

VISVESVARAYA TECHNOLOGICAL UNIVERSITY



BELAGAVI – 590018, Karnataka

INTERNSHIP REPORT

ON

“Stockport | Predictive Sentiment Analysis”

Submitted in partial fulfilment for the award of degree(21EC094)

BACHELOR OF ENGINEERING Electronics and Communication

Submitted by:

**NAVYA M
1SJ21EC094**



Conducted at
Compssoft Technologies



SJC INSTITUTE OF TECHNOLOGY
Department of Electronics and Communication

VTU Affiliated AICTE Approved

Accredited by NBA and NACC, New Delhi

P.B. No:20, B.B. Road, Chickballapur-562101

SJC INSTITUTE OF TECHNOLOGY
Department of Electronics and Communication
VTU Affiliated AICTE Approved
Accredited by NBA and NAAC, New Delhi
P.B. No:20, B.B. Road, Chickballapur-562101

CERTIFICATE

This is to certify that the Internship titled “**Stockport | Predictive Sentiment Analysis**” carried out by **Miss. Navya M**, a bonafide student of **SJC Institute of Technology**, in partial fulfillment for the award of **Bachelor of Engineering**, in **Electronics And Communication** under Visvesvaraya Technological University, Belagavi, during the year 2022-2023. It is certified that all corrections/suggestions indicated have been incorporated in the report.

The project report has been approved as it satisfies the academic requirements in respect of Internship prescribed for the course Internship / Professional Practice (21CSI85)

Signature of Guide

Signature of HOD

Signature of Principal

External Viva:

Name of the Examiner

Signature with Date

1) _____

2) _____

D E C L A R A T I O N

I, **Navya M**, Third year student of Electronics and Communication Engineering, SJC Institute of Technology - 562101, declare that the Internship has been successfully completed, in **COMPSOFT TECHNOLOGIES**. This report is submitted in partial fulfillment of the requirements for award of Bachelor Degree in Electronics and Communication, during the academic year 2022-2023.

Date: _____ :

Place: Chickaballapur

USN: 1SJ21EC094

NAME: Navya M

OFFER LETTER



Date: 26th October, 2023

Name: Navya M
USN: 1SJ21EC094
Placement ID: 23OCTMLBONE

Dear Student,

We would like to congratulate you on being selected for the **Machine Learning with Python (Research Based)** Internship position with **Compsoft Technologies**, effective Start Date **26th October, 2023**. All of us are excited about this opportunity provided to you!

This internship is viewed as being an educational opportunity for you, rather than a part-time job. As such, your internship will include training/orientation and focus primarily on learning and developing new skills and gaining a deeper understanding of concepts of **Machine Learning with Python (Research Based)** through hands-on application of the knowledge you learn while you train with the senior developers. You will be bound to follow the rules and regulations of the company during your internship duration.

Again, congratulations and we look forward to working with you!

Sincerely,

Nithin K. S
Project Manager
COMPSOFT TECHNOLOGIES
No. 363, 19th main road,
1st Block Rajajinagar
Bangalore - 560010

ACKNOWLEDGEMENT

This Internship is a result of accumulated guidance, direction and support of several important persons. We take this opportunity to express our gratitude to all who have helped us to complete the Internship.

We express our sincere thanks to our Principal, for providing usadequate facilities to undertake this Internship.

We would like to thank our Head of Dept – branch code, for providing us an opportunity to carry out Internship and for his valuable guidance and support.

We would like to thank our (Lab assistant name) Software Services for guiding us during the period of internship.

We express our deep and profound gratitude to our guide, Guide name, Assistant/Associate Prof, for her keen interest and encouragement at every step in completing the Internship.

We would like to thank all the faculty members of our department for the support extended during the course of Internship.

We would like to thank the non-teaching members of our dept, for helping us during the Internship.

Last but not the least, we would like to thank our parents and friends without whose constant help, the completion of Internship would have not been possible.

NAME: Navya M
USN :1SJ21EC094

ABSTRACT

The Stock market forecasters focus on developing a successful approach to predict stock prices. The vital idea to successful stock market prediction is not only achieving best results but also to minimize the inaccurate forecast of stock prices. This paper attempts to design and implement a predictive system for guiding stock market investment. The novelty of our approach is the combination of both sensex points and Really Simple Syndication (RSS) feeds for effective prediction. Our claim is that the sentiment analysis of RSS news feeds has an impact on stock market values. Hence RSS news feed data are collected along with the stock market investment data for a period of time. Using our algorithm for sentiment analysis, the correlation between the stock market values and sentiments in RSS news feeds are established. This trained model is used for prediction of stock market rates.

Table of Contents

Sl no	Description	Page no
1	Company Profile	
2	About the Company	
3	Introduction	
4	System Analysis	
5	Requirement Analysis	
6	Design Analysis	
7	Implementation	
8	Snapshots	
9	Conclusion	
10	References	

CHAPTER 1

COMPANY PROFILE

1. COMPANY PROFILE

A Brief History of Compsoft Technologies

Compsoft Technologies, was incorporated with a goal "To provide high quality and optimal Technological Solutions to business requirements of our clients". Every business is a different and has a unique business model and so are the technological requirements. They understand this and hence the solutions provided to these requirements are different as well. They focus on clients requirements and provide them with tailor made technological solutions. They also understand that Reach of their Product to its targeted market or the automation of the existing process into e-client and simple process are the key features that our clients desire from Technological Solution they are looking for and these are the features that we focus on while designing the solutions for their clients.

Sarvamoola Software Services. is a Technology Organization providing solutions for all web design and development, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Sarvamoola Software Services. specialize in ERP, Connectivity, SEO Services, Conference Management, effective web promotion and tailor-made software products, designing solutions best suiting clients requirements.

Compsoft Technologies, strive to be the front runner in creativity and innovation in software development through their well-researched expertise and establish it as an out of the box software development company in Bangalore, India. As a software development company, they translate this software development expertise into value for their customers through their professional solutions.

They understand that the best desired output can be achieved only by understanding the clients demand better. Compsoft Technologies work with their clients and help them to define their exact solution requirement. Sometimes even they wonder that they have completely redefined their solution or new application requirement during the brainstorming session, and here they position themselves as an IT solutions consulting group comprising of high caliber consultants.

They believe that Technology when used properly can help any business to scale and achieve new heights of success. It helps Improve its efficiency, profitability, reliability; to put it in one sentence " Technology helps you to Delight your Customers" and that is what we want to achieve.

CHAPTER 2

ABOUT THE COMPANY

2. ABOUT THE COMPANY



Compsoft Technologies is a Technology Organization providing solutions for all web design and development, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Compsoft Technologies specialize in ERP, Connectivity, SEO Services, Conference Management, effective webpromotion and tailor-made software products, designing solutions best suiting clients requirements. The organization where they have a right mix of professionals as a stakeholders to help us serve our clients with best of our capability and with at par industry standards. They have young, enthusiastic, passionate and creative Professionals to develop technological innovations in the field of Mobile technologies, Web applications as well as Business and Enterprise solution. Motto of our organization is to “Collaborate with our clients to provide them with best Technological solution hence creating Good Present and Better Future for our client which will bring a cascading a positive effect in their business shape as well”. Providing a Complete suite of technical solutions is not just our tag line, it is Our Vision for Our Clients and for Us, We strive hard to achieve it.

Products of Compsoft Technologies.

Android Apps

It is the process by which new applications are created for devices running the Android operating system. Applications are usually developed in Java (and/or Kotlin; or other such option) programming language using the Android software development kit (SDK), but other development environments are also available, some such as Kotlin support the exact same Android APIs (and bytecode), while others such as Go have restricted API access.

The Android software development kit includes a comprehensive set of development tools. These include a debugger, libraries, a handset emulator based on QEMU, documentation, sample code, and tutorials. Currently supported development platforms include computers running Linux (any modern desktop Linux distribution), Mac OS X 10.5.8 or later, and Windows 7 or later. As of March 2015, the SDK is not available on Android itself, but software development is possible by using specialized Android applications.

Web Application

It is a client–server computer program in which the client (including the user interface and client- side logic) runs in a web browser. Common web applications include web mail, online

retail sales, online auctions, wikis, instant messaging services and many other functions. web applications use web documents written in a standard format such as HTML and JavaScript, which are supported by a variety of web browsers. Web applications can be considered as a specific variant of client-server software where the client software is downloaded to the client machine when visiting the relevant web page, using standard procedures such as HTTP. The Client web software updates may happen each time the web page is visited. During the session, the web browser interprets and displays the pages, and acts as the universal client for any web application. The use of web application frameworks can often reduce the number of errors in a program, both by making the code simpler, and by allowing one team to concentrate on the framework while another focuses on a specified use case. In applications which are exposed to constant hacking attempts on the Internet, security-related problems can be caused by errors in the program.

Frameworks can also promote the use of best practices such as GET after POST. There are some who view a web application as a two-tier architecture. This can be a “smart” client that performs all the work and queries a “dumb” server, or a “dumb” client that relies on a “smart” server. The client would handle the presentation tier, the server would have the database (storage tier), and the business logic (application tier) would be on one of them or on both. While this increases the scalability of the applications and separates the display and the database, it still doesn’t allow for true specialization of layers, so most applications will outgrow this model. An emerging strategy for application software companies is to provide web access to software previously distributed as local applications. Depending on the type of application, it may require the development of an entirely different browser-based interface, or merely adapting an existing application to use different presentation technology. These programs allow the user to pay a monthly or yearly fee for use of a software application without having to install it on a local hard drive. A company which follows this strategy is known as an application service provider (ASP), and ASPs are currently receiving much attention in the software industry.

Security breaches on these kinds of applications are a major concern because it can involve both enterprise information and private customer data. Protecting these assets is an important part of any web application and there are some key operational areas that must be included in the development process. This includes processes for authentication, authorization, asset handling, input, and logging and auditing. Building security into the applications from the beginning can be more effective and less disruptive in the long run.

Web design

It encompasses many different skills and disciplines in the production and maintenance of websites. The different areas of web design include web graphic design; interface design; authoring, including standardized code and proprietary software; user experience design; and

search engine optimization. The term web design is normally used to describe the design process relating to the front-end (client side) design of a website including writing mark up. Web design partially overlaps web engineering in the broader scope of web development. Web designers are expected to have an awareness of usability and if their role involves creating mark up then they are also expected to be up to date with web accessibility guidelines. Web design partially overlaps web engineering in the broader scope of web development.

Departments and services offered

Compsoft Technologies plays an essential role as an institute, the level of education, development of student's skills are based on their trainers. If you do not have a good mentor then you may lag in many things from others and that is why we at Compsoft Technologies gives you the facility of skilled employees so that you do not feel unsecured about the academics. Personality development and academic status are some of those things which lie on mentor's hands. If you are trained well then you can do well in your future and knowing its importance of Compsoft Technologies always tries to give you the best.

They have a great team of skilled mentors who are always ready to direct their trainees in the best possible way they can and to ensure the skills of mentors we held many skill development programs as well so that each and every mentor can develop their own skills with the demands of the companies so that they can prepare a complete packaged trainee.

Services provided by Compsoft Technologies.

- Core Java and Advanced Java
- Web services and development
- Dot Net Framework
- Python
- Selenium Testing
- Conference / Event Management Service
- Academic Project Guidance
- On The Job Training
- Software Training

CHAPTER 3

INTRODUCTION

3. INTRODUCTION

Introduction to ML

Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

Machine learning is an important component of the growing field of data science. Through the use of statistical methods, algorithms are trained to make classifications or predictions, and to uncover key insights in data mining projects. These insights subsequently drive decision making within applications and businesses, ideally impacting key growth metrics. As big data continues to expand and grow, the market demand for data scientists will increase. They will be required to help identify the most relevant business questions and the data to answer them.

Machine learning algorithms are typically created using frameworks that accelerate solution development, such as TensorFlow and PyTorch. Machine learning tend to be used interchangeably, it's worth noting the nuances between the two.

Machine learning models fall into three primary categories.

- Supervised machine learning
- Un-Supervised machine learning
- Semi-Supervised machine learning

Problem Statement

Stockport | Predictive Sentiment Analysis. The Stock market forecasters focus on developing a successful approach to predict stock prices. The vital idea to successful stock market prediction is not only achieving best results but also to minimize the inaccurate forecast of stock prices. This paper attempts to design and implement a predictive system for guiding stock market investment. The novelty of our approach is the combination of both sensex points and Really Simple Syndication (RSS) feeds for effective prediction.

CHAPTER 4

SYSTEM ANALYSIS

4. SYSTEM ANALYSIS

1. Existing System

Real-Time Twitter Sentiment Analysis for stocks based on which the future movement of the market is predicted Goal: Understand the working of Sentiment analysis and Improve the accuracy.

2. Proposed System

The system explores the task of automatic identification of news opinions with the help of RSS news feeds and predicts the stock market movement whether goes up or down. The Fig. 2 describes the system for prediction and analysis of stock market using RSS news feed.

3. Objective of the System

In this process, it removes the incorrect, incomplete, improperly formatted, or duplicated data. Dirty data can cause confusion in the data set. Hence, this module cleans the data by filling missing values, smoothing the noisy data, identifying and removing the outliers. After preprocessing, the data are passed to the next module.

CHAPTER 5

REQUIREMENT ANALYSIS

5. REQUIREMENT ANALYSIS

Hardware Requirement Specification

- Operating System: Windows 10, Linux
- Processor: Intel(R) Core(TM) i5-5200U
- RAM: 8.00 GB
- Video Memory
- Device: Webcam

Software Requirement Specification

- IDE: Jupiter Notebook
- Frame work: Jupiter
- Database: MYSQL

CHAPTER 6

DESIGN ANALYSIS

6. DESIGN & ANALYSIS

```
import pandas as pd

from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix,
precision_recall_curve, roc_curve, auc
from nltk.corpus import stopwords
from nltk.stem import PorterStemmer
from sklearn.pipeline import make_pipeline
import matplotlib.pyplot as plt
import seaborn as sns

# Download NLTK resources
import nltk
nltk.download('stopwords')

# Larger dataset (replace this with your own dataset)
data = {
    'text': ['I love Stockport!', 'Stockport is a great place.', 'I dislike Stockport.', 'Stockport is not
good.'] * 1000,
    'label': ['positive', 'positive', 'negative', 'negative'] * 1000
}

df = pd.DataFrame(data)

# Preprocess the text data
stop_words = set(stopwords.words('english'))
stemmer = PorterStemmer()

def preprocess_text(text):
    words = [stemmer.stem(word) for word in text.split() if word.lower() not in stop_words]
```

```

return ' '.join(words)

df['processed_text'] = df['text'].apply(preprocess_text)

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(df['processed_text'], df['label'], test_size=0.2,
random_state=42)

# Build a predictive sentiment analysis model
model = make_pipeline(TfidfVectorizer(), MultinomialNB())
model.fit(X_train, y_train)

# Make predictions on the test set
y_pred = model.predict(X_test)

# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)
classification_rep = classification_report(y_test, y_pred)
conf_matrix = confusion_matrix(y_test, y_pred)

print(f'Accuracy: {accuracy}')
print('Classification Report:')
print(classification_rep)

# Plot confusion matrix
sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues', xticklabels=model.classes_,
yticklabels=model.classes_)
plt.xlabel('Predicted')
plt.ylabel('True')
plt.title('Confusion Matrix')
plt.show()

# Plot precision-recall curve
precision, recall, _ = precision_recall_curve(y_test, model.predict_proba(X_test)[:, 1],

```

```
pos_label='positive')
plt.plot(recall, precision, marker='.')
plt.xlabel('Recall')
plt.ylabel('Precision')
plt.title('Precision-Recall Curve')
plt.show()

# Plot ROC curve
fpr, tpr, _ = roc_curve(y_test, model.predict_proba(X_test)[:, 1], pos_label='positive')
roc_auc = auc(fpr, tpr)
plt.plot(fpr, tpr, color='darkorange', lw=2, label='ROC curve (area = {:.2f})'.format(roc_auc))
plt.plot([0, 1], [0, 1], color='navy', lw=2, linestyle='--')
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('Receiver Operating Characteristic (ROC) Curve')
plt.legend(loc="lower right")
plt.show()
```

CHAPTER 7

IMPLEMENTATION

7. IMPLEMENTATION

Implementation is the stage where the theoretical design is turned into a working system. The most crucial stage in achieving a new successful system and in giving confidence on the new system for the users that it will work efficiently and effectively.

The system can be implemented only after thorough testing is done and if it is found to work according to the specification. It involves careful planning, investigation of the current system and its constraints on implementation, design of methods to achieve the change over and an evaluation of change over methods as a part from planning.

Two major tasks of preparing the implementation are education and training of the users and testing of the system. The more complex the system being implemented, the more involved will be the system analysis and design effort required just for implementation.

The implementation phase comprises of several activities. The required hardware and software acquisition is carried out. The system may require some software to be developed. For this, programs are written and tested. The user then changes over to his new fully tested system and the old system is discontinued.

TESTING

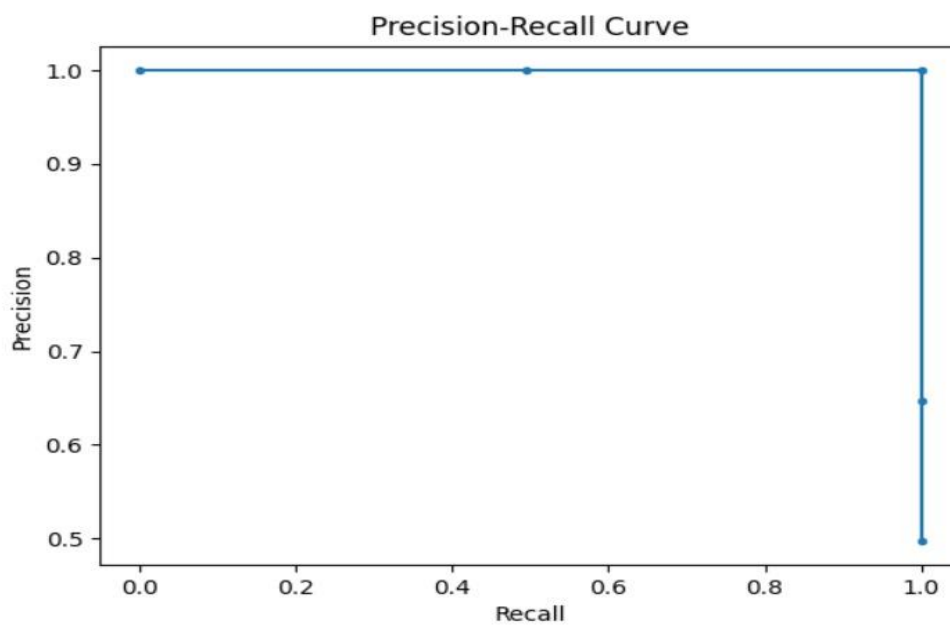
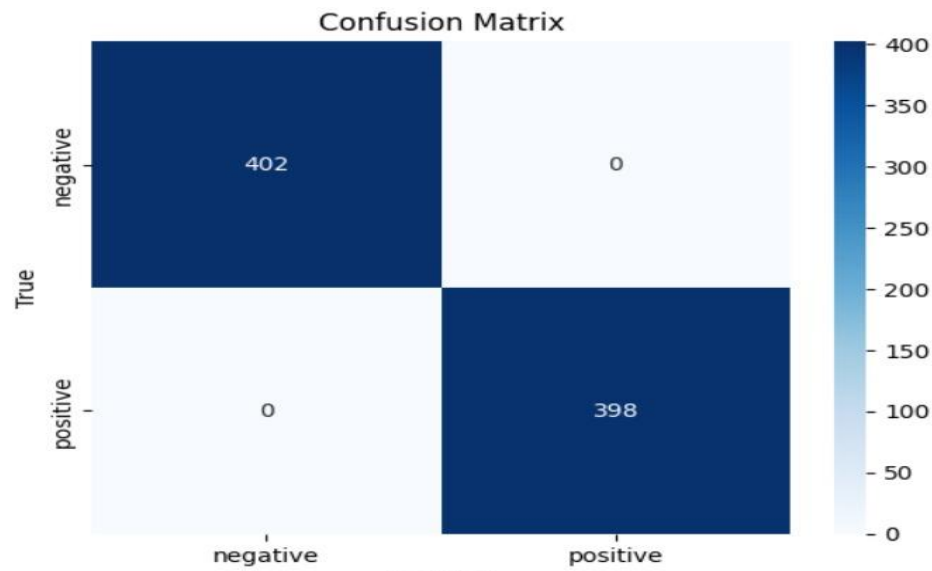
The testing phase is an important part of software development. It is the Information zed system will help in automate process of finding errors and missing operations and also a complete verification to determine whether the objectives are met and the user requirements are satisfied. Software testing is carried out in three steps:

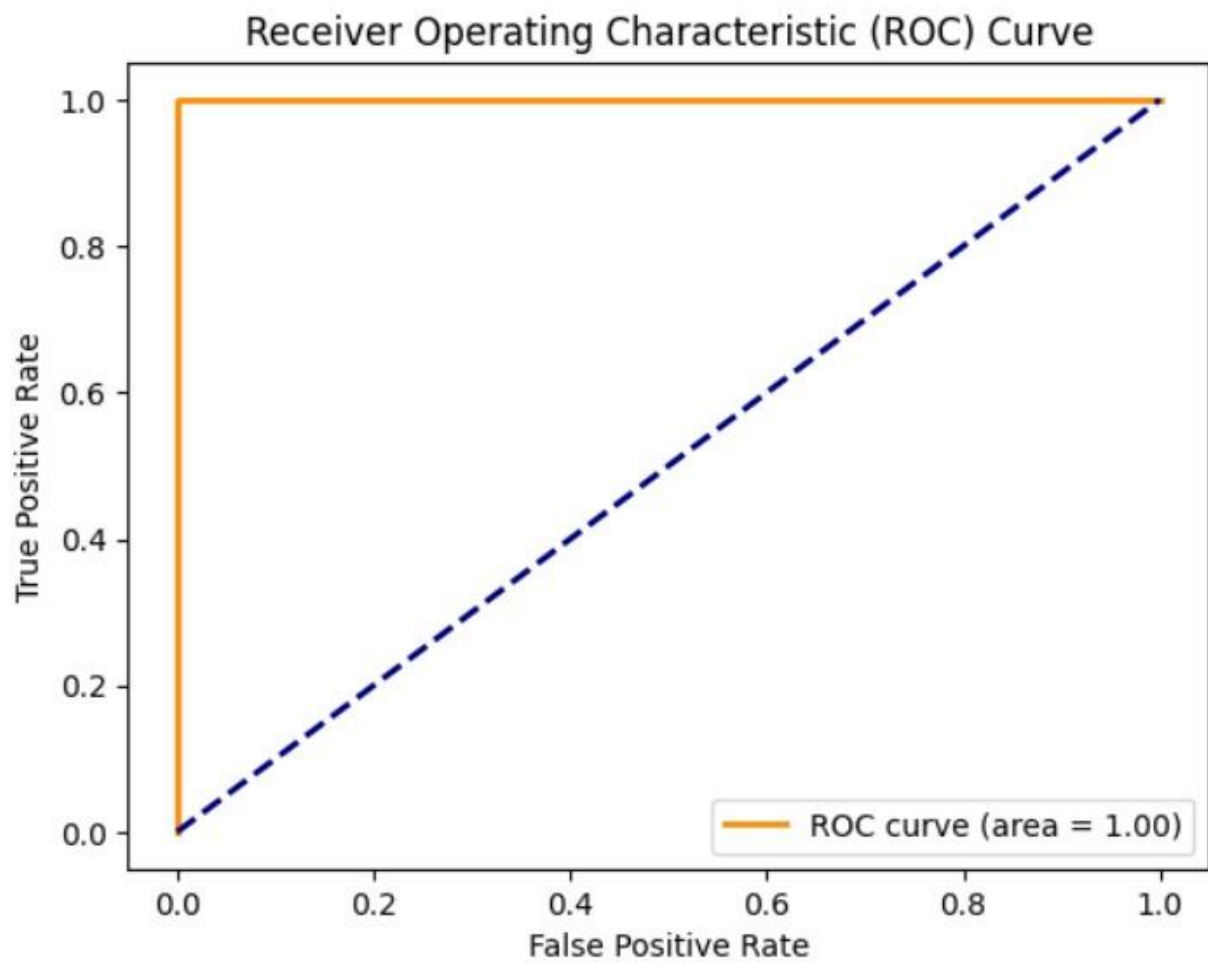
1. The first includes unit testing, where in each module is tested to provide its correctness, validity and also determine any missing operations and to verify whether the objectives have been met. Errors are noted down and corrected immediately.
2. Unit testing is the important and major part of the project. So errors are rectified easily in particular module and program clarity is increased. In this project entire system is divided into several modules and is developed individually. So unit testing is conducted to individual modules.
3. The second step includes Integration testing. It need not be the case, the software whose modules when run individually and showing perfect results, will also show perfect results when run as a whole.

CHAPTER 8

SNAPSHOTS

8. SNAPSHOTS





```
[nltk_data] Package stopwords is already up-to-date!
```

```
Accuracy: 1.0
```

```
Classification Report:
```

	precision	recall	f1-score	support
negative	1.00	1.00	1.00	402
positive	1.00	1.00	1.00	398
accuracy			1.00	800
macro avg	1.00	1.00	1.00	800
weighted avg	1.00	1.00	1.00	800

CHAPTER 9

CONCLUTION

9. CONCLUSION

The package was designed in such a way that future modifications can be done easily. The following conclusions can be deduced from the development of the project:

- ❖ Automation of the entire system improves the efficiency
- ❖ It provides a friendly graphical user interface which proves to be better when compared to the existing system.
- ❖ It gives appropriate access to the authorized users depending on their permissions.
- ❖ It effectively overcomes the delay in communications.
- ❖ Updating of information becomes so easier
- ❖ System security, data security and reliability are the striking features.
- ❖ The System has adequate scope for modification in future if it is necessary.

10. REFERENCE

- [1] A. Balahur, Mijail Kabadjov, J. Steinberger, R. Steinberger, A. Montoyo, “Challenges and Solutions in the Opinion Summarization”, *Journal of Intelligent Information Systems*, Springer, Vol.39, No.2, pp. 375-398, 2012.
- [2] Q. Li, Y. Chen, L. L. Jiang and P. Li and H. Chen, “A Tensor-Based Information Framework for Predicting the Stock Market”, *ACM Transactions on Information Systems*, Vol.34, No.2, pp.1-30, 2016.
- [3] D. Nguyen and M. Le, “A Two-Stage Architecture for Stock Price forecasting by combining SOM and fuzzy-SVM”, *International Journal of Computer Science and Information Security*, Vol. 12, No. 8, pp. 1-6, 2014.
- [4] B. Johan, H. Mao and X. Zeng, “Twitter Mood Predicts The Stock Market”, *Journal of Computational Science*, Vol.2, No.1, pp.1-8, 2011.
- [5] H. Kunhuang & T.H.K Yu, “The Application of Neural Networks to Forecast Fuzzy Time Series”, *Physical A: Statistical Mechanics and Its Applications*, Vol. 363, No.2, pp. 481-491, 2006.
- [6] M. Karamibekr, A. A. Ghorbani, “Sentence Subjectivity Analysis in Social Domains” In: *Proc. of IEEE/WIC/ACM International Conf. On Web Intelligence (WI) and Intelligent Agent Technology (IAT)*, Vol. 3, No. 1, pp.268- 275, 2013.
- [7] N. L. D. Khoa, K. Sakakibara and I. Nishikawa, “Stock Price Forecasting using Back Propagation Neural Networks with Time and Profit Based Adjusted Weight Factors”, In: *Proc. of SICE-ICASE International Joint Conf. in Bexco, Busan, Korea*, Vol.1, No.1, pp. 5484 - 5488, 2006.
- [8] D. Yan, G. Zhou, X. Zhao, Y. Tian and F. Yang, “Predicting Stock using Micro blog Moods”, *Journal on China Communications*, Vol. 13, No. 10, pp.244- 257, 2016.

