VISVESVARAYA TECHNOLOGICAL UNIVERSITY



BELAGAVI – 590018, Karnataka INTERNSHIP REPORT

ON

"Sentiment Analysis using ML"

Submitted in partial fulfilment for the award of degree(18CSI85)

BACHELOR OF ENGINEERING IN COMPUTER SCIENCE

Submitted by:

NAME	USN
KAUSARJAHAN A SHAIKH	2GP19CS013
MAHIMAPRIYA EDWIN LOBO	2GP19CS015
SHEJAL J GAONKAR	2GP19CS022





Department of Computer Science & Engineering

GOVERNMENT ENGINEERING COLLEGE MAJALI,

KARWAR-581345 DIST UTTAR KANNADA KARNATAKA

Department of Computer Science & Engineering GOVERNMENT ENGINEERING COLLEGE MAJALI, KARWAR-581345 DIST UTTAR KANNADA KARNATAKA



CERTIFICATE

This is to certify that the Internship titled "Sentiment Analysis using ML" carried out by Kausarjahan A Shaikh, a bonafide student of Government Engineering College Majali, in partial fulfillment for the award of Bachelor of Engineering, in Computer Science 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 (18CSI85)

Signature of Guide	Signature of HOD	Signature of Principal
	External Viva:	
Name of the Examiner		Signature with Date
1)		
2)		

DECLARATION

I, **KAUSARJAHAN A SHAIKH**, final year student of Computer Science Branch, Government Engineering College Karwar, declare that the internship has been successfully completed, in **COMPANY NAME**. This report is submitted in partial fulfillment of the requirements for award of Bachelor Degree in Branch name, during the academic year 2022-2023.

Date:

Place:

USN: 2GP19CS013

NAME: KAUSARJAHAN A SHAIKH

OFFER LETTER





Date: 2nd September, 2022

Name: Kausarjahan A Shaikh

USN: 2GP19CS013

Dear Student,

We would like to congratulate you on being selected for the **Machine Learning WithPython(Research Based)** Internship position with **Varcons Technologies Pvt Ltd**, effective Start Date 2nd **September, 2022**, 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 parttime 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,

Spoorthi H C **Director**Varcons Technologies Pvt Ltd
213, 2st Floor,
18 M G Road, Ulsoor,
Bangalore-560001

Department of Computer Science & Engineering GOVERNMENT ENGINEERING COLLEGE MAJALI,

KARWAR-581345 DIST UTTAR KANNADA KARNATAKA



CERTIFICATE

This is to certify that the Internship titled "Sentiment Analysis using ML" carried out by Mahimapriya Edwin Lobo, a bonafide student of Government Engineering College Majali, in partial fulfillment for the award of Bachelor of Engineering, in Computer Science 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 (18CSI85)

Signature of Guide	Signature of HOD	Signature of Principal
	External Viva:	
Name of the Examiner		Signature with Date
1)		
2)		

DECLARATION

I, MAHIMAPRIYA EDWIN LOBO, final year student of Computer Science Branch, Government Engineering College Karwar, declare that the Internship has been successfully completed, in COMPANY NAME. This report is submitted in partial fulfillment of the requirements for award of Bachelor Degree in Branch name, during the academic year 2022-2023.

Date::

Place:

USN: 2GP19CS015

NAME: MAHIMAPRIYA EDWIN LOBO

OFFER LETTER





Date: 2nd September, 2022

Name: Mahimapriya Edwin Lobo

USN: 2GP19CS015

Dear Student,

We would like to congratulate you on being selected for the **Machine Learning WithPython(Research Based)** Internship position with **Varcons Technologies Pvt Ltd**, effective Start Date **2**nd **September**, **2022**, 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 parttime 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,

Spoorthi H C **Director**Varcons Technologies Pvt Ltd
213, 2st Floor,
18 M G Road, Ulsoor,
Bangalore-560001

Department of Computer Science & Engineering GOVERNMENT ENGINEERING COLLEGE MAJALI, KARWAR-581345 DIST UTTAR KANNADA KARNATAKA



CERTIFICATE

This is to certify that the Internship titled "Sentiment Analysis using ML" carried out by Shejal J Gaonkar, a bonafide student of Government Engineering College Majali, in partial fulfillment for the award of Bachelor of Engineering, in Computer Science 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 (18CSI85)

Signature of Guide	Signature of HOD	Signature of Principal
	External Viva:	
Name of the Examiner		Signature with Date
1)		
2)		

DECLARATION

I, **SHEJAL J GAONKAR**, final year student of Computer Science Branch, Government Engineering College Karwar, declare that the Internship has been successfully completed, in COMPANY NAME. This report is submitted in partial fulfillment of the requirements for award of Bachelor Degree in Branch name, during the academic year 2022-2023.

Date::

Place:

USN: 2GP19CS022

NAME: SHEJAL J GAONKAR

OFFER LETTER





Date: 2nd September, 2022

Name: Shejal J Gaonkar

USN: **2GP19CS022**

Dear Student,

We would like to congratulate you on being selected for the **Machine Learning WithPython(Research Based)** Internship position with **Varcons Technologies Pvt Ltd**, effective Start Date **2**nd **September**, **2022**, 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 parttime 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,

Spoorthi H C **Director**Varcons Technologies Pvt Ltd
213, 2st Floor,
18 M G Road, Ulsoor,
Bangalore-560001

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 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, forhelping 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.

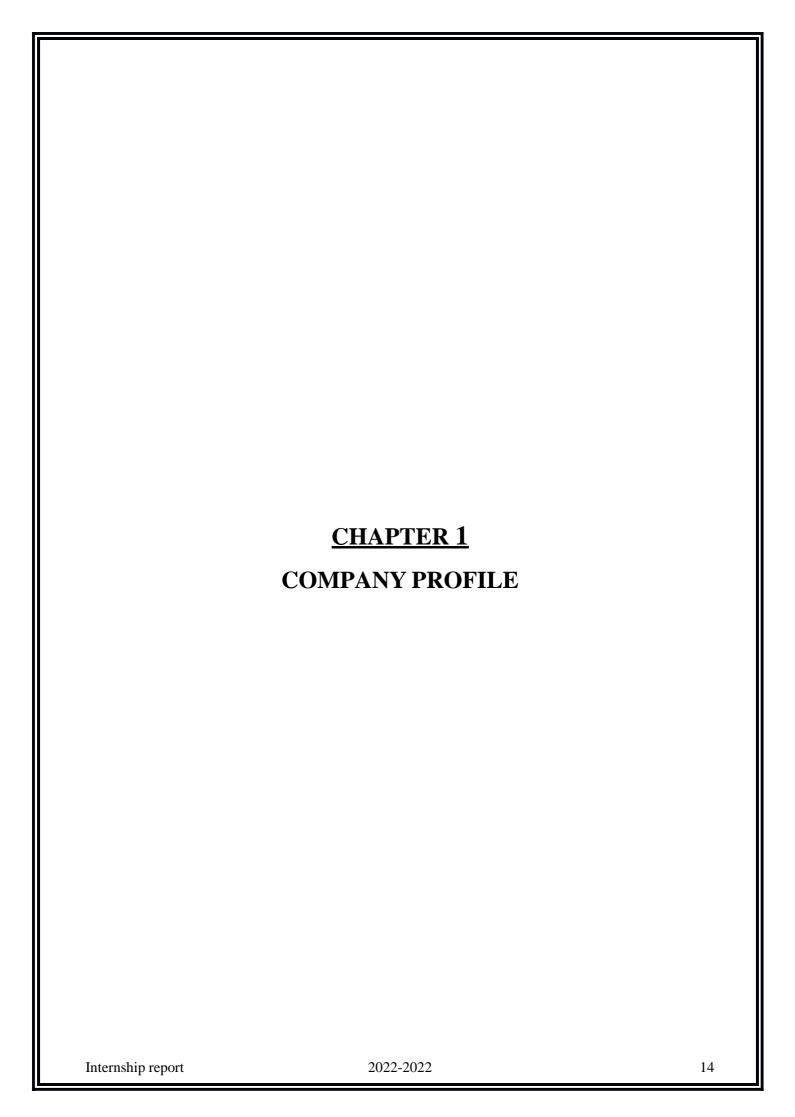
NAME USN
KAUSARJAHAN A SHAIKH 2GP19CS013
MAHIMAPRIYA EDWIN LOBO 2GP19CS015
SHEJAL J GAONKAR 2GP19CS022

ABSTRACT

Sentiment analysis or opinion mining is the computational study of people's opinions, sentiments, attitudes, and emotions expressed in written language. It is one of the most active research areas in natural language processing and text mining in recent years. Its popularity is mainly due to two reasons. First, it has a wide range of applications because opinions are central to almost all human activities and are key influencers of our behaviors. Whenever we need to make a decision, we want to hear others' opinions. Second, it presents many challenging research problems, which had never been attempted before the year 2000. Part of the reason for the lack of study before was that there was little opinionated text in digital forms. It is thus no surprise that the inception and the rapid growth of the field coincide with those of the social media on the Web. In fact, the research has also spread outside of computer science to management sciences and social sciences due to its importance to business and society as a whole.

Table of Contents

Sl no	Description	Page no
1	Company Profile	15
2	About the Company	17
3	Introduction	22
4	System Analysis	24
5	Requirement Analysis	25
6	Design Analysis	27
7	Implementation	30
8	Snapshots	33
9	Conclusion	38
10	References	39



1. COMPANY PROFILE

A Brief History of Varcons technologies pvt ltd

Varcons technologies pvt ltd, 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 eclient 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.

Varcons technologies pvt ltd, 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. Varcons technologies pvt ltd work with their clients and help them todefiine their exact solution requirement. Sometimes even they wonder that they have completely redefined their solution or new application requirement during the brainstormingsession, 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 itin one sentence "Technology helps you to Delight your Customers" and that is what we wantto achieve.

	CHAPTER 2 ABOUT THE COMPANY	
Internship report 2022-2022 16		

2. ABOUT THE COMPANY



Varcons technologies pvt ltd 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, Varcons technologies pvt ltd specialize in ERP, Connectivity, SEO Services, Conference Management, effective web promotion 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 Varcons technologies pvt ltd.

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 zutorials. 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 Internship report 2022-2022

softwaredevelopment 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 specifific 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 specifified 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 is 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

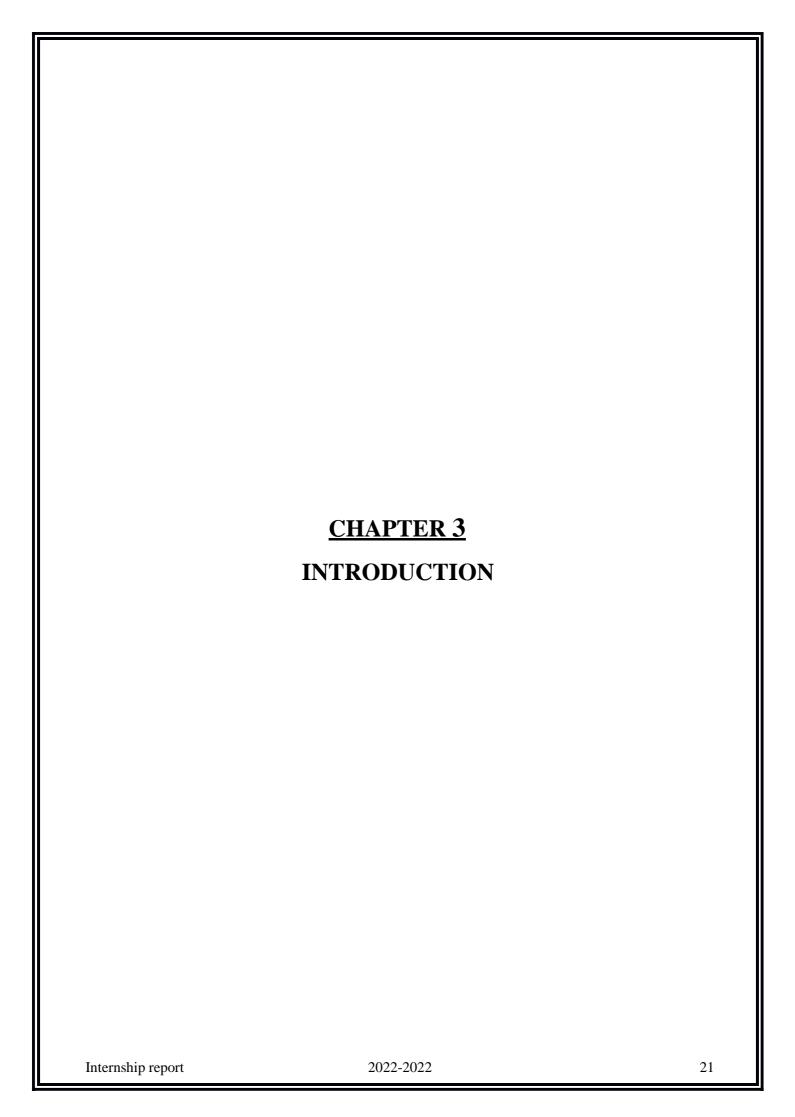
Varcons technologies pvt ltd plays an essential role as an institute, the level of education, development of student's skills are based on their trainers. If you do not have a good mentor then you may lag in many things from others and that is why we at Varcons technologies pvt ltd gives you the facility 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 Varcons technologies pvt ltd 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 Varcons technologies pvt ltd.

- 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



3. <u>INTRODUCTION</u>

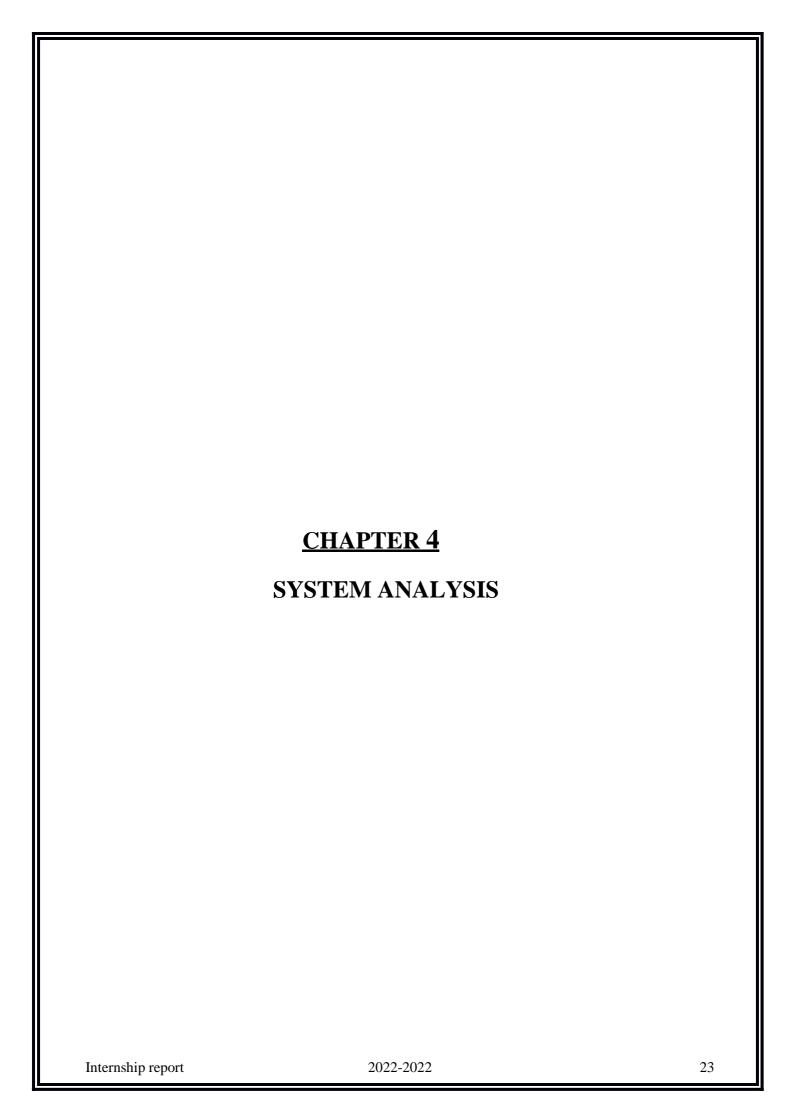
> Introduction to ML

Machine learning is programming computers to optimize a performance criterion using example data or past experience. We have a model defined up to some parameters, and learning is the execution of a computer program to optimize the parameters of the model using the training data or past experience. The model may be predictive to make predictions in the future, or descriptive to gain knowledge from data. The field of study known as machine learning is concerned with the question of how to construct computer programs that automatically improve with experience. Formal definition was given by Tom M. Mitchell:

"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E."

Problem Statement

Since the Advent of the Internet, humans have used it as a communication tool in the form of mostly text messages and nowadays video and audio streams and as we increase our dependence on technology it becomes increasingly important to better gauge human sentiments expressed with the help of technology. However, in this textual communication data, we lose the access to sentiments or the emotions conveyed behind a sentence, as we often use our hands and facial expressions to express our intent behind the statement. From this textual data, we can gain insights into the individual. Insights which can be used for multiple different uses such as content recommendation based on current mood, market segmentation analysis and psychological analysis in humans. In this project, we have attempted to classify human sentiment into three categories namely positive, negative and neutral. Which helps us better understand human thinking and gives us an insight which can be used in a variety of ways as stated above.



4. SYSTEM ANALYSIS

> Existing System

Traditionally, analysis of sentiment has been done on text, but a large amount of data is now being uploaded as reviews, images, emoticons, and videos. By inspecting these data, the sentiment of the public toward a specific matter can be analyzed, examined and discovered. Over the years, people considered emoticons as a medium of communication that is used in texts or solely to dedicate one's sentiment in an efficient manner. The existing work on sentiment analysis can be classified from different points of views: technique used, view of the text, level of detail of text analysis, rating level, etc.

> Proposed System

In this project we find how to identify tweets with a negative, positive and neutral sentiment. To do so, we'll create a sentiment analyzer to classify positive and negative tweets in text format. Though we'll be using our classifier for Twitter data analysis, it can also be used to analyze text data from other sources as well. We are going to look at datasets, various text processing, and embedding techniques, and then employ a machine learning model to process our data.

a) Objective of the System

The objective of sentiment analysis is to accurately extract people's opinions from a large number of unstructured review texts and classifying them into sentiment classes, i.e., positive, negative, or neutral. Sometimes "highly positive" and "highly negative" are also considered. Sentiment analysis is also called opinion mining, which refers to the use of natural language processing and text mining to identify the emotional information from text materials. Sometimes, social media behavior analysis also plays an important role in sentiment analysis and recommendation systems. Through social media, users can read online news, comment, like and share posts, blogs, and videos, and write reviews or opinions about products.

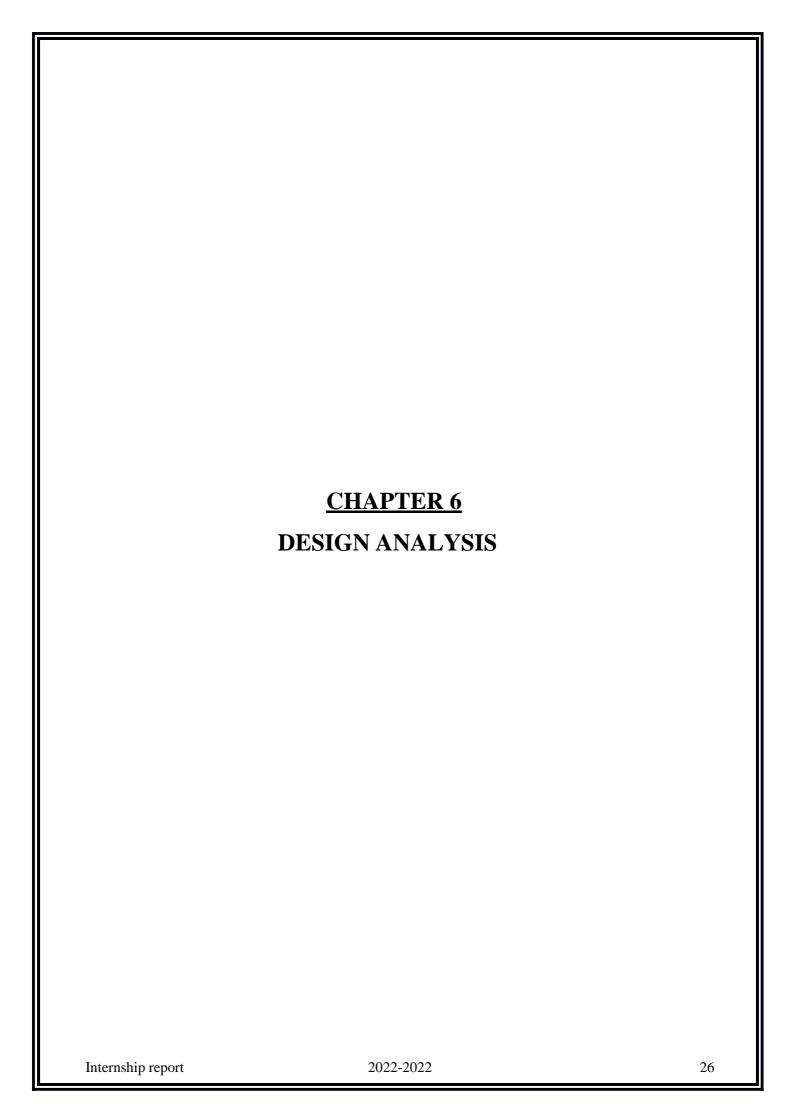
5. <u>REQUIREMENT ANALYSIS</u>

***** Hardware Requirement Specification

- Operating system Windows 8 or later, macOS Sierra 10.12 or later,
- Processor Intel Pentium 4 or later, Intel, Intel Pentium 4 or later
- Memory 2 GB minimum, 4 GB recommended
- Screen resolution 1280x1024 or larger
- Application window size 1024x680 or larger
- Internet connection Required

Software Requirement Specification

- Python Editor vs code, Jupyter notebook, Google colab
- Web Browser Chrome, Firefox, Safari
- Libraries used
 - i. Tweepy
 - ii. TextBlob
- Twitter API



6. DESIGN & ANALYSIS

> Sentiment Analysis:

Sentiment analysis is a natural language processing technique that identifies the polarity of a given text. There are different flavors of sentiment analysis, but one of the most widely used techniques labels data into positive, negative and neutral.

> Tweepy: a Python Library for the Twitter API

Tweepy is a Python library for accessing the Twitter API. It is great for simple automation and creating twitter bots. Tweepy has many features.

- Get tweets from timeline.
- Creating and deleting Tweets.
- Initialization:

```
import tweepyauth = tweepy.OAuthHandler(consumer_key,
consumer_secret)
auth.set_access_token(access_token, access_token_secret)
api = tweepy.API(auth)
```

- Tweepy supports the OAuth 1.0a User Context, OAuth 2.0 Bearer Token (App-Only), and OAuth 2.0 Authorization Code Flow with PKCE (User Context) authentication methods.
- The simplest way to authenticate as your developer account is to generate an access token and access token secret through your app's Keys and Tokens tab under the Twitter Developer Portal Projects & Apps page.

> TextBlob: Simplified Text Processing

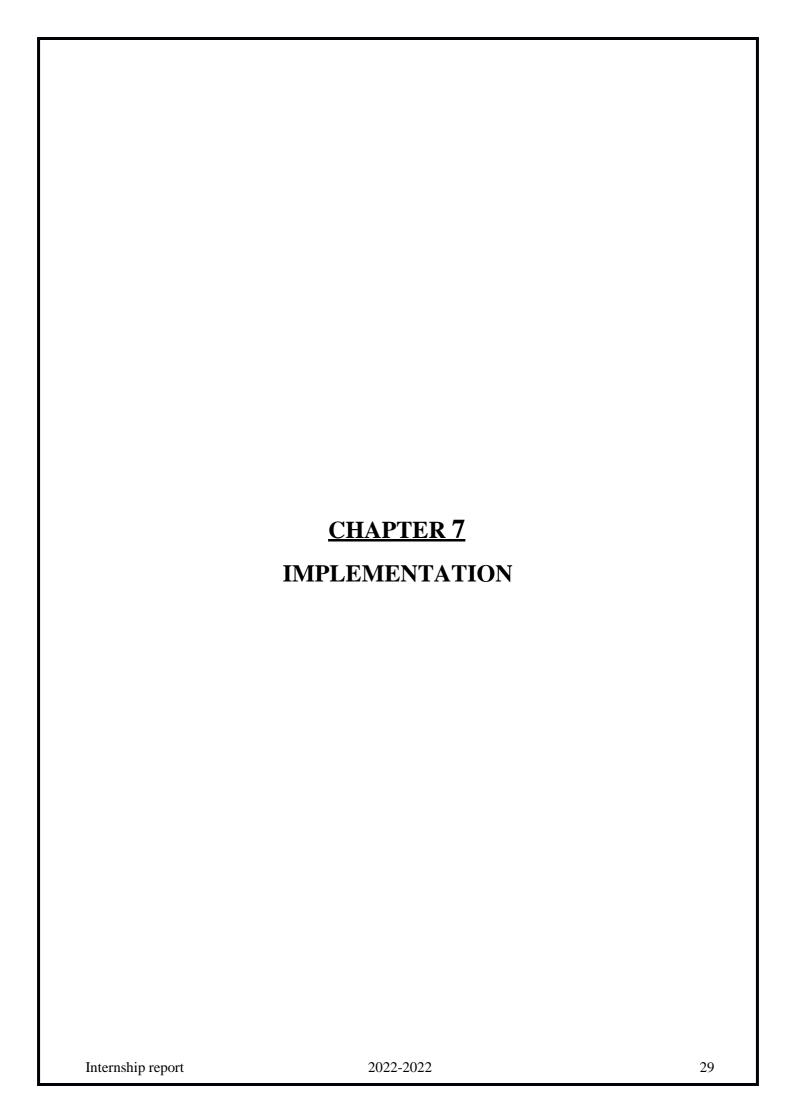
- TextBlob is a Python (2 and 3) library for processing textual data.
- It provides a simple API for diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more.
- TextBlob stands on the giant shoulders of NLTK and pattern, and plays nicely with both

Features:

- Noun phrase extraction
- Part-of-speech tagging
- Sentiment analysis
- Classification (Naive Bayes, Decision Tree)
- Tokenization (splitting text into words and sentences)
- Word and phrase frequencies
- Parsing
- **n**-grams
- Word inflection (pluralization and singularization) and lemmatization
- Spelling correction
- Add new models or languages through extensions
- WordNet integration

> Twitter API:

- The Twitter API is a set of programmatic endpoints that can be used to understand or build the conversation on Twitter.
- This API allows you to find and retrieve, engage with, or create a variety of different resources including the following:
 - a) Tweets
 - b) Users
 - c) Spaces
 - d) Direct Messages
 - e) Lists
 - f) Trends
 - g) Media



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 it constraints on implementation, design of methods to achieve the change over and an evaluation of change over methods 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.

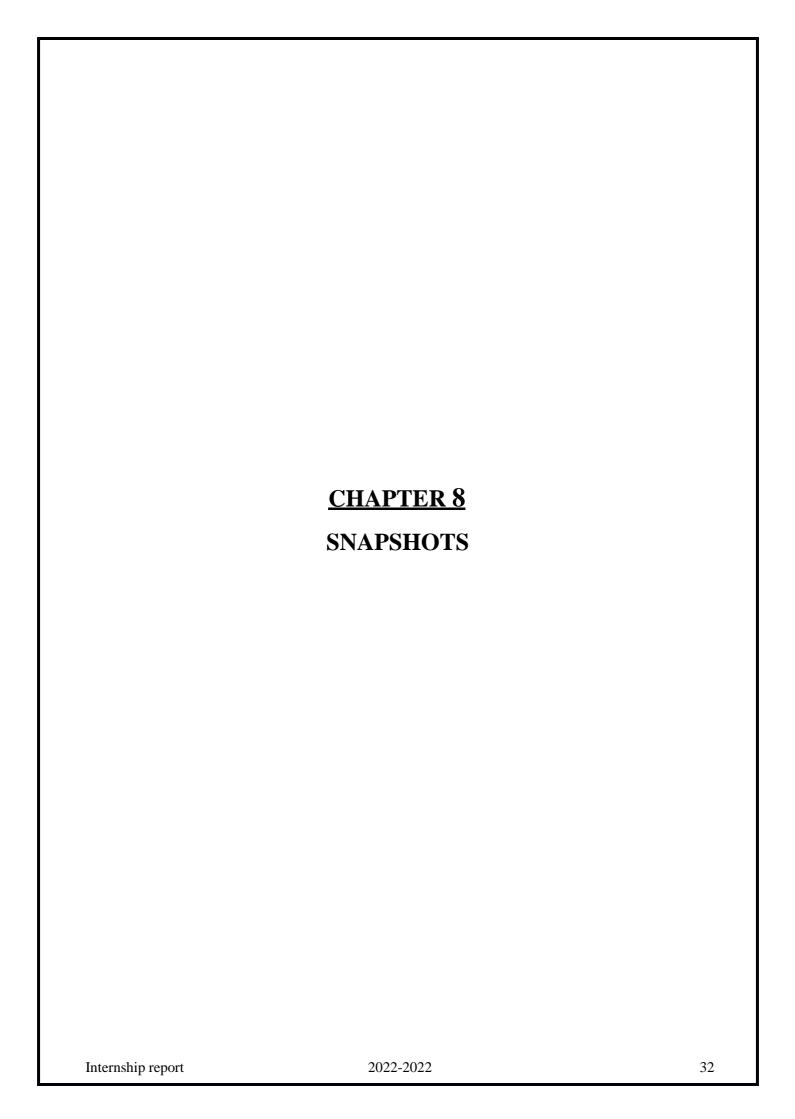
The implementation phase comprises of several activities. The required hardware and software acquisition is carried out. The system may require some software to be developed. For this, programs are written and tested. The user then changes over to his new fully tested system and the old system is discontinued.

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:

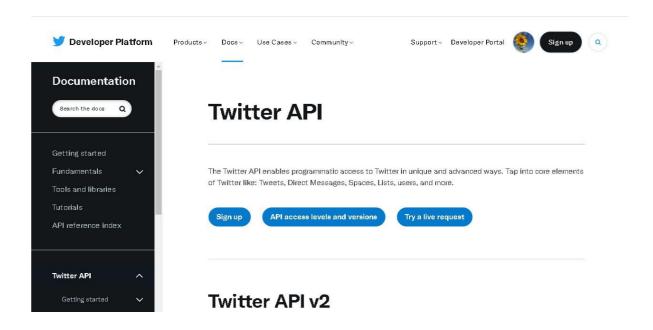
- 1. The first includes unit testing, where in each module is tested to provide its correctness, validity and also determine any missing operations and to verify whether the objectives have been met. Errors are noted down and corrected immediately.
- Unit testing is the important and major part of the 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.
_	

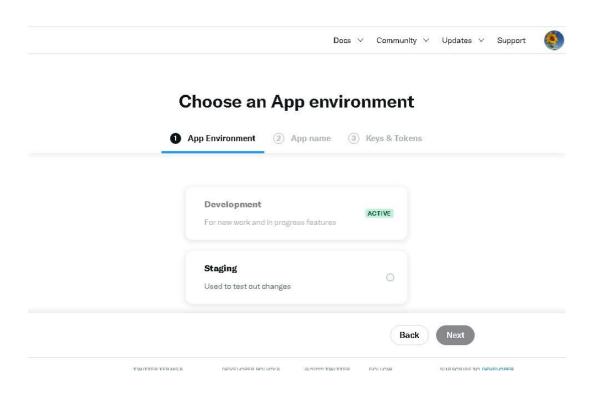


8. SNAPSHOTS

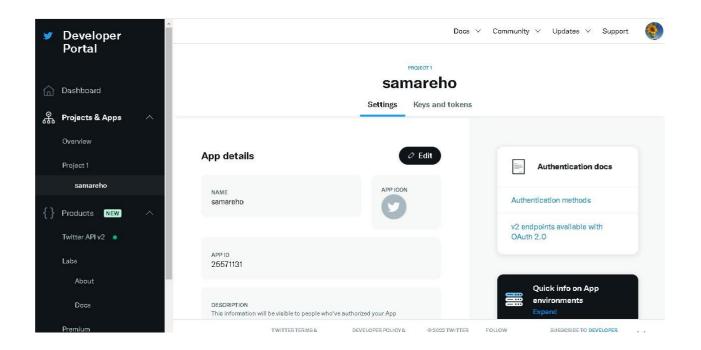
a) Creating a account at https://developer.twitter.com/



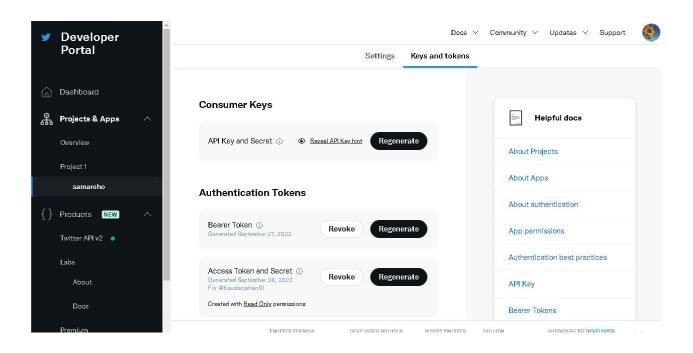
b) Creating an app in the Twitter development portal.



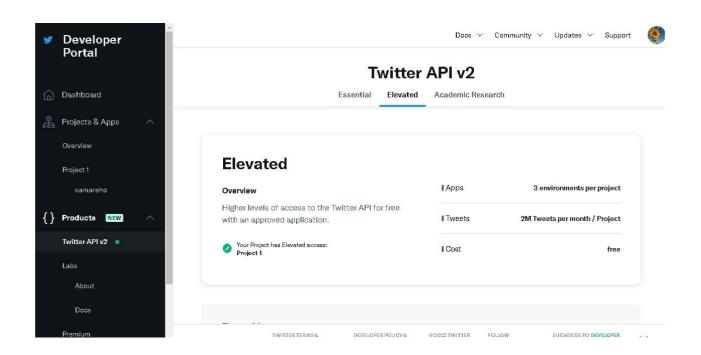
c) Creating a project.



d) Generating keys and tokens.



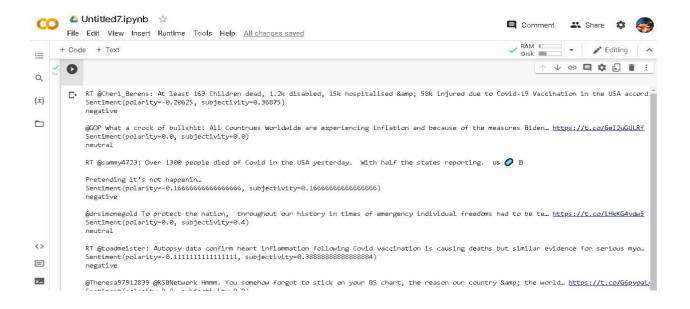
e) Applying for elevated access.

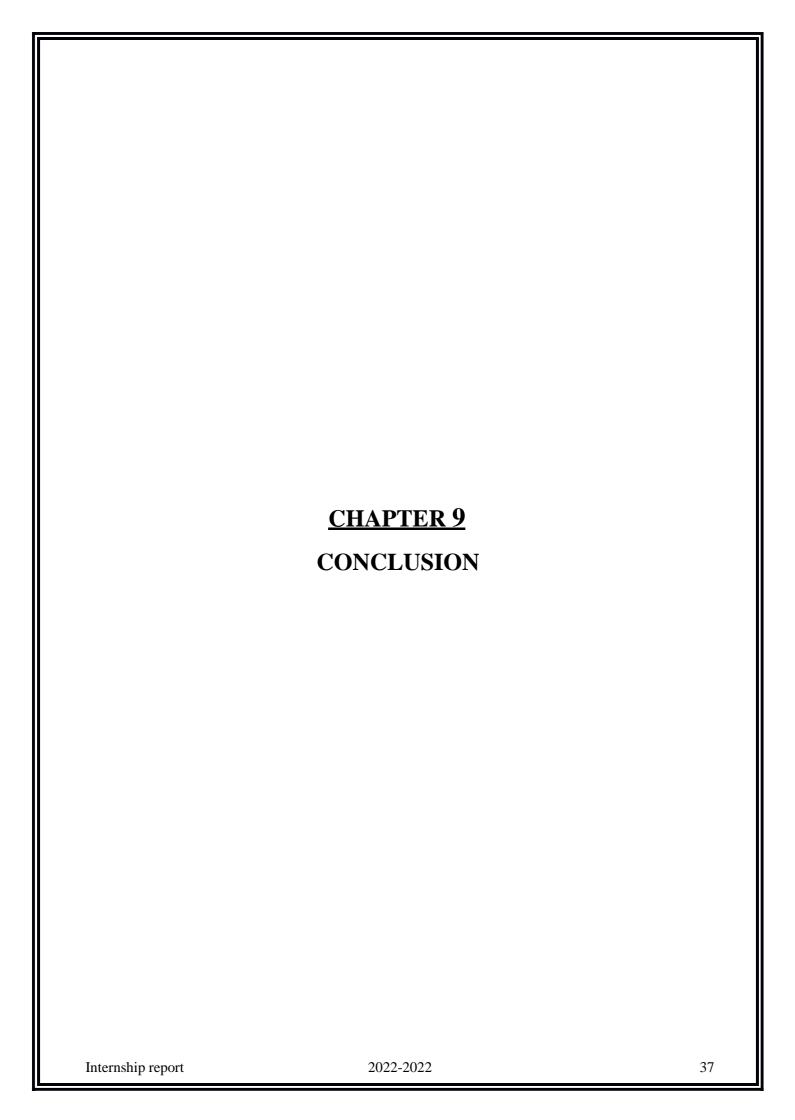


f) Including access keys in the code snippet.



g) Output: Tweets are fetched from Twitter and the sentiment analysis is done.





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] Ansari GJ, Shah JH, Yasmin M, Sharif M, Fernandes SL (2018) A novel machine learning approach for scene text extraction.
- [2] https://www.tweepy.org/
- [3] https://textblob.readthedocs.io/en/dev/
- [4] https://developer.twitter.com/en
- [5] https://en.wikipedia.org/wiki/Sentiment analysis