Project report on

Stox Akashvani : Stock Market Prediction Website

by

Jaikishan Bagul [2020510001]

Tanish Gangwal [2020510020]

Prajwal Warkhade [2020510069]

Under the guidance of

Internal Supervisor

Prof. Nikhita Mangaonkar

Department of Master in Computer Applications Sardar Patel Institute of Technology

University of Mumbai

2021 – 22

PROJECT APPROVAL CERTIFICATE

This is to certify that the following students

Jaikishan Bagul [2020510001]

Tanish Gangwal [2020510020]

Prajwal Warkhade [2020510069]

Have satisfactorily carried out work on the project entitled

"**Stox Akashvani : Stock Market Prediction Website**"

Towards the fulfillment of mini project, as laid down by University of Mumbai during year 2021-22.

Project Guide H.O.D. Principal (Prof. Nikhita Mangaonkar) (Dr. Pooja Raundale) (Dr. B.N. Chaudhari)

PROJECT EVALUATION CERTIFICATE

This is to certify that the following students

Jaikishan Bagul [2020510001]

Tanish Gangwal [2020510020]

Prajwal Warkhade [2020510069]

Have successfully completed the Project report on **"Stox Akashvani : Stock Market Prediction Website"**, which is found to be satisfactory and is approved

At

SARDAR PATEL INSTITUTE OF TECHNOLOGY,
ANDHERI (W), MUMBAI.

INTERNAL EXAMINER EXTERNAL EXAMINER

Table of Contents

ABSTRACT

Stox Akashvani is a stock market prediction website. Stock market prediction can be instrumental in determining the future value of a company's stock. It is imperative to say that a successful prediction of a stock's future price could yield significant profit which would be beneficial for those who invested in the pipeline of things including stock market prediction.The model uses the information pertaining to the stocks and contemplates the previous model accuracy to innovate the approach used in our project. The experimental evaluation is based on the historical data set of National Stock Exchange (NSE). The proposed approach aims to provide models like Stacked LSTM and ARIMA which perform better than its contemporaries which have been achieved to a certain extent.

Objectives

- To build a user friendly and informative website.
- To provide information about the company's stocks of NSE.
- To provide trends, weekly, yearly sentiments of stocks.
- To create a indicators for Buy and Sell.
- To create a graph for projected stock market analysis.
- To provide information of American stock market as well such as S&P 500, NASDAQ, etc.
- To provide information of Crypto Currencies such Bitcoin, Etherum, etc. ● To create a news window for latest news in stock market.

List of figures

1.Introduction

1.1 Problem Definition:

In the field of quantitative trading, predicting the future security returns lies in the center of the industry, as the future trading strategy is always deployed and created based on our view of the financial market in the future. The trading area has two main different methods, namely fundamental analysis and quantitative trading. Fundamental method makes the trading decision based on the subjective view on the industry or company's future direction, it mainly relies on the public information such as market news, corporate statistics as well as a series of financial statement releases. On the other hand, the quantitative trading strategy uses mathematical models to make the decision hence avoids the interruption of human subjectivity and emotion.

1.2 Objective and Scope:

The main goal of this project is to develop website where a user see their stock's graph and get the sentiments of the stock. User also get the fundamental and quantitative analysis of stock. The project can be mainly divided into two parts:the first part of the prediction is that predict the weekly analysis and second part is predict yearly analysis. We have additional modules i.e., news of latest trend in stock market, company's information, etc.

1.3 System Requirements:

- Smartphone with Minimum Android version 4.4
- Any Web Browser
- Internet connectivity
- i3 Processor Based Computer or higher
- Memory: 4 GB RAM or more
- Hard Drive: 500 GB or more

2.Tools and Techniques

2.1 Tools used:

- Visual Studio
- Google Colab

2.2 Techniques used:

- Python
- CSS/HTML/JS

2.3 Project management plans:

2.3.1 Tasks

The major tasks are:

- Create a good user interface.

- Show predicted data from the analysis of fundamental analysis and quantitative trading.

2.3.2 Information Gathering

During information gathering about personality test it was observed that it will be a only depends on personality test. According to Erikson’stheory (stage 5:- Identity VS Confusion), identitydevelopment primarily relates to career identity, which is mainly developed during adolescence. Therefore I referred many theories for analysis for caree and select best test set for this.

3.Project Analysis and Design

3.1 Methodologies Accepted:

It would be better if we used the incremental model since the career analysis test can be easily

Project Management is the process of scoping, planning, organizing and controlling the development of an acceptable system at a minimum cost and within a specified time frame.

Choosing a model: Among the various popular models for developing a model, we have chosen Spiral model.

The *spiral model* combines the idea of iterative development with the systematic, controlled aspects of the waterfall model. Spiral model is a combination of iterative development process model and sequential linear development model i.e. waterfall model with very high emphasis on risk analysis. It allows for incremental releases of the product, or incremental refinement througheach iteration around the spiral.

The spiral model has four phases. A software project repeatedly passes through these phasesin iterations called Spirals.

Identification: This phase starts with gathering the business requirements in the baseline spiral. In the subsequent spirals as the product matures, identification of system requirements, subsystemrequirements and unit requirements are all done in this phase. This also includes understanding thesystem requirements by continuous communication between the customer and the system analyst.At the end of the spiral the product is deployed in the identified market.

Design: Design phase starts with the conceptual design in the baseline spiral and involves architectural design, logical design of modules, physical product design and final design in the subsequent spirals.

Construct or Build: Construct phase refers to production of the actual software product at everyspiral. In the baseline spiral when the product is just thought of and the design is being

developeda POC (Proof of Concept) is developed in this phase to get customer feedback. Then in the subsequent spirals with higher clarity on requirements and design details a working model of the software called build is produced with a version number. These builds are sent to customer for feedback.

Evaluation and Risk Analysis: Risk Analysis includes identifying, estimating, and monitoring technical feasibility and management risks, such as schedule slippage and cost overrun. After testing the build, at the end of first iteration, the customer evaluates the software and provides feedback.

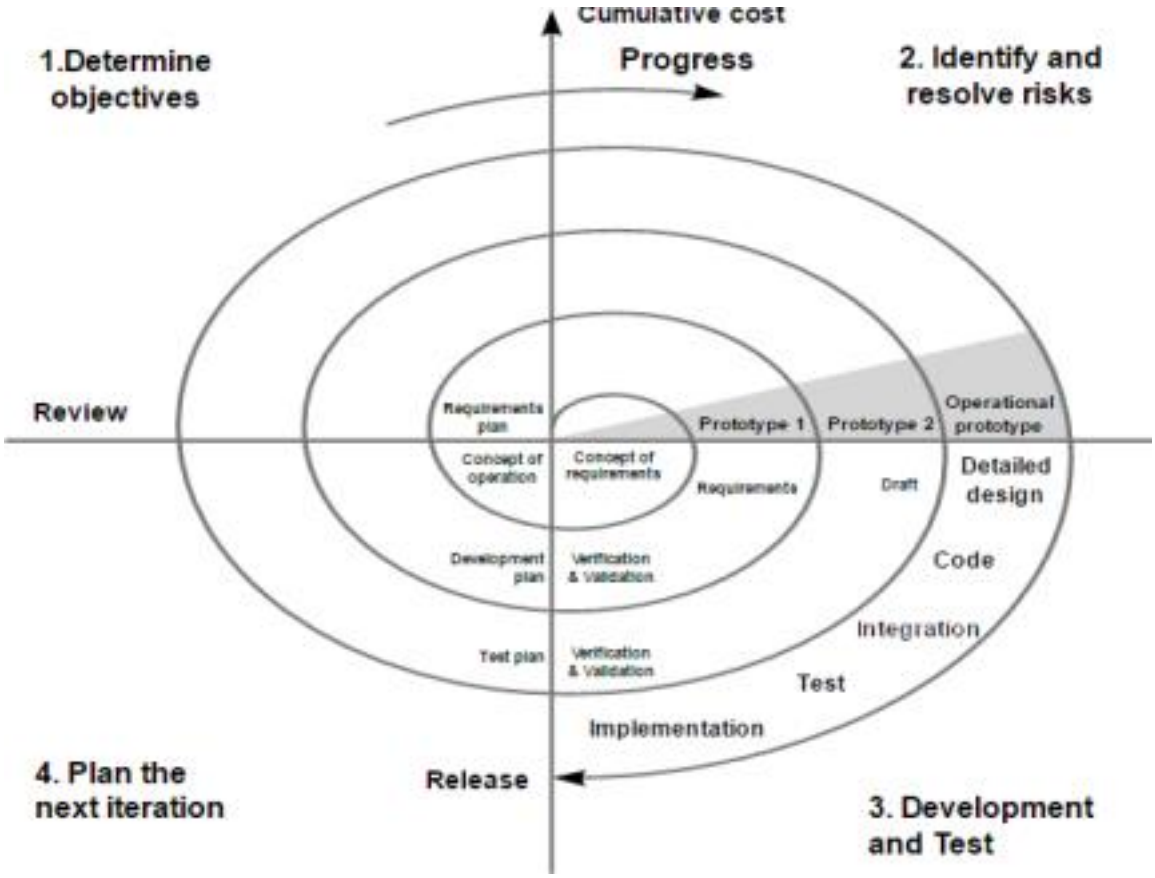

Fig 3.1 : Spiral Model

Advantages of spiral model:

- Changing requirements can be accommodated.

- Allows for extensive use of prototypes.
- Requirements can be captured more accurately.
- Users see the system early.
- Development can be divided into smaller parts and more risky parts can be developed earlier which helps better risk management.

When to use the spiral model:

- When costs there is a budget constraint and risk evaluation is important.
- For medium to high-risk projects.
- Customer is not sure of their requirements which is usually the case.
- Requirements are complex and need evaluation to get clarity.
- New product line which may be released in phases to get enough customer feedback.
- Significant changes are expected in the product during the development cycle.

3.2 UML Diagram:

3.2.1 Use case diagram

Use case diagrams are a set of use cases, actors and their relationships. They represent the use case view of a system. A use case represents a particular functionality of a system. So a use case diagram is used to describe the relationships among the functionalities and their internal/external controllers. These controllers are known as actors.

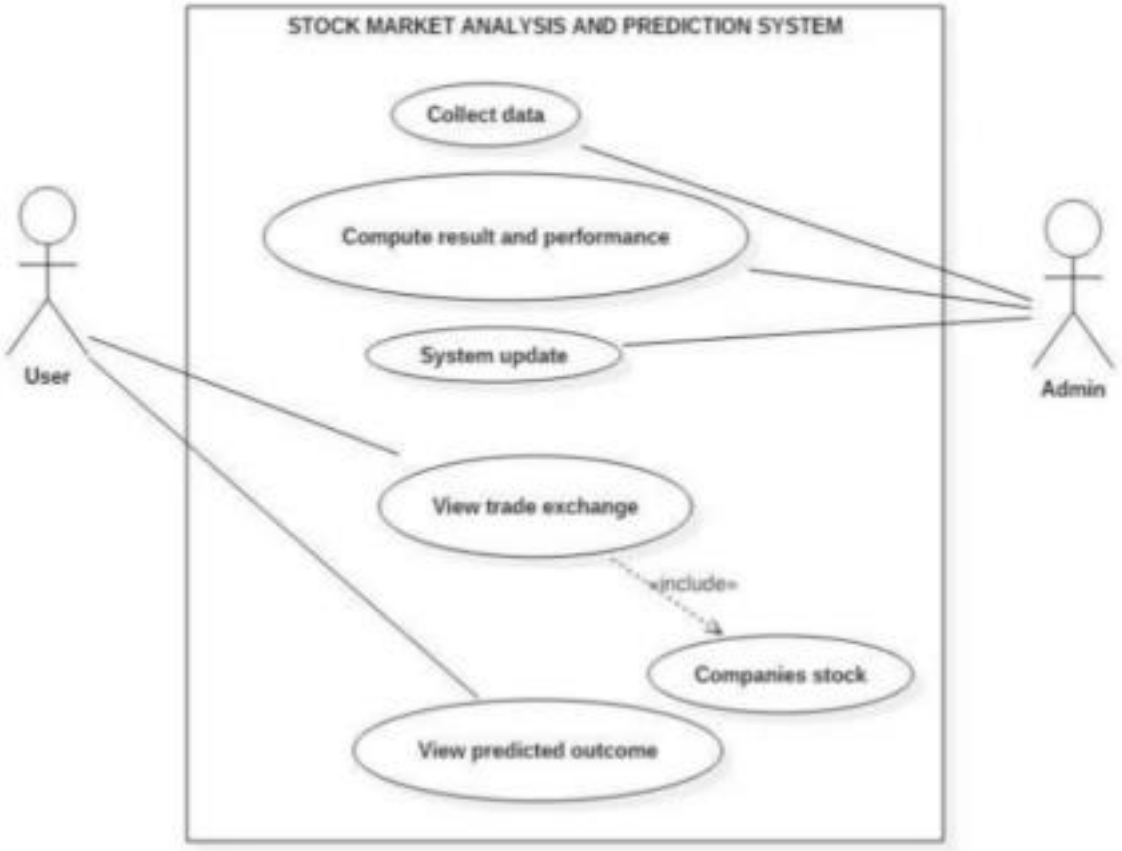

Fig 3.2.1(a). Use case diagram

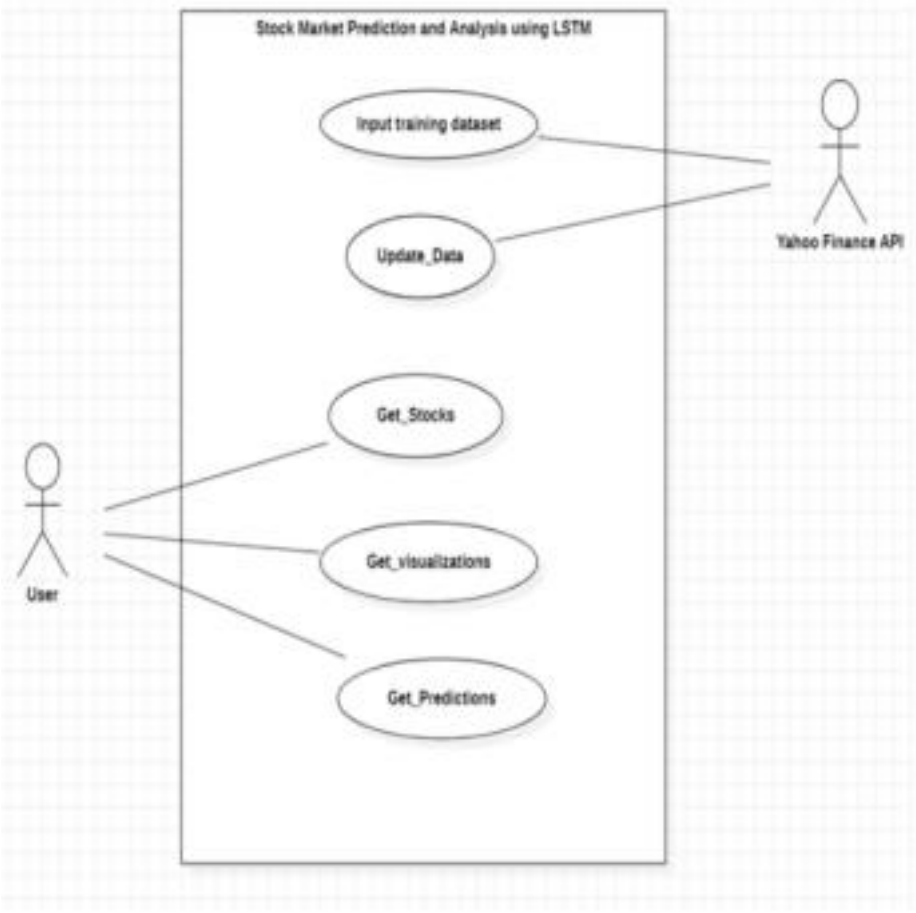

Fig 3.2.1(b). Use case diagram

Object diagram for stock market prediction and analysis using LSTM

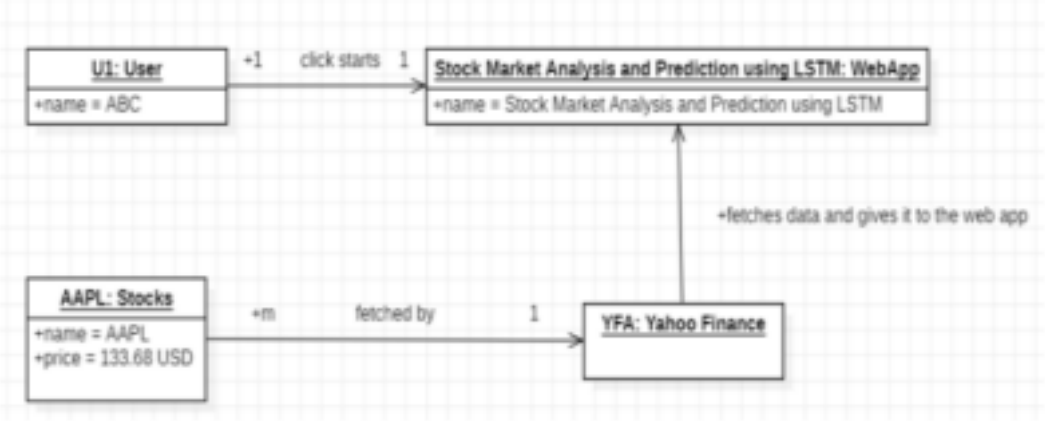

Fig3.2.1(c). Object diagram

3.2.2 Deployment diagram

Deployment diagrams are used to visualize the topology of the physical components of a

system where the software components are deployed. Deployment diagrams are used for describing the hardware components where software components are deployed. Componentdiagrams and deployment diagrams are closely related. Component diagrams are used to describe the components and deployment diagrams shows how they are deployed in hardware.

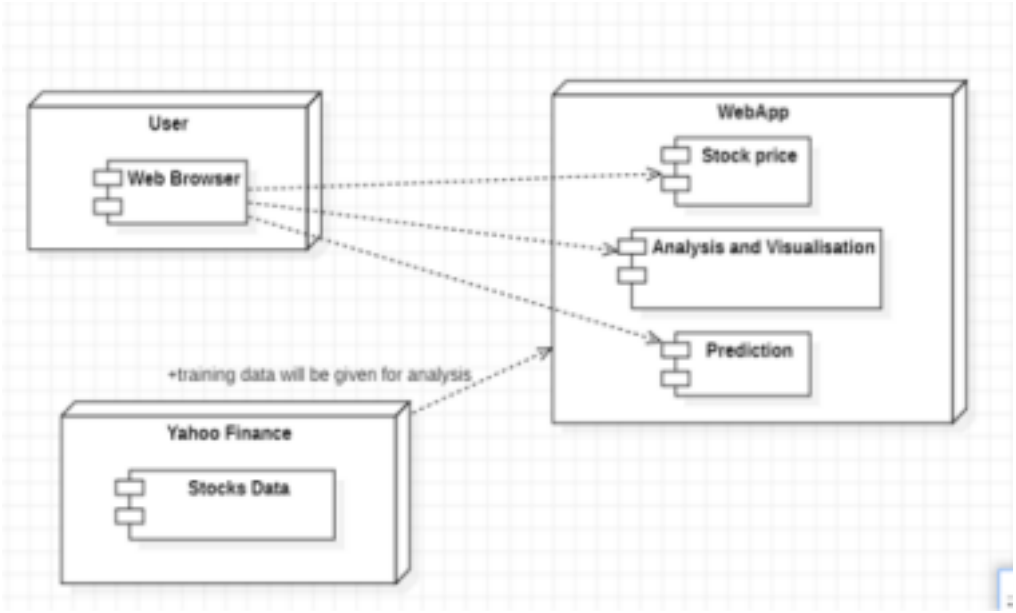

Fig 3.2.2. Deployment diagram

3.2.3 Component diagram

Component diagrams are different in terms of nature and behavior. Component diagrams areused to model physical aspects of a system. Component diagrams are used to visualize the organization and relationships among components in a system. These diagrams are also used to make executable systems. It does not describe the functionality of the system but it describes the components used to make those functionalities. Component diagrams can also be described as a static implementation view of a system. Static implementation represents the organization of the components at a particular moment.

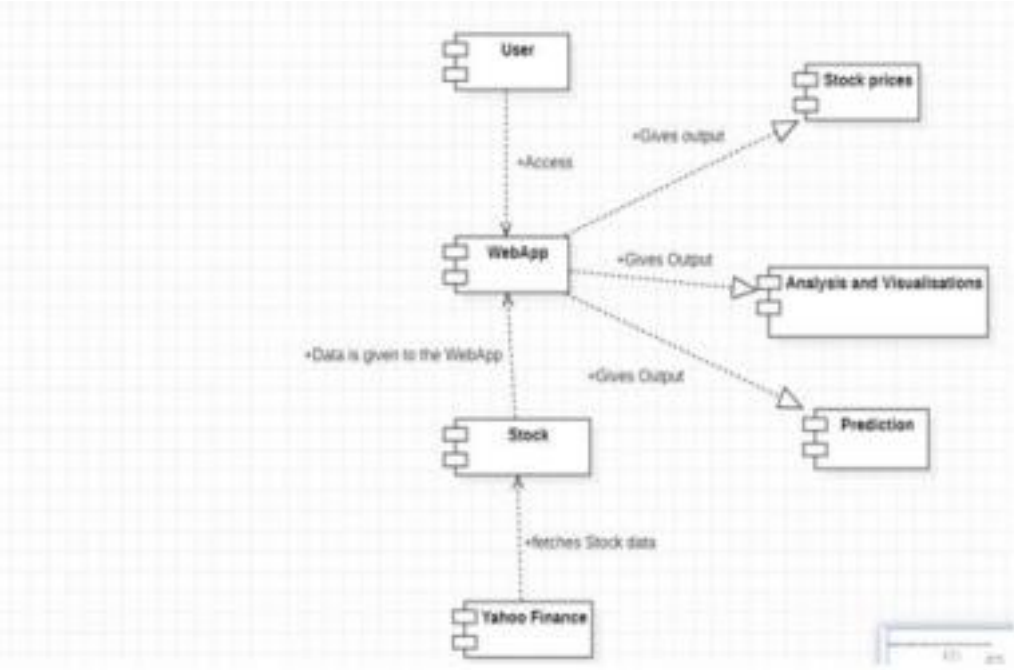

Fig 3.2.3. Component diagram

3.2.4 Activity diagram

Activity diagram is another important diagram in UML to describe dynamic aspects of the system. Activity is a particular operation of the system. Activity diagrams are not only used for visualizing dynamic nature of a system but they are also used to construct the executable system by using forward and reverse engineering techniques. The only missing thing in activity diagram is the message part. It does not show any message flow from one activity to another. Activity diagram is some time considered as the flow chart. Although the diagrams looks like a flow chart but it is not. It shows different flow like parallel, branched, concurrent and single.

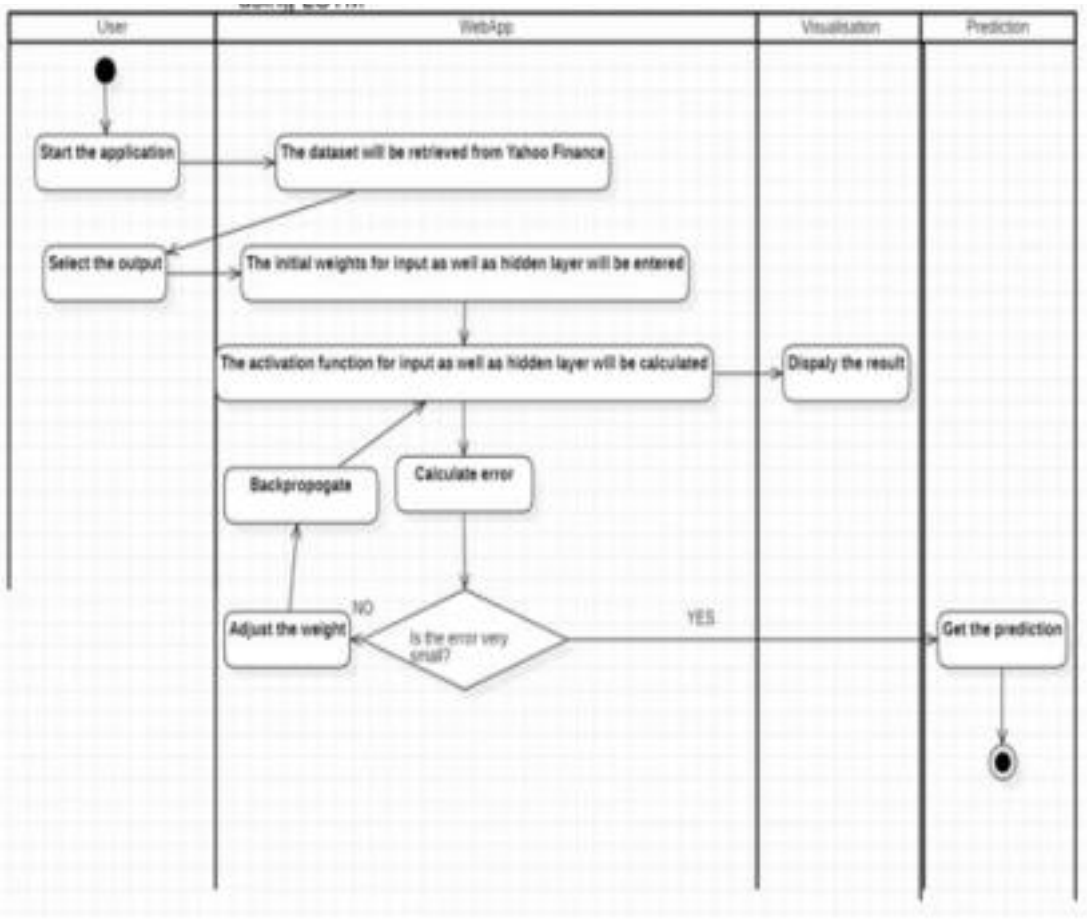

Fig 3.2.4. Activity diagram

3.3 Gantt chart Diagrams:

Gantt chart is a type of a bar chart that is used for illustrating project schedules. Gantt chartscan be used in any projects that involve effort, resources, milestones and deliveries. Gantt charts allow project managers to track the progress of the entire project. Through Gantt charts, the project manager can keep a track of the individual tasks as well as of the overall project progression. Gantt charts can be successfully used in projects of any scale. When using Gantt charts for large projects, there can be an increased complexity when tracking the tasks.

Name	Begin date	End date	Duration	ID	Predecessors
Requirement Gathering (...	10/3/22	16/3/22	5	0	
Questionnaire	10/3/22	14/3/22	3	2	
Research	14/3/22	16/3/22	3	9	
Analysis (PA)	17/3/22	23/3/22	5	6	0
Analysing current pro...	17/3/22	21/3/22	3	12	9
Feasability Check	21/3/22	23/3/22	3	14	
Design (PA)	24/3/22	14/4/22	16	19	6
Design database	24/3/22	28/3/22	4	21	9
Software Design	29/3/22	1/4/22	4	23	
Interface Design	1/4/22	6/4/22	4	25	
Create Design Interface	6/4/22	8/4/22	3	36	
Development (PA)	15/4/22	28/4/22	10	27	19
Develop System Mod...	15/4/22	20/4/22	4	28	22
Integrate System Mod...	20/4/22	25/4/22	4	29	
Perform Initial Testing	25/4/22	28/4/22	4	30	
Testing (PA)	29/4/22	11/5/22	9	36	27
Perform System Testing	29/4/22	3/5/22	3	38	

Fig 3.3(a): Gantt Chart(Part-1)

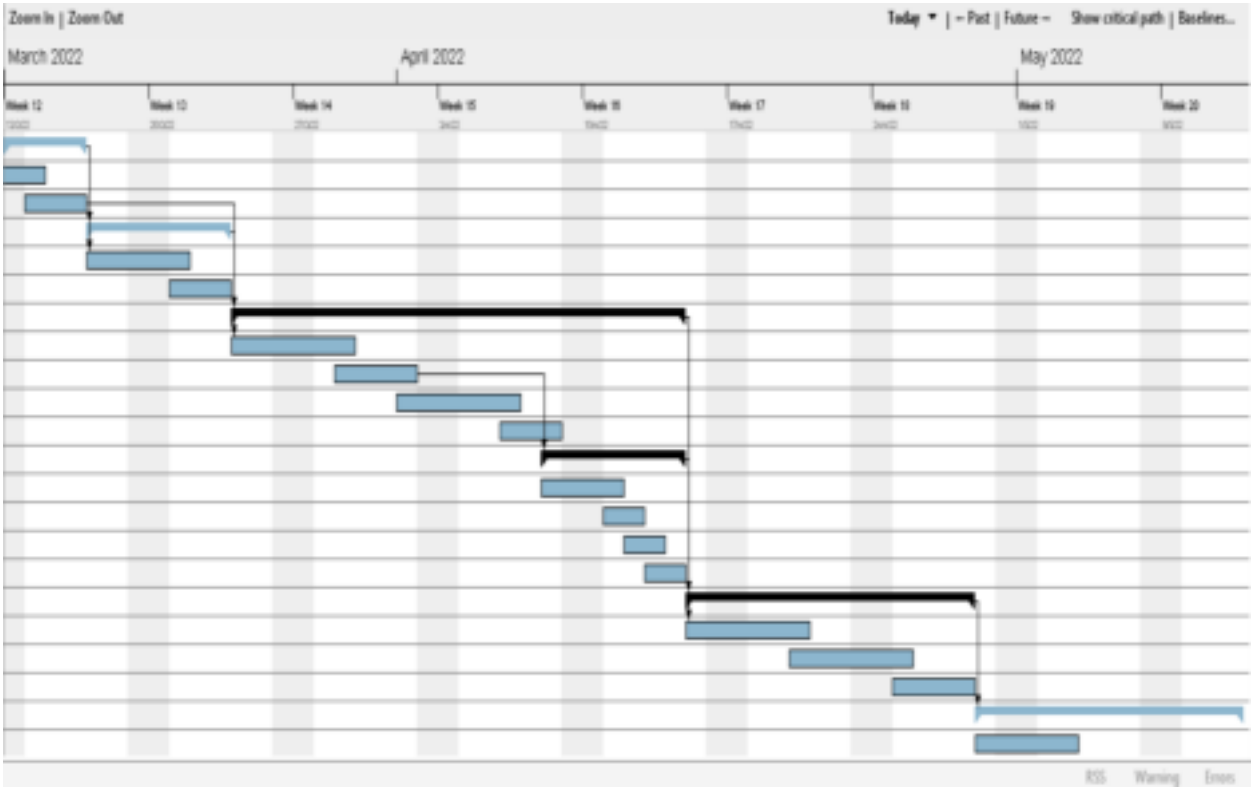

Fig 3.3(b): Gantt Chart(Part-2)

3.4 PERT chart

A PERT chart presents a graphic illustration of a project as a network diagram consisting of numbered nodes (either circles or rectangles) representing events, or milestones in the project linked by labelled vectors (directional lines) representing tasks in the project. The direction of thearrows on the lines indicates the sequence of tasks. The PERT chart is sometimes preferred over the Gantt chart, another popular project management charting method, because it clearly illustratestask dependencies. On the other hand, the PERT chart can be much more difficult to interpret, especially on complex projects.

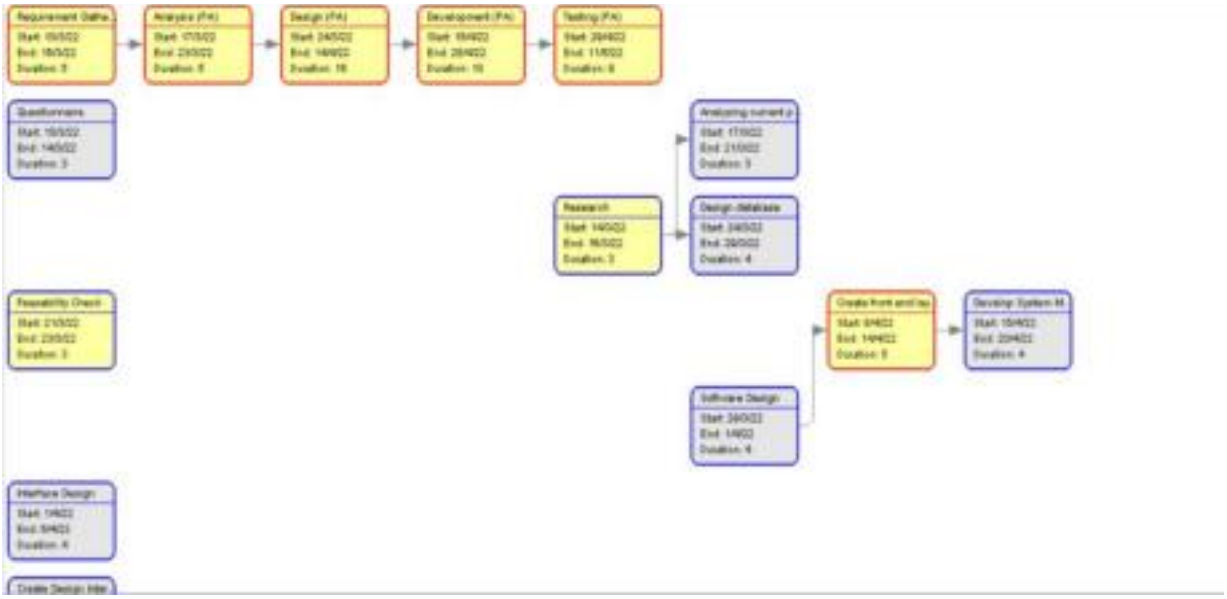

Fig 3.4(a): PERT Chart(Part-1)

Fig 3.4(b): PERT Chart (Part-2)

3.5 Code Implementation

```
%%writefile stockapp.py
import streamlit as st
from datetime import date

import yfinance as yf
```

```
from fbprophet import Prophet
from fbprophet.plot import plot_plotly
from plotly import graph_objs as go

START = "2015-01-01"
TODAY = date.today().strftime("%Y-%m-%d")

st.title('Stox Akashavani')

stocks = ('GOOG', 'AAPL', 'MSFT', 'GME','FB','TSLA','TATAMOTORS.NS','PAYTM
.NS','SCI','LICI.NS','ZOMATO.NS','IDEA.NS','TATASTEEL.NS','TATAPOWER.NS',
' TTML.NS','CDSL.NS','MOTHERSUMI.NS')
selected_stock = st.selectbox('Select dataset for prediction', stocks)

n_years = st.slider('Years of prediction:', 1, 4)
period = n_years * 365


@st.cache
def load_data(ticker):
    data = yf.download(ticker, START, TODAY)
    data.reset_index(inplace=True)
    return data


data_load_state = st.text('Loading data...')
data = load_data(selected_stock)
data_load_state.text('Loading data... done!')

st.subheader('Raw data')
st.write(data.tail())

# Plot raw data
def plot_raw_data():
    fig = go.Figure()
    fig.add_trace(go.Scatter(x=data['Date'], y=data['Open'],
name="stock_o pen"))
```



```
    fig.add_trace(go.Scatter(x=data['Date'], y=data['Close'],
name="stock_ close"))
    fig.layout.update(title_text='Time Series data with Rangeslider',
xaxi s_rangeslider_visible=True)
    st.plotly_chart(fig)

plot_raw_data()
```

```
# Predict forecast with Prophet.
df_train = data[['Date','Close']]
df_train = df_train.rename(columns={"Date": "ds", "Close": "y"})

m = Prophet()
m.fit(df_train)
future = m.make_future_dataframe(periods=period)
forecast = m.predict(future)

# Show and plot forecast
st.subheader('Forecast data')
st.write(forecast.tail())

st.write(f'Forecast plot for {n_years} years')
fig1 = plot_plotly(m, forecast)
st.plotly_chart(fig1)

st.write("Forecast components")
fig2 = m.plot_components(forecast)
st.write(fig2)
```


4.Project Implementation

4.1 Snapshot of application:

Fig.: 4.1.1: UI

Fig.: 4.1.2: UI

Fig.: 4.1.3: UI

Fig.: 4.1.4: UI

Fig.: 4.1.5: UI

Fig.: 4.1.6: UI

Fig.: 4.1.7: UI

Fig.: 4.1.8: UI

Fig.: 4.1.9: UI

Fig.: 4.1.10: UI

Fig.: 4.1.11: UI

Fig.: 4.1.12: UI

5. Literature Review

In the last few decades forecasting of stock returns has become an important field of research. In most of the cases the researchers had attempted to establish a linear relationship between the input macroeconomic variables and the stock returns. After the discovery of nonlinearity in the stock market index returns, many literatures have come up in nonlinear statistical modeling of the stock returns, most of them required that the nonlinear model be specified before the estimation is done. But since stock market return is noisy, uncertain, chaotic and nonlinear in nature, ANN has evolved out to be better technique in capturing the structural relationship between a stock's performance and its determinant factors more accurately than many other statistical techniques.

In literature, different sets of input variables are used to predict stock returns. In fact, different input variables are used to predict the same set of stock return data. Some researchers used input data from a single time series where others considered the inclusion of heterogeneous market information and macro-economic variables.

Banerjee(2014)The Indian stock exchange is that the center of interest for several economists, investors and researchers and thus it's quite necessary for them to possess a transparent understanding of the current standing of the market. the foremost reliable way to forecast the long run is to undertake to know the current and therefore, consequently the prior objective as the analysis of the current scenario of the Indian exchange therefore on perceive and check out to make a stronger future scope for investment. On this context, they need collected information on the monthly closing stock indices of sensex for 6 years (2007-2012). The analysis of the performance of the Indian exchange for 6 years with reference to time presents them a suitable time series ARIMA model (1,0, I) that helps them in predicting the approximate values of the longer term indices. Out of the initial six completely different models, they select ARIMA(I ,0, I) because the best model supported the actual fact that it

satisfies all the conditions for the goodness of work not like the remainder.

In LSTM, a conventional recurrent neural network, during the gradient back-spread stage, the gradient sign can end up being increased countless occasions by the weight network related with

the associations between the neurons of the recurrent secret layer. This implies that the extent of loads in the change framework can have a solid effect on the learning cycle.

6.Future enhancements

• Potential improvement can be made to our data collection and analysis method. • Future

research can be done with possible improvement such as more refined data and more accurate algorithm.

• Implementation of discussion forums and twitter sentiments.

7.Limitations

- Questions are limited.
- Its based on personality test theories and RAISEC test, So I can’t says it this is 100% perfect application for career analysis.

8.Conclusion

Stock market trades have gotten famous, urging scientists to discover predictions utilizing new advances or strategies. Appropriate predictive strategies can help analysts, investors and anybody

managing the stock market. To help predict the stock record, a less blunder of the predictive model is required which may consider the handling of the input information. Working with neural networks that are comprised of LSTM (Long-Short Term Memory Models) components has its limitations such as RNN to be very difficult to train even after adding LSTM (long/short memory). The main reason was the model's inability to decide which information to save or discard when the input stream grew larger and many other factors that contribute to the use of time embedding transformer architecture to predict Stocks.

9.Bibliography

- https://ieeexplore.ieee.org/document/6970973

• https://groww.in/
• https://upstox.com/
• https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9165760 •
https://ieeexplore.ieee.org/document/8593076/authors#authors •
https://towardsdatascience.com/lstm-for-google-stock-price-prediction-e35f5cc84165 •
https://neptune.ai/blog/predicting-stock-prices-using-machine-learning