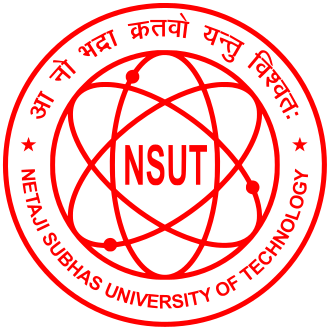
**PSYCHOLOGICAL AID FOR**

**VICTIMS**

**IN EMERGENCY**

**Soft Computing Project Report**

****

**DEPARTMENT OF COMPUTER SCIENCE ENGINEERING**

**NETAJI SUBHAS UNIVERSITY OF TECHNOLOGY**

**DWARKA, NEW DELHI**

**June, 2020**

***Submitted by Group No. P12***

**Vishal Kumar, 2019PCS2006**

**Anubhav Dhankhar, 2017UCO1550**

**Mansi Joshi,** **2017UCO1576**

**Prabhat Kumar, 2017UCO1621**

**CANDIDATE’S DECLARATION**

We hereby certify that the work which is being presented in the **Soft Computing Project** entitled **“Psychological Aid for Victims in Emergency”** in partial fulfilment for the award of the Degree of Computer Science Engineering affiliated to **Netaji Subhas University of Technology Dwarka, New Delhi** and submitted to the Department of Computer Science Engineering. It is an authentic record of our own work carried out during a period from **January 2020 to June 2020**.

**Date: - 29th June, 2020**

**Vishal Kumar (2019PCS2006)**

**Anubhav Dhankhar (2017UCO1550)**

**Mansi Joshi (2017UCO1576)**

**Prabhat Kumar (2017UCO1621)**

**ACKNOWLEDGEMENT**

We would like to express gratitude to professors **Shampa Chakraverty** and **Kanika Bhatia** for providing invaluable guidance and suggestions throughout the session of **Soft Computing Project**.

We express our sincere thanks to professors for encouraging and allowing us to present the **Soft Computing Project** **“Psychological Aid for Victims in Emergency”** at department premises for partial fulfilment of the requirements leading to award Degree of Computer Science Engineering. We take this opportunity to thank professors who have helped in our project.

**ABSTRACT**

Sentiment Analysis is a kind of Natural Language Processing to determine traits and emotions of a person by using text data. The Dataset is created with the help of Social Media APIs that collect information from messages, tweets and profiles. Sentiment Analysis can help victims by giving suggestions of recommended psychologists based on traits and emotions.

**TABLE OF CONTENTS**

**CHAPTER NO. TITLE PAGE NO.**

**ACKNOWLEDGMENT**

**ABSTRACT**

**CHAPTER 1 INTRODUCTION**

1.1 Problem Statement 6

**CHAPTER 2 PRE-TEST**

2.1 Skills of team members as per project (Pre-test) 7

**CHAPTER 3 PROGRESS ACHIEVED**

3.1 Division of Work 9

3.2 Task Completed 9

**CHAPTER 4 DATASET**

4.1 Dataset Used 10

4.2 API Employed 10

4.3 How to get started with the Twitter APIs 10

4.4 Pre-processing of Dataset 11

**CHAPTER 5 MODELS AND SOFT COMPUTING**

5.1 Sentiment Analysis using TextBlob 12

5.2 Workflow 12

**CHAPTER 6 WORKING OF SENTIMENT ANALYSIS**

6.1 Analysing the Sentiment of a Tweet 13

**CHAPTER 7 LEARNING FROM THE PROJECT**

7.1 Challenges 15

7.2 Learning 15

**CHAPTER 8 POST-TEST**

8.1 Skills of team members as per project (Post-test) 16

**REFERENCES**

CHAPTER-1

**INTRODUCTION**

In this project report, we propose a Natural Language Processing approach to compute personality traits. The Twitter API segments tweets of a location in tokens, then TextBlob Library performs Sentiment Analysis to find emotion of tweets.

1.1 Problem Statement

Design an application identify the types of psychologist for counselling, gathering personality traits through social media accounts, mapping of victims to psychologists on the basis of their mental state and the type of loss incurred.

CHAPTER-2

**PRE-TEST**

2.1 Skills of team members as per project (Pre-test)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skills | Panda | HTML | Bootstrap | OO Modelling | JavaScript |
| Vishal Kumar (2019PCS2006) | W | S | W | M | M |
| Anubhav Dhankhar (2017UCO1550) | W | S | W | M | M |
| Mansi Joshi (2017UCO1576) | M | S | W | M | W |
| Prabhat Kumar (2017UCO1621) | W | M | W | M | M |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skills | Data Structures | ER  Modelling | MySQL | NN | Software Testing |
| Vishal Kumar (2019PCS2006) | M | S | M | M | M |
| Anubhav Dhankhar (2017UCO1550) | S | M | M | W | W |
| Mansi Joshi (2017UCO1576) | S | M | W | S | W |
| Prabhat Kumar (2017UCO1621) | S | M | M | W | M |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skills | Software  Engineering | NumPy | Software Automation | Android | GitHub |
| Vishal Kumar (2019PCS2006) | M | M | W | W | W |
| Anubhav Dhankhar (2017UCO1550) | M | W | W | W | W |
| Mansi Joshi (2017UCO1576) | M | S | W | W | M |
| Prabhat Kumar (2017UCO1621) | M | W | W | W | W |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skills | Genetic Algorithm | Algorithms | Version  Control | SQL | Python |
| Vishal Kumar (2019PCS2006) | M | M | W | M | M |
| Anubhav Dhankhar (2017UCO1550) | W | M | W | M | M |
| Mansi Joshi (2017UCO1576) | W | S | S | M | S |
| Prabhat Kumar (2017UCO1621) | W | S | W | M | S |

CHAPTER-3

**PROGRESS ACHIEVED**

3.1 Division of Work

VISHAL KUMAR, ANUBHAV DHANKHAR:

Integrating sentiments and tweets of a location and study how to classify tweets.

MANSI JOSHI:

Pre-processing of dataset, sentiment analysis of tweets and their classification, Plotting of results.

PRABHAT KUMAR:

Find and explored various API’s that can be used for this work, Extracted the live tweets using TWEEPY API, added the filter to get location wise tweets.

3.2 Tasks Completed

1. Learnt to use the Twitter API.
2. Obtained tweets for given coordinates.
3. Sentiment Analysis and Classification of a given tweet achieved.
4. Tried integrating Step 2 and Step 3.

CHAPTER-4

**DATASET**

4.1 Dataset Used

The twitter API has been used to fetch tweets that are used as a dataset in this project. Live tweets have been extracted and the process is dynamic. Tweets are short texts and contain relevant topics, making it a good dataset for Sentiment Analysis. It is also a great dataset to analyze user engagement with their respective locations, since geo locations can be identified.

4.2 API Employed

Tweepy is a Python library for accessing the Twitter API. It is great for simple automation and creating twitter bots. The Twitter API exposes dozens of HTTP endpoints that can be used to retrieve, create and delete tweets, retweets and likes. The Twitter streaming API is used to download twitter messages in real time. We can filter tweets using its geolocation after identifying the bounding box of its longitude and latitude.

4.3 How to get started with the Twitter APIs

1. Apply and receive approval for a Twitter developer account.
2. Create a Twitter developer app.
3. Generate your app’s API keys and user’s access tokens.
4. Generate your app’s bearer token.
5. Apply and receive access to the desired API.
6. Find the documentation, libraries, code examples, and other resources that you need to make your first successful request.

4.4 Pre-processing of Dataset

* The raw data having polarity is highly susceptible to inconsistency and redundancy.
* Remove all URLs (e.g. www.xyz.com), hash tags (e.g. #topic), targets (@username)
* Correct the spellings (sequence of repeated characters is to be handled)
* Replace all the emoticons with their sentiment
* Remove all punctuations, symbols, numbers
* Remove Stop Words Expand Acronyms (we can use an acronym dictionary)
* Remove Non-English Tweets

CHAPTER-5

**Models and Soft Computing Technology**

5.1 Sentiment Analysis using TextBlob

TextBlob is a python library and offers a simple API to access its methods and perform basic NLP tasks.

**Sentiment Analysis** is basically the process of determining the attitude or the emotion of the writer, i.e., whether it is positive, negative or neutral. The sentiment function of TextBlob returns polarity. Cleaning of tweets to remove special characters.

Polarity is a value which lies in the range of [-1,1] where positive value means positive statement, negative value means negative statement, and 0 means neutral statement.

5.2 Workflow

* Inputs - User feeds in the name of the place and the number of tweets required for his study.
* Tweepy client makes the API call to fetch tweets of that location.
* Fetched tweets saved in an output file.
* Cleaning of tweets to remove special characters.
* Textblob returns a polarity for each tweet.
* Classification of tweets and plotting of results.

CHAPTER-6

**WORKING OF SENTIMENT ANALYSIS**

6.1 Analysing the Sentiment of a Tweet

Figure 1: Sample Query

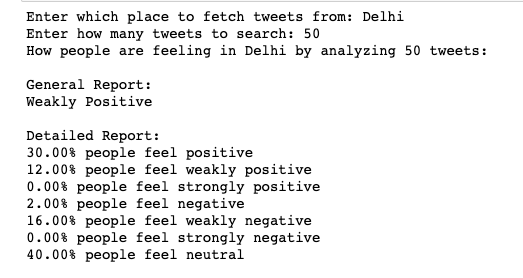


Figure 2: Pie Chart Classification

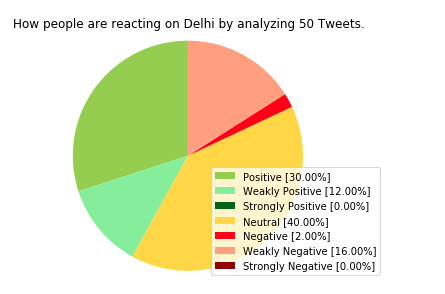


Figure 3: Tweet Text

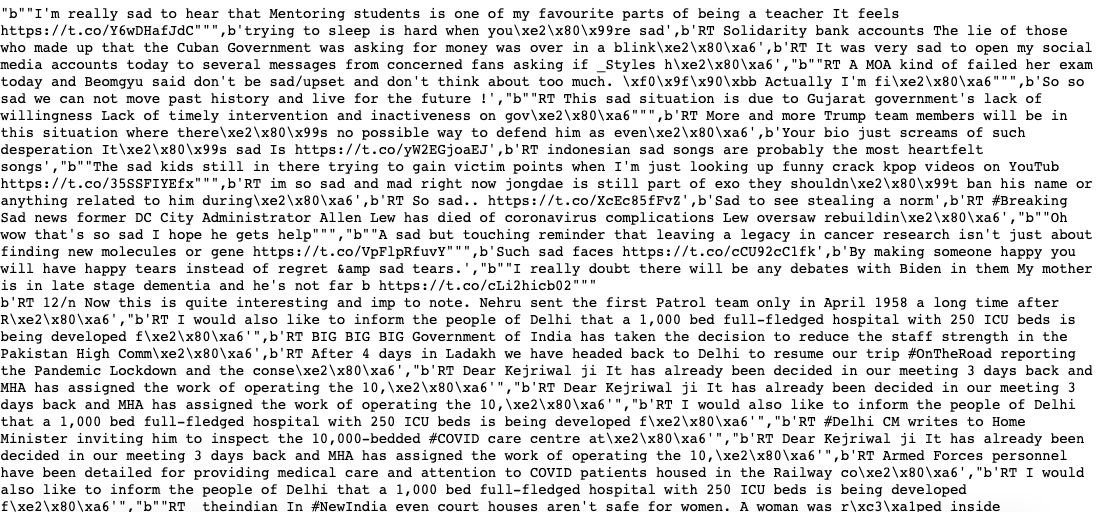


Figure 4: List of Places

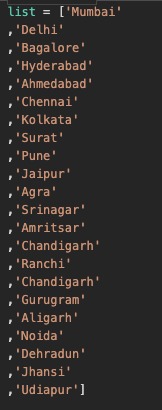


Figure 5: First 3 highly negative cities



Figure 6: Suggesting a Psychologist



Figure 7: Suggesting a Psychologist



CHAPTER-7

**LEARNING FROM THE PROJECT**

7.1 Challenges

* Developer account approval
* A lot of unstructured data.
* Fake profiles on twitter.
* Only 1% tweets are geotagged.
* Execution error when trying to integrate “tt.py” and Sentiment module.
* There is a limit of accessing tweets using API.
* Lack of communication with all the group members making the project somewhat cumbersome.
* Ongoing internships.

7.2 Learning

Vishal Kumar - Python, NLP using TextBlob

Anubhav Dhankhar - Python, NLP

Mansi Joshi - Python, Use of Twitter APIs and tweepy Client

Sentiment Analysis

Prabhat Kumar - Python, Using API, Cleansing of data

CHAPTER-8

**POST-TEST**

8.1 Skills of team members as per project (Post-test)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skills | Panda | HTML | Bootstrap | OO Modelling | JavaScript |
| Vishal Kumar (2019PCS2006) | **M** | S | W | M | M |
| Anubhav Dhankhar (2017UCO1550) | **M** | **M** | W | M | M |
| Mansi Joshi (2017UCO1576) | M | S | W | M | **M** |
| Prabhat Kumar (2017UCO1621) | **M** | M | W | M | M |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skills | Data Structures | ER  Modelling | MySQL | NN | Software Testing |
| Vishal Kumar (2019PCS2006) | M | S | M | M | M |
| Anubhav Dhankhar (2017UCO1550) | S | M | M | W | W |
| Mansi Joshi (2017UCO1576) | S | M | W | S | W |
| Prabhat Kumar (2017UCO1621) | S | M | M | W | M |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skills | Software  Engineering | NumPy | Software Automation | Android | GitHub |
| Vishal Kumar (2019PCS2006) | M | M | W | W | W |
| Anubhav Dhankhar (2017UCO1550) | M | W | W | W | W |
| Mansi Joshi (2017UCO1576) | M | S | W | W | **S** |
| Prabhat Kumar (2017UCO1621) | M | **M** | W | W | W |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skills | Genetic Algorithm | Algorithms | Version  Control | SQL | Python |
| Vishal Kumar (2019PCS2006) | M | M | W | M | M |
| Anubhav Dhankhar (2017UCO1550) | **M** | M | W | M | M |
| Mansi Joshi (2017UCO1576) | **M** | S | S | M | S |
| Prabhat Kumar (2017UCO1621) | **M** | S | W | M | S |

**REFERENCES**

[1] Tweepy Documentation - docs.tweepy.org

[2] Twitter API - developer.twitter.com/en/docs

[3] TextBlob - textblob.readthedocs.io/en/dev/

[4] Stack Overflow - stackoverflow.com

[5] How to Track Tweets by Geographic Location - www.bmc.com/blogs/track-tweets-location

[6] Vishal A. Kharde and S.S. Sonawane, “Sentiment Analysis of Twitter Data: A Survey of

Techniques”, International Journal of Computer Applications, Vol 139 - No.11, April 2016

[7] Barbara Plank, Dirk Hovy - Personality Traits on Twitter or. How to get 1,500 Personality Tests in a Week

[8] Sentiment Analysis - monkeylearn.com/sentiment-analysis/