TABLE OF CONTENTS

CONTENTS

ABSTRACT

Sentiment analysis is a topic of great interest and development since it has many practical applications. Companies use sentiment analysis to automatically analyze survey responses, product reviews, social media comments, and the like to get valuable insights about their brands, products, and services. Our model is basically the process of identifying and categorizing opinions expressed in a piece of text, in order to determine whether the person's attitude towards a particular topic is positive, negative, or neutral. Sentiment analysis is helpful in social media monitoring and reviews analysis, where brands need to extract public opinion out of a vast amount of data. Sentiment analysis leverages machine learning and natural language processing techniques to assign sentiment scores to pieces of writing. It focuses on the textual data or information and the live twitter analysis. This project uses python programming importing various libraries. A web application to differentiate upon finding out the sentiment response between the twitter live hashtags and textual data considering the adjectives.

INTRODUCTION

Sentiment analysis is the present and future of emerging and established brands. By using sentiment analysis, brands can get a definite feel of how customers perceive their products and services. Sentiment analysis can transform how you analyze, interpret, and evaluate customer feedback. Depending on how detailed you want the sentiment analysis to be, you can extract text from a paragraph, sentence, or a complete document. This analysis detects emotions in the text. It used to identify the happy, sad, or angry types of emotions. It is an algorithm based sentiment analysis tool that helps to understand what a person feels a certain way. Massive volume of data is always a significant concern for most businesses. Emails, social media, chats, articles, and support tickets come in big numbers, and most are unorganized.

LITERATURE SURVEY

- **Opinion Mining**

Opinion mining refers to the broad area of natural language processing, text mining, computational linguistics, which involves the computational study of sentiments, opinions and emotions expressed in text . Although, view or attitude based on emotion instead of reason is often colloquially referred to as a sentiment . Hence, lending to an equivalent for opinion mining or sentiment analysis. stated that opinion mining has many application domains including accounting, law, research, entertainment, education, technology, politics, and marketing. In earlier days many social media have given web users avenue for opening up to express and share their thoughts and opinions

- **Twitter**

Twitter is a popular real time microblogging service that allows users to share short information known as tweets which are limited to 140 characters. Users write tweets to express their opinion about various topics relating to their daily lives. Twitter is an ideal platform for the extraction of general public opinion on specific issues.A collection of tweets is used as the primary corpus for sentiment analysis, which refers to the use of opinion mining or natural language processing. Twitter, with 500 million users and million messages per day, has quickly became a valuable asset for organizations to invigilate their reputation and brands by extracting and analyzing the sentiment of the tweets by the public about their products, services market and even about competitors.highlighted that, from the social media generated opinions with the mammoth growth of the world wide web, super volumes of opinion texts in the form of tweets, reviews, blogs or any discussion groups and forums are available for analysis, thus making the world wide web the fastest, most comprising and easily accessible medium for sentiment analysis.

- **Microblogging with E-commerce**

A microblogging platform such as Twitter is similar to a conventional blogging platform; just single posts are shorter . Twitter has a small number of words which are designed for the quick transmission of information or exchange of opinion . However, small business or large organizations are initiating the potential of microblogging platforms that have been developed a few years' time for promoting foreign trade websites by using a foreign microblogging platform as Twitter marketing . The instant of sharing, interactive, community-oriented features are opening an e-commerce, launched a new bright spot which it can be shown that microblogging platform has enabled companies do brand image, product important sales channel, improve product sales, talk to consumer for a good interaction and other business activities involved. said, in fact, the companies manufacturing such products have started to poll theses microblogs to get a sense of general sentiment for a product. Many times these companies study user reactions and reply to users on microblogs .

- **Social Media**

defined a social media as a group of Internet-based applications that create on the ideological and technological foundations of Web2.0 which is allowed to build and exchange user generated contents. In a discussion of Internet World Start, identified that a trend of internet users is increasing and continuing to spend more time with social media by the total time spent on mobile devices and social media in the U.S.across PC increased by 37 percent to 121billion minutes in 2012, compared to 88 billion minutes in 2011. On the other hand, businesses use social networking sites to find and communicate with clients, business can be demonstrated to damage productivity caused by social networking . As social media can be posted so easily to the public, it can harm private information to spread out in the social world . On the contrary, discussed that the benefits of participating in social media have gone beyond simply social sharing to build an organization's reputation and bring in career opportunities and monetary income. In addition, social media is also being used for advertisement by companies for promotions, professionals for searching, recruiting, social learning online and electronic commerce. Electronic commerce or E-commerce refers to the purchase and sale of goods or services online which can via social media, such has Twitter which is convenient due to its 24-hours availability, ease of customer service and global reach . Among the reasons why business tends to use more social media is for getting insight into consumer

behavioral tendencies, market intelligence and presenting an opportunity to learn about customer review and perceptions.

- **Twitter Sentiment Analysis**

The sentiment can be found in the comments or tweet to provide useful indicators for many different purposes . Also, stated that a sentiment can be categorized into two groups, which is negative and positive words. Sentiment analysis is a natural language processing technique to quantify an expressed opinion or sentiment within a selection of tweets Sentiment analysis refers to the general method to extract polarity and subjectivity from semantic orientation which refers to the strength of words and polarity text or phrases. There has two main approaches for extracting sentiment automatically which are the lexicon-based approach and machine-learning-based approach.

1. Lexicon-based Approach

Lexicon-based methods make use of predefined list of words where each word is associated with a specific sentiment . The lexicon methods vary according to the context in which they were created and involve calculating orientation for a document from the semantic orientation of texts or phrases in the documents. Besides, it also states that a lexicon sentiment is to detect word-carrying opinion in the corpus and then to predict opinion expressed in the text. has shown the lexicon methods which have a basic paradigm which are:

i. Preprocess each tweet, post by remove punctuation

ii. Initialize a total polarity score (s) 0

iii. Check if token is present in a dictionary, then If token is positive, s will be positive(+) If token is negative, s will be negative (-)

iv. Look at the total polarity score of tweet post If s > threshold, tweet post as positive If s< threshold, tweet post as negative.

2. Machine-learning-based Approach

Machine learning methods often rely on supervised classification approaches where sentiment detection is framed as a binary which are positive and negative. This approach requires labeled data to train classifiers. This approach, it becomes apparent that aspects of the local context of a word need to be taken into account such as negative (e.g. Not beautiful) and intensification (e.g. Very beautiful). However, [20] showed a basic paradigm for create a feature vector is: i. Apply a part of speech tagger to each tweet post ii. Collect all the adjective for entire tweet posts iii. Make a popular word set composed of the top N adjectives iv. Navigate all of the tweets in the experimental set to create the following:

• Number of positive words

• Number of negative words

• Presence, absence or frequency of each word

SOFTWARES USED

Front End

- HTML
- CSS
- Java Script

HTML

- HTML stands for Hyper Text Markup Language.HTML is the standard markup language for creating Web pages.
- HTML describes the structure of a Web page.

CSS

- CSS describes how HTML elements are to be displayed on screen, paper, or in other media. CSS saves a lot of work.
- CSS is used to define styles for your web pages, including the design, layout and variations in display for different devices and screen sizes.

Java Script

- JavaScript is the programming language of the web.
- JavaScript enables interactive web pages and is an essential part of web applications.
- The vast majority of websites use it for client-side page behavior and all major web browsers have a dedicated JavaScript engine to execute it.

Back End

- Python

- Flask

Python

- Python is an interpreted, high-level and general-purpose programming language.
- Python's design philosophy emphasizes code readability with its notable use of significant whitespace.
- Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.
- It supports multiple programming paradigms,including structured (particularly, procedural), object-oriented and functional programming.

Flask

- Flask is a web application framework written in Python.
- Flask supports the extensions to add such functionality to the application.
- A Flask application is started by calling the run() method.
- However, while the application is under development, it should be restarted manually for each change in the code. To avoid this inconvenience, enable debug support.
- The server will then reload itself if the code changes.
- It will also provide a useful debugger to track the errors if any, in the application.

METHODOLOGY

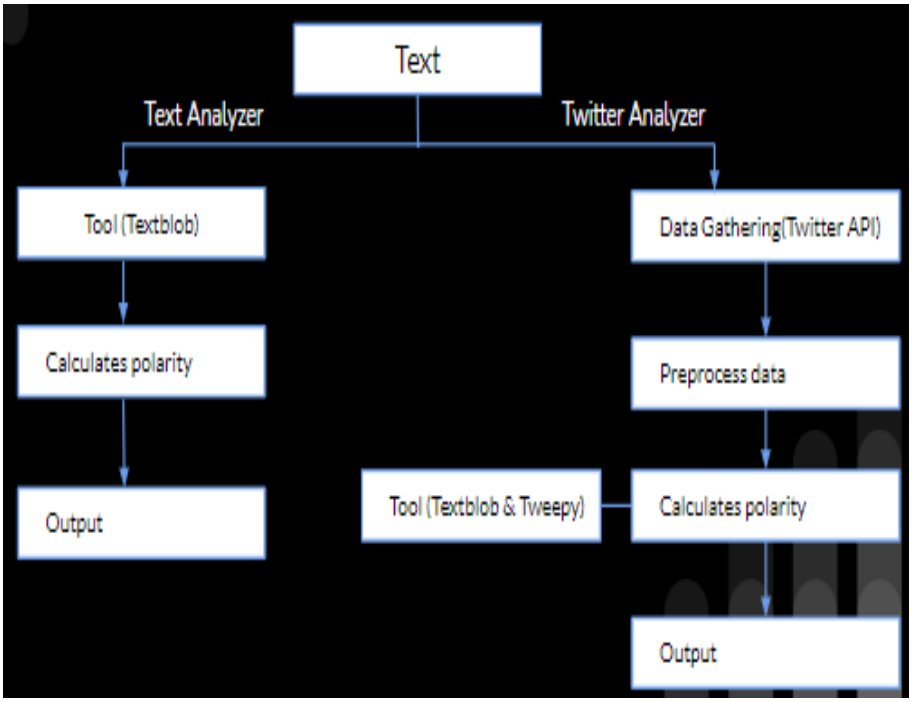
This project has been divided into 2 phases.First, literature study is conducted, followed by system development. Literature study involves conducting studies on various sentiment analysis techniques and methods that are currently in use. In phase 2, application requirements and functionalities are defined prior to its development. Also, architecture and interface design of the program and how it will interact are also identified. In developing the Twitter Sentiment Analysis application, several tools are utilized, such as Python Shell 3.9 and VSCode.

1. Proposed Architecture

As our goal is to achieve sentiment analysis for data provided from Twitter. We are going to build a classifier which consists of different python libraries.

We have taken some text as input from the user. Here we have 2 analyzers in the web application . One was text analyzer and another was twitter analyzer . Text analyzer gives sentiment behind the input while Twitter Analyzer will take the hashtag and analyzes the number of positive tweets and negative tweets.

2. FLOW CHART



Firstly we will learn what happens in text analyzer :

1 . Input is taken from the user and using some NLP library like Textblob we will calculate the polarity and gives the score of polarity as output

And then lets us learn what happens in twitter analyzer :

1. After user enters a keyword, we will take it as hashtag and this twitter API will collect all twitter post with that keyword.
2. After necessary preprocessing, all the tweets are send to calculate the polarity using textblob tool and gives output.

IMPLEMENTATION

For this web application, we designed the home page using html, css and javascript

CODE:

The below screenshots are the code we designed to deploy the application.

The below app.route is used to change the location or the directory we want to refer. To implement it we have used home and home1 function using the render_template method that redirects to the html page designed.

POST and GET are the http methods, the server receives data from POST method, value "nm" parameter obtained from form data and passed as part of the variable to USER url and browser displays the output.

Else if user in the session the page return to user, else returns to login page.

```
8   @app.route("/")
9   def home():
10      return render_template("sample1.html")
11
12  @app.route("/home")
13  def home1():
14      return render_template("sample1.html")
15
16
17  @app.route("/login", methods=["POST", "GET"])
18  def login():
19      if request.method == "POST":
20          session.permanent = True
21          user = request.form["nm"]
22          session["user"] = user
23          return redirect(url_for("user"))
24      else:
25          if "user" in session:
26              return redirect(url_for("user"))
27          return render_template("login.html")
28
```

The below code used to find the text polarity using python libraries and code is used to display the output as positive, negative or neutral by linking the html page login1.html, the values are passed to the html values and the output is displayed by giving text as input. The python textblob is imported which is used to process the textual information

```
@app.route("/user")
def user():
    if "user" in session:
        user = session["user"]
        y = user
        from textblob import TextBlob
        edu = TextBlob(y)
        x=edu.sentiment.polarity
        z = x*100
        z = str(z)
        if x<0:
            b = "Negative"
            return render_template('login1.html', value1=z, value2=b,value3=y)
        elif x == 0:
            d = "Neutral"
            return render_template('login1.html', value1=z, value2=d,value3=y)
        elif x>0 and x<=1:
            f = "Positive"
            return render_template('login1.html', value1=z, value2=f,value3=y)
    else:
        return redirect(url_for("login"))
```

The below code used to find the twitter sentiment polarity using python libraries such as textblob and tweepy and code is used to display the output as pie chart which contains value of positive and negative tweets by linking the html page login3.html, the values are passed to the html values and the output is displayed by giving text as input.

```python
@app.route("/user1")
def user1():
    if "user" in session:
        user = session["user"]
        import tweepy
        from textblob import TextBlob
        consumer_key = 'YPyCQ50LpE7qDJJf7pKPtbNTg'
        consumer_key_secret = '6jf3xFI6qghRbhS1NYfUE8gHG5YThZ6axqN1dcGlXtGBeLrUDR'
        access_token = '1333107433552506882-YHASmsWdHamoWKL1fP8hLtWPzeXhrR'
        access_token_secret = 'W1OoHb4eaBAMp5fCbMIunbRdfdJfqwiXJSYHQhZutj5OD'
        auth = tweepy.OAuthHandler(consumer_key, consumer_key_secret)
        auth.set_access_token(access_token, access_token_secret)
        api = tweepy.API(auth)
        public_tweets = api.search(user)
        p=0
        n=0
        for tweet in public_tweets:
            edu = TextBlob(tweet.text)
            x=edu.sentiment.polarity
            if x>0:
                p +=1
            else:
                n +=1
        a = ((p)/(p+n))*100
        b = 100-a
        return render_template('login3.html', value1=a, value2=b,value3=user)
    else:
        return redirect(url_for("login"))
```

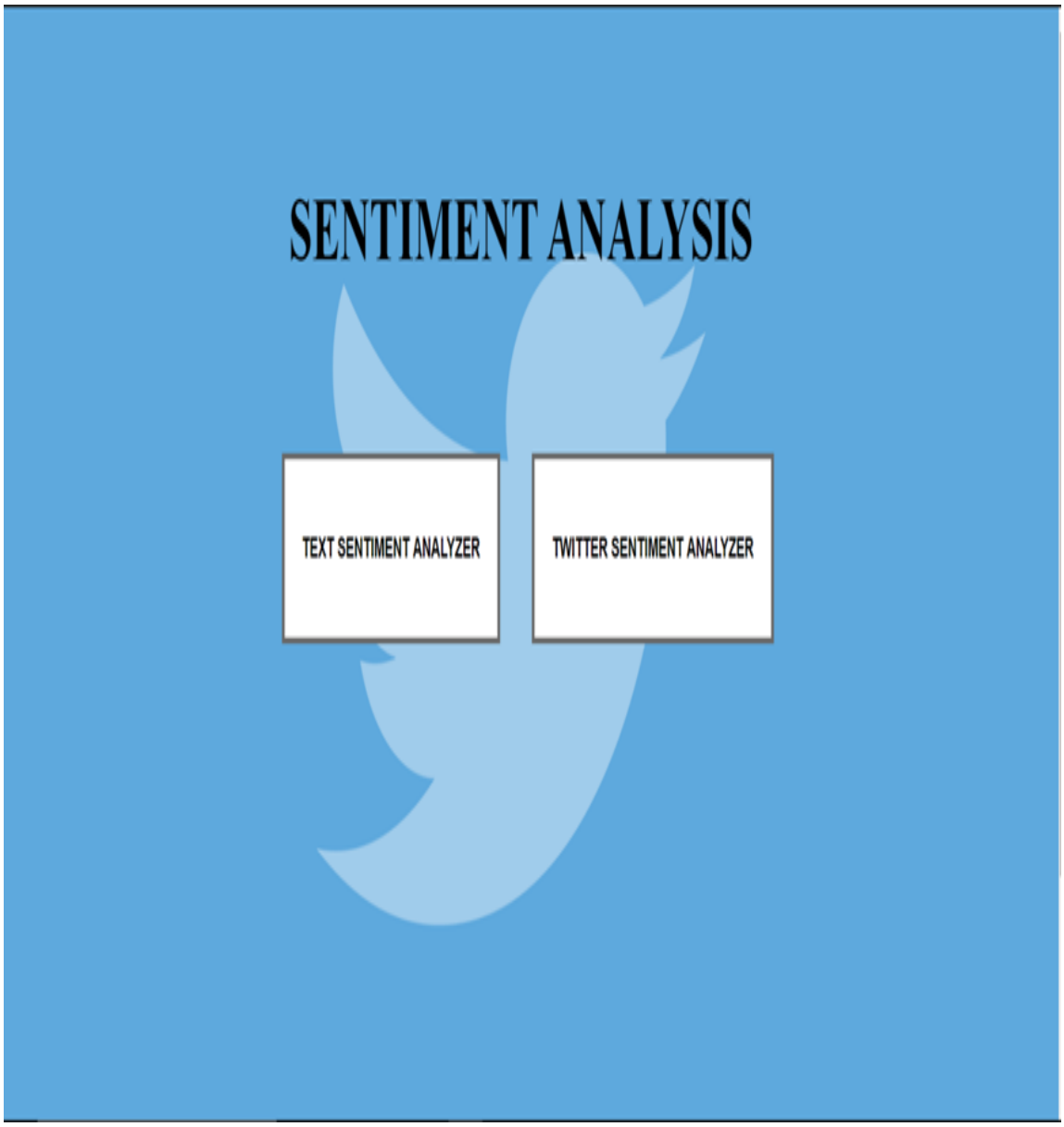
Here app.route("/back") is used to return to the home page, we defined the logout function in order to carry out the process. It leaves the current session and return to home page.

The name "main" mentioned is the main function that we use in every programming language, it's the start of the application. Run method is made to start the application.

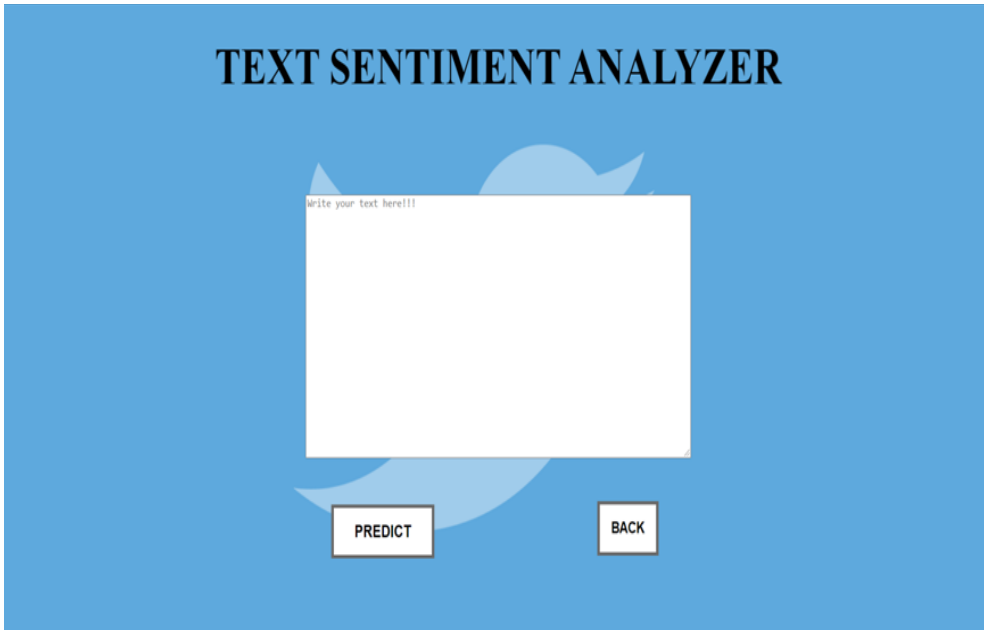
```python
@app.route("/back")
def logout():
    session.pop("user", None)
    return redirect(url_for("home"))


if _name_ == "_main_":
    app.debug = True
    app.run()
```

RESULT

Home Page

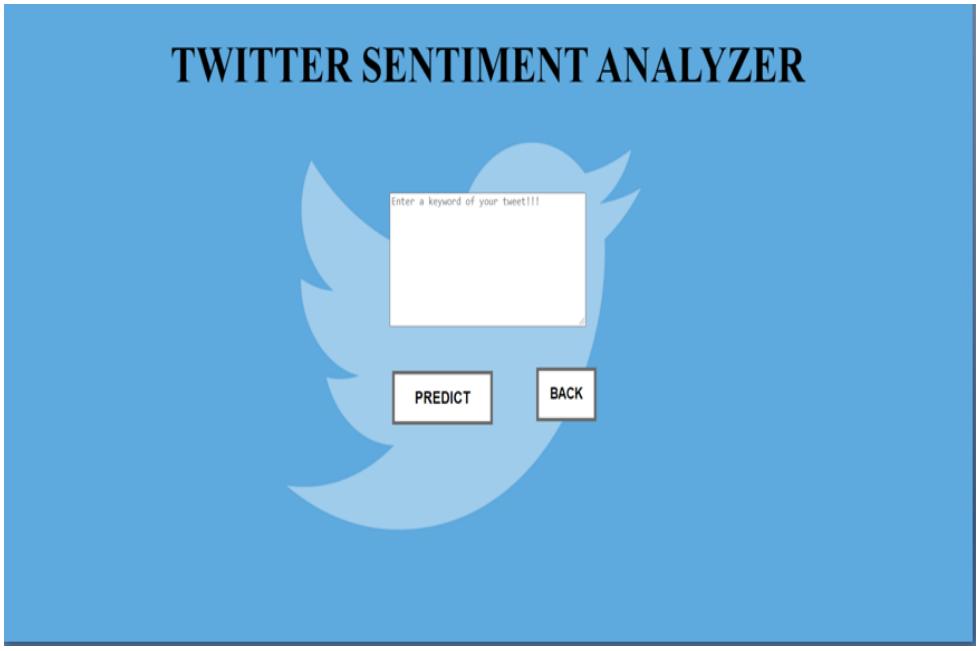


Text Analyzer home page

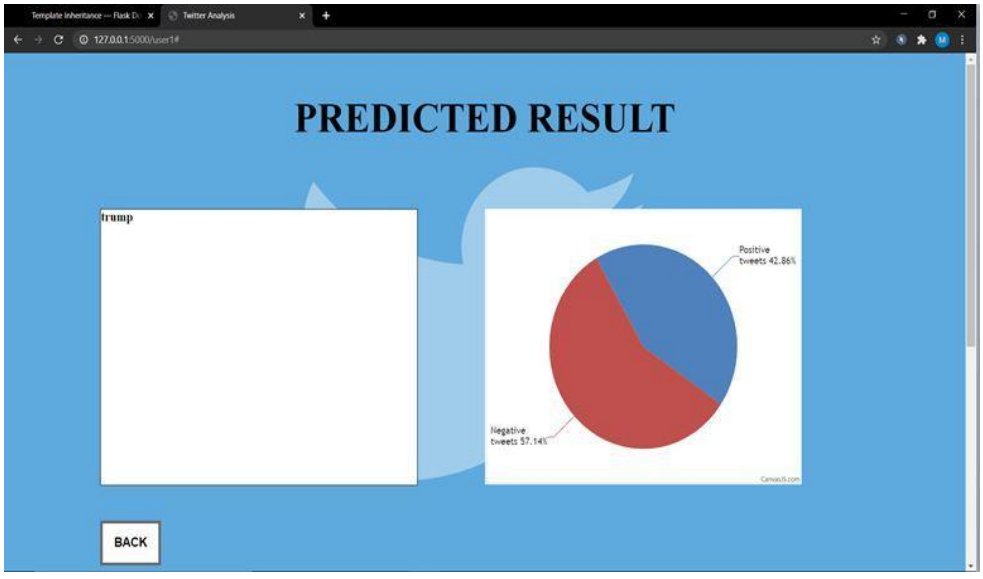


Text analyzer output

Twitter analyzer home page



Twitter analyzer output page



CONCLUSION

We presented results for sentiment analysis on Twitter and Text data using python. Sentiment analysis is a part of NLP, text can be classified by sentiment referred to as polarity. Applying sentimental analysis to extract the sentiment became an important work for many organizations and even individuals. The model created presents the approach for sentiment analysis to uncover the sentiment, we extracted the opinion words by combining the adjectives in the text and the hashtags from twitter. The project involved data visualization of polarity and subjectivity scores calculated with TextBlob. The code snippet shows a straightforward implementation of TextBlob and Tweepy on tweets streamed from twitter in real-time. This model can be improved further by using machine learning and deep learning concepts. Classification and feature extraction models can be implemented. The application can be converted to real-time user friendly by making it compatible with mobiles.

FUTURE SCOPE

Right now we have worked with only the very simplest models; we can improve those models by adding extra information like closeness of the word with a negation word. We could specify a window prior to the word (a window could for example be of 2 or 3 words) under consideration and the effect of negation may be incorporated into the model if it lies within that window. The closer the negation word is to the unigram word whose prior polarity is to be calculated, the more it should affect the polarity.

REFERENCES

1. Introduction to Sentiment Analysis, https://www.edureka.co/blog/sentiment-analysis-methodology/
2. Twitter Approches, https://monkeylearn.com/blog/sentiment-analysis-of-twitter/
3. Text Analysis, https://monkeylearn.com/blog/sentiment-analysis-deep-learning/
4. Python Text Libraries, https://textblob.readthedocs.io/en/dev/
5. Twitter api, http://docs.tweepy.org/en/latest/getting_started.html#introduction