSION

Predicting the stock market was a time-consuming and laborious procedure a few years or even a decade ago. However, with the application of machine learning for stock market forecasts, the procedure has become much simpler.

With the machine learning algorithms LSTM, SVR, RNN, KNN,GRU and LSTM with GRU on the finance dataset. All the techniques have shown an improvement in the accuracy of predictions, there by yielding positive results.

Use of recently introduced machine learning techniques in the prediction of stocks have yielded promising results and thereby marked the use of them in profitable exchange schemes. It has led to the conclusion that it is possible to predict stock market with more accuracy and efficiency using machine learning techniques.

CHAPTER 11

FUTURE ENHANCEMENT

Further adding Sentiment analysis involves analyzing news articles, social media posts, and other textual data to gauge market sentiment and investor emotions. By assessing the positive or negative sentiment surrounding a stock, sentiment analysis models attempt to predict the impact of public opinion on stock prices.

Developing ANDROID application and IOS application.

CHAPTER 12

REFERENCES

- Usmani, S. H. Adil, K. Raza and S. S. A. Ali, "Stock market prediction using machine learning techniques," 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, 2016, pp. 322-327.
- K. Raza, "Prediction of Stock Market performance by using machine learning techniques," 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, 2017, pp 1-55
- H. Gunduz, Z. Cataltepe and Y. Yaslan, "Stock market direction prediction using deep neural networks," 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, 2017, pp. 1-4.
- Faria, G., & Verona, F. (2018). Forecasting stock market returns by summing the frequency-decomposed parts. Journal of Empirical Finance,pp. 45
- Goyal, A., & Welch, I. (2003). Predicting the equity premium with dividend ratios. Management Science, 49, 639-654.
- Goyal, A., & Welch, 1. (2008). A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies, 21, 1455-1508.
- Guo, H. (2006). Time-varying risk premia and the cross section of stock returns. Journal of Banking & Finance, 30, 2087-2107.
- Haven, E., Liu, X., & Shen, L. (2012). De-noising option prices with the wavelet method. European Journal of Operational Research, 222(1), 104-112.
- Inoue, A., & Kilian, L. (2004). In-sample or out-of-sample tests of predictability: Which one should we use? Economic Review, 23, 371-402