TWITTER SENTIMENT ANALYSIS

Project report submitted in partial fulfillment of the requirement for the degree of

**Bachelor of
Technology In
Computer Science and
Engineering By
Robin Singh (161240)**

Under the

supervision of

Dr. Hari Singh

To

Department of Computer Science & Engineering and Information Technology

Jaypee University of Information Technology Waknaghat, Solan

173234, Himachal Pradesh

Candidate's Declaration

I hereby declare that the work presented in this report entitled TWITTER SENTIMENT ANALYSIS in fulfilment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering/Information Technology submitted in the department of Computer Science & Engineering and Information Technology, Jaypee University of Information Technology, Waknaghat is an authentic record of my own work carried out over a period from August 2018 to June 2020 under the supervision of Dr. Hari Singh, Associate Professor, Computer Science and Engineering/Information Technology.

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

Robin Singh, 161240

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

Dr. Hari Singh
Assistant Professor
(Senior Grade)

Computer Science and Engineering / Information Technology

Dated:

ACKNOWLEDGEMENT

I have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals and organizations. I would like to extend our sincere thanks to all of them.

I am highly indebted to Dr. Hari Singh for his guidance and constant supervision as well as for providing necessary information regarding the project and also for their support in completing the project.

I would like to express our gratitude towards our parents and Jaypee University of Information Technology for their kind cooperation and encouragement which helped us in completion of this project.

Our thanks and appreciations also go to our colleague in developing the project and people who have willingly helped us out with their abilities.

TABLE OF CONTENTS

Title   **Page no.**

List of Abbreviations

- NLP: Natural Language Processing

- NLTK: Natural Language Toolkit

- JSON: JavaScript Object notation

- AI: Artificial Intelligence

- ANN: Artificial Neural Network

- SVM: Support Vector Machine

- API: Application Programming Interface

- MDM: Mobile Device management

List of Figures

ABSTRACT

Social media have received more attention nowadays. Public and private opinion about a wide variety of subjects are expressed and spread continually via numerous social media. Twitter is one of the social media that is gaining popularity. Twitter offers organizations a fast and effective way to analyze customers' perspectives toward the critical to success in the market place. Developing a program for sentiment analysis is an approach to be used to computationally measure customers' perceptions. This paper reports on the design of a sentiment analysis, extracting a vast amount of tweets. Python is used in this development along with various modules such as Tweepy, numpy, pandas and Textblob. Results classify customers' perspective via tweets into positive and negative, which is represented in a pie chart and tabular form.

Chapter 1: Introduction

In this chapter we have discussed about the general idea around this project, specified the problem statement, objectives, methodology and the general organization of the project.

1.1 Introduction:

Sentiment Analysis is a machinery based method of interpreting text and crucial the feelings of the text into good, bad or neutral. Performing arts Sentiment Analysis on Twitter knowledge will facilitate corporations acquire qualitative insights to know however folks are talking concerning their whole. With over thirty million active users, causation daily average of five hundred million tweets, Twitter has become one in every of the highest social media platforms for news, data, and interaction with brands and cogent figures round the world. Therefore, it's no surprise that corporations contemplate this micro blogging platform a necessary channel for his or her selling strategy and to supply client service. Twitter permits businesses to succeed in a broad audience and connect with customers while not intermediaries. On the drawback, it is damage a whole's name if negative content concerning the brand suddenly goes infectious agent – you would possibly find yourself with associate surprising PR crisis on your hands. this is often one in every of the explanations why social listening that's observation spoken communication and feedback in social media ― has become an important method in social media selling.

Monitoring Twitter permits corporations to know their audience, keep it up prime of what's being said concerning their whole and their competitors, and see new trends within the trade. However, once it involves analyzing Twitter knowledge, quantitative metrics just like the range of mentions or retweets don't seem to be enough to induce a full image of a scenario. What counts is having the ability to understand the significance of these

mentions. Are users talking completely or negatively a few product or some topic? And that's precisely what sentiment analysis determines. It provides qualitative insights on what's being aforesaid a few topic or whole.

In this project, we'll take a more in-depth check up on sentiment analysis and the way you'll be able to use it to research Twitter knowledge. We'll have a depth practice of the complete method, from recommending some tools to gather information from Twitter, to obtaining you up with the steps involved and running with Twitter sentiment analysis program.

Definition:

Sentiment analysis (a.k.a opinion mining) is that the machine-controlled method of distinctive and extracting the subjective data that written language. this could be either an opinion, a judgment, or a sense a few specific topic or subject. the foremost common variety of sentiment analysis is termed 'polarity detection' and consists in classifying a press release as 'positive', 'negative' or 'neutral'.

For example, allow us to take this sentence: "I don't realize the app useful: it's extremely slow and perpetually crashing". A sentiment analysis model would mechanically tag this as Negative. A sub-field of tongue process, sentiment analysis has been obtaining plenty of attention in recent years thanks to its several exciting apps in a very sort of fields, starting from business to political studies.

Natural Language Processing: natural language processing could be a field in machine learning with the flexibility of a pc to know, analyze, manipulate, and doubtless generate human language.

Some options of NLP:

• Data Retrieval (Google finds relevant and similar results).

• Computational linguistics (Google Translate interprets language from one language to another).

• Sentiment Analysis (Hater news provides U.S. the sentiment of the user).

• Spam Filter (Gmail filters spam emails separately).

• Auto-Predict (Google Search predicts user search results).

• Auto-Correct (Google Keyboard and grammatically correct words otherwise spelled wrong).

• Speech Recognition.

(Natural Language Toolkit) NLTK: NLTK could be a well-liked ASCII text file package in Python. NLTK provides all common natural language processing Tasks.

Thousands of text documents are processed for sentiment (and different options as well as named entities, topics, themes, etc.) in seconds, compared to the hours it might take a team of individuals to manually complete a similar task.

Thanks to sentiment analysis, corporations will perceive the name of their whole. By analyzing social media posts, product reviews, client feedback, or NPS responses (among different sources of unstructured business data), they'll bear in mind of however their customers feel concerning their product. they'll additionally track specific topics and find relevant insights on however folks are talking concerning those topics.

Sentiment analysis is especially helpful for social media observation as a result of it goes on the far side metrics that concentrate on the quantity of likes or retweets, and provides a qualitative purpose of read. Let's say a corporation has simply launched a replacement

product feature and you notice a pointy increase in mentions on Twitter. However, receiving plenty of mentions doesn't essentially mean a decent factor. Are customers tweeting additional as a result of their expressing good items concerning this new product feature? Or, are customers really whiney concerning the feature having many bugs? Performing arts Twitter sentiment analysis is a wonderful thanks to perceive the tone of these mentions and procure period of time insights on however users are perceiving your new product.

1.2 Problem Statement:

The main goal of the project is to perform sentimental analysis on the tweets of a particular user and i.e. determine whether the sentiments/feelings associated with a particular tweet are positive, negative or neutral. Also, to perform various kinds of graphical analysis in the data i.e. subjectivity, no, of likes, retweets etc.

1.3 Objectives:

The main objectives of this project embody crucial the feelings related to the varied tweets and obtaining and basic graphical analysis of assorted tweet attributes over an amount of your time. This could be useful in crucial the opinion of an outsized quantity of individuals.

1.4 Methodology:

Methods of Sentiment Analysis:

• Information assortment

Consumers sometimes specific their sentiments on public forums just like the blogs, discussion boards, product reviews also as on their personal logs – Social network sites like Facebook and Twitter. Sentiments are asserted in numerous method, with totally contrasting terminology and meaning of writing, creating the info immense and scrambled. Manual analysis of sentiment information is nearly not possible. Therefore, special programming languages like 'R' are accustomed method and analyze the info.

• Text Preparation

Preparing the text is however anyway separating the extricated data before examination. It includes differentiating and deleting non-text content and content that is extraneous to the world of study from the information that is present.

1. Analysing the sentiment

At this stage, every sentence of the review is checked for sound judgment basis. Sentences with subjective expressions are kept and tweets that conveys objective expressions are discarded.

2. Classification

Sentiments are broadly speaking classified into 2 things i.e positive and negative. At this stage of sentiment analysis methodology, every tweet detected is assessed into positive, negative and neutral.

3. Results

The main plan of sentiment analysis is to convert JSON text into a pandas data frame. On completing the project, we will visualize the data using matplotlib and will display the sentiments.

1.5 Organization:

"Sentiment Analysis is outlined as a scientific analysis of on-line expressions ". There are five steps to research sentiment information and here's the graphical illustration of the methodology to try and do constant.
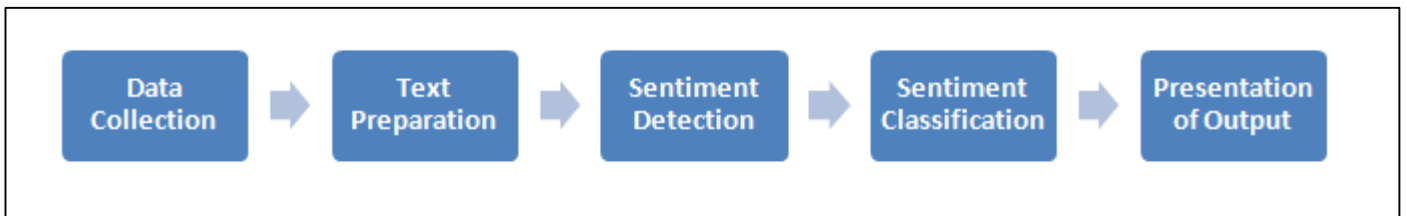

Fig 1: Steps of sentiment analysis

Chapter 2: Literature Survey

Here, in this part, we have examined about the different segments utilized in our venture and we have referenced the separate research paper from where the content has been cited. The different segments incorporate Opinion mining, Twitter, smaller scale blogging, python, online life and ways to deal with wistful examination.

2.1 Opinion Mining

Assessment mining alludes to the expansive space of language process, content mining, etymology, that includes the procedure investigation of estimations, sentiments and feelings communicated in content [8]. In another view, read or point upheld feeling as opposed to reason is normally conversationally named as a notion [8]. Subsequently, transfer to steady for assessment mining or supposition examination. [9] express that supposition mining has a few application areas just as bookkeeping, law, inquire about, redirection, training, innovation, legislative issues, and selling. In prior days a few web based life have provided internet clients path to distribute their position and feelings [10].

2.2 Twitter

Twitter might be a widely used ongoing social website that grants clients to express their views alluded to as tweet that are restricted to one hundred forty characters [11]. Clients compose tweets to communicate their sentiments about numerous things in regards to their everyday life. Twitter is a stage for the retrieval of broad feelings on explicit point [9,10]. Many of the tweets is utilized in light of the fact that the essential corpus for slant investigation, that alludes to the use of assessment mining or language process [1]. Twitter, with enormous clients and million messages for each day, has immediately turned into an important quality for links to investigate their name and types by removal

and tearing down the slope of the tweets by the entire group of people concerning their product. [2] showed that, from the internet based life results in sentiments with the large progress of the globe wide internet, big volume of conclusion messages inside the style of tweets, audits, web journal or any talk group and gatherings are available for consideration, hence making the globe wide net the fastest.

2.3 Microblogging with E-commerce

A stage like Twitter is similar to a standard online journal stage just single posts are short [13]. Twitter has limited for a little scope of words that are meant for the short transmission of info. or trade of supposition [7]. In any case, less business or large association are commencement to the capability of microblogging as an internet business selling device [3]. However, microblogging stage possesses been created numerous years' energy for advancing remote exchange site by utilizing an outside microblogging stage as Twitter selling [3].The moment of sharing, intelligent, network situated alternatives are hole A web based business, propelled a {brand new} brilliant recognize that it are regularly indicated that microblogging stage has empowered firms do mark picture, item imperative deals channel, improve item deals, visit customer for a fair connection and diverse business exercises concerned [2,3] [14]. [13] stated, truth be told, the organizations delivering such product have begun to study big online journals to start a method for regular interpretation for a product. Constantly these organizations render user responses and answer to users on social sites [14].

2.4 Social Media

[15] Mentions an online life as an association of net-put together applications that make with respect to the thoughtful and innovative establishments of Web2.0 that is permitted to make and trade of client created substance. In an enormously exchange of Internet beginning, realizes that a number of web clients is growing and kept on expanding with

online life by the whole time spent on smartphones and net based life in the America across Personal computer hyperbolic by 37% to one hundred twenty-one billion minutes in 2012, contrasted with eighty-eight billion min. in 2011. On the other side, organizations utilize person to person communication locales to search out and speak with buyers, business are frequently incontestable mischief to profitability brought about by interpersonal interaction [17]. As media are frequently report so just to the general population, it will harm individual information to showed inside the social world [11].

Despite what might be expected, [18] referenced that the upsides of participating in online life have gone on the far side only social sharing to make association's name and produce in profession openings and budgetary monetary profit. also, [15] referenced that the web-based social networking is furthermore being utilized for advertising by firms for advancements, experts for watching out, selecting, social learning on-line and electronic trade. Electronic trade or E-business alludes to the procurement and closeout of items or administrations on-line which may through online networking, such has Twitter that is helpful in view of its 24-hours accessibility, simple customer administration and world reach [19].

Among the clarifications of why business will in general utilize a great deal of internet based life is for acquiring knowledge into customer conduct propensities, showcase insight and blessing an opportunity to find out about customer survey and discernments.

2.5 Twitter Sentiment Analysis

The notions are frequently found inside the remarks or tweet to deliver supportive pointers for a few very surprising capacities [20]. Additionally, [12] and [36] unequivocal that a conclusion is frequently characterized into 2 groups, that is negative and positive words. Assumption Analysis might be a language procedure strategy to measure a communicated conclusion or sentiment inside an assortment of tweets [8]. Feeling examination alludes to the last system to remove extremity and judgment from etymology direction that alludes to the quality of words and extremity content or expressions [19].

There have been 2 fundamental methodologies for separating feeling precisely that are the lexicon methods and Artificial Intelligence methods [19-23].

1. Lexicon based

Lexicon method assemble utilisation of already mentioned rundown of texts any place each text is identified with a specific slant. The vocab ways are different as per the settings during which they were made and include hard direction for an archive from the etymology direction of writings or expressions inside the reports [19]. Also, [24] express that a dictionary-supposition is to discover word-conveying conclusion inside the set thus to foresee assessment communicated in the content. [20] has demonstrated the vocabulary ways that have a fundamental worldview which are:

i. Pre-interpret each tweet, post by remove accentuation

ii. Introduce a total extreme score (s) equivalent zero - > s=0

iii. Check in the event that token is blessing in a very wordbook, at that point

If that token is sure, s will be certain (+)

Else if it is negative, s will be less than zero (-)

iv. Approve the entire polar value of twitter tweet.

On the off chance that s > edge, tweet post as positive

In the event that s < edge, tweet post as negative

Be that as it may, [21] featured one preferred position of inclining based philosophy, is that it's the ability to adjust and make prepared models for explicit capacities and settings. In differentiation, partner degree openness of named information and accordingly the low importance of the methodology of most recent information that is cause naming information might be costly or perhaps safeguard for a couple of errands [21].

2. AI based Approach

AI ways ordinarily regard administered arrangement approaches any place feeling recognition is surrounded as a twofold that are sure and negative [24]. This approach needs marked information to mentor classifiers [21]. This methodology, it becomes evident that parts of the local setting of a word should be mulled over like negative (for example Not wonderful) and increase (for example frightfully delightful) [19].

In any case, [20] demonstrated a fundamental worldview for produce an element vector is:

i. Affix a locale of voice tag to each twitter tweet.

ii. Gather entire descriptor for whole twitter tweet.

iii. Fabricate a famous word combination made out of most noteworthy words.

iv. Explore the entirety of tweet inside training example to make the accompanying:

• Assortment of +ve texts

• Assortment of –ve texts

• Existence, nonattendance or recurrence of every word [19] gave some case of switch nullification, refutation only to invert the extremity of the dictionary: regularly evolving lovely (+3) into not staggering (- 3). extra models:

She isn't tremendous (6-5=1) anyway not horrendous (- 6+5=-1) either. For this situation, the nullification of an intensely negative or positive value mirrors a blended point of view that is appropriately caught inside the moved worth. Be that as it may, [21] has referenced the restriction of AI based way to deal with be extra fitting for Twitter than the lexical principally based approach.

Moreover, [20] communicated that AI ways will produce a set assortment of the chief frequently occurring in style words that named a number cost for the benefit of the recurrence of the word inside the Twitter.

2.6 Techniques of Sentiment Analysis

The etymology thoughts of substances separated from tweets is wont to live the general connection of a lot of elements with a given assessment extremity [12]. Extremity alludes to the chief fundamental sort, that will be that if a book or sentence is certain or negative [25]. Be that as it may, supposition examination has strategies in conveyance extremity, for example:

1. Natural Language Processing

NLP strategies are upheld AI and especially applied math discovering that uses a general learning algorithmic guideline joined with an outsized example, a corpus, of information to get familiar with the establishments [24]. Opinion examination has been taken care of as an etymological correspondence process meant normal language handling, at a few degrees of unpleasantness. Extending from being an archive level characterization task

[17], it's been taken care of at the sentence level [18] and extra as of late at the expression level [13]. Normal language preparing might be a field in registering that includes making PCs get that implies from human language and contribution as the most straightforward method for interfacing with the $64000 world.

2. Case-based Reasoning

CBR is one in everything about procedures available to actualize conclusion investigation. CBR is comprehended by reviewing the past with progress settled issues and utilize similar answers for resolve this firmly associated issues. [15] known some of the advantages of abuse CBR that CBR needn't bother with a specific area model then incitement turns into an assignment of social occasion care chronicles and CBR framework will learn by accomplishment new data as test cases. Furthermore, the utilization of data procedures assembles the upkeep of enormous sections of information simpler [15].

3.Artificial Neural Network

[13] Mentioned that ANN or alluded to as neural system might be a maths method that joins and interleaves group of counterfeit neurons. It will strategy data abuse the associations way to deal with calculation.

4. Support Vector

SVM is used to discover the feelings of twitter post. [10] communicated SVM is in a situation to concentrate and investigate to get up to (70%-81.3%) of exactness upon check case. [29] gathered instructing information from 3 totally unique Twitter notion recognition sites that fundamentally utilize some pre-manufactured assessment dictionaries to name each tweet as positive or negative. abuse SVM prepared from these

vociferous labelled information, they acquired eighty-one.3% in opinion grouping precision.

2.7 Application Programming Interface

Speculative chemistry API achieves higher than the others as far as the standard and furthermore the measure of the extricated elements. As time passed the Python Twitter API is made by gathered tweets [30]. Python will precisely compute recurrence of texts being retweeted every one hundred second, arranged the most noteworthy 200 texts upheld there-tweeting recurrence, and keep them inside the chose data. Since the Python TwitterAPI exclusively encased Twitter posts for the preeminent ongoing 6 days, gathered the data required to be keep during a very surprising data [14].

2.8 Python

It was founded by Guido Van Rossum in Netherlands, 1989 that has been open in 1991.It's a programing language that is open and settles a workstation drawback that is giving a direct gratitude to work out an answer. [22] referenced that Python might be referred to as a scripting language. Besides, [22] and [23] moreover upheld that really Python might be a just portrayal of language because of it might be one composed and run on a few stages. also, [24] referenced that Python might be a language that is pleasant for composing a model because of Python is a littler sum time exceptional and managing model gave, differentiation with various programming dialects.

Numerous specialists are voice correspondence that Python is practical, notably for an extravagant venture, Python is fitting for interpersonal companies or news steaming comes that most perpetually are an electronic which is driving a huge data. [34] gave the clarification that because of Python will deal with and deal with the memory utilized. Other than Python makes a generator that licenses partner tedious technique for things,

each thing in turn and grant program to snatch supply data each thing in turn to go each

through the complete procedure chain.

Chapter 3: System Development

In this chapter we have given a basic idea about NLP which is the domain under which our project lies. We have also mentioned about the language used and platform. We discussed about the different modules in this project such as tweepy, numpy, pandas.

3.1 NLP

Natural Language process (NLP) is that the intersection of applied science, Linguistics and Machine Learning that's attached the interaction between computers and humans in tongue.
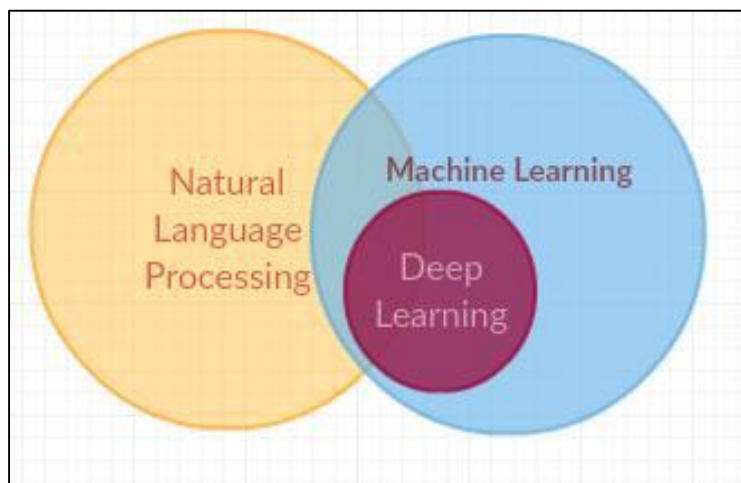

Fig 2: Venn diagram for NLP

NLP is far toward empowering PCs to grasp and deliver human idiom. Uses of informatics systems are used in separating of text, machine interpretation and Voice Agents like Alexa and Siri. Informatics is one amongst the fields that are profited from

the advanced methodologies in Machine Adapting, significantly from Profound Learning methods.

Regular idiom making ready methodology utilize the characteristic dialect tool cabinet for creating the principle organize in python tasks to figure with human dialect information. this is often easier to-use by giving the interfaces to a minimum of one than forty corpora and lexicon resources, for portrayal, for half passages sentences and to urge the words in its distinctive frame Marking, parsing, and gloss thinking for current reasoning quality basic idiom handling libraries, and for dynamic discourse. The NLTK will utilize a huge instrument space and can create some facilitate for people with the complete basic idiom taking care of system. this can assist people with "part sentences from sections, to half up words, seeing the syntactical segments of these words, denoting the elemental subjects, doing this it serves to your machine by acknowledging the most factor to the substance.

3.2 Platform Used

Windows 10

Windows ten is outlined because the Microsoft that works with the actual framework for PCs, tablets, and inserted gadgets etc. Microsoft discharged Windows ten is follow-up to Windows eight. it had been aforesaid on Gregorian calendar month that the window ten are invigorated rather than discharging it and framework as a successor.

When the window ten is chosen or received will be updated by inheritance squarely from window seven, eight or window ten. While not activity meddling and also the framework design methodology. For maintenance shoppers run the windows ten that helps in exchanging the applying on the past software package and setting to window 10. Shoppers pickup and fill or refresh window ten. With the assistance of window refresh partner window ten will be redesigned to physically begin associate degree overhaul for Windows.

Windows ten is employed to focus on add capacities that through which IT offices allows to utilize mobile phones the board (MDM) programming to anchor and management gadgets helps in running operating framework. For given boarding programming as an example, Microsoft Framework Centre Arrangement Chief. Microsoft Windows ten is employed for varied validation advances, as an example, good cards and tokens. Further, Windows Hi has the biometric verification to Windows ten, wherever shoppers will sign on with a novel finger impression, or facial acknowledgment.

The framework is employed to include virtualization-based security tools, as an example, Secluded shopper Mode, Windows Safeguard Gizmo Watch and Windows shielded Qualification Monitor. The Windows ten is employed to stay the highlights of explicit data, procedures and shopper certifications separated attempting to resolve the matter from any strike. Windows ten is newer version for Bit Locker secret writing to substantiate data between clients' gadgets, reposting instrumentation, messages and cloud administrations. Windows eight came up with the new plan and gave touch-empowered motion driven UI like those on cell phones and tablets, however there wasn't abundant interpretation of well to customary work space and digital computer PCs, significantly in business settings. In Windows ten, Microsoft venture to deal with this issue and totally different behavior of Windows eight, as an example, associate degree absence of massive business neighborly highlights. The declaration of Windows ten in Gregorian calendar month 2014 from Microsoft was created and window business executive was made that point. There was the discharge from Microsoft to Windows ten by seeing the full population in Gregorian calendar month 2015. then shoppers discovered that Windows ten is cordial than Windows eight as a result of it had been additional typical interface, that echoes the work space partaking format of Windows seven. The Windows ten consecrate Refresh, that clad in August 2016, created some modifications to the assignment bar and start Menu. It to boot bestowed program augmentations in Edge and gave client's access to Cortana on the bolt screen. In Apr 2017, Microsoft discharged the

Windows ten manufacturers Refresh, that created Windows hello facial acknowledgment innovation faster and enabled shoppers to spare tabs in Microsoft Edge to examine later.

The Windows ten fall manufacturers refresh appeared in Gregorian calendar month 2017, adding Windows Safeguard journey monitor to secure against zero-day assaults. The refresh likewise enabled shoppers and IT to place applications running out of sight into vitality productive mode to safeguard battery life and enhance execution.

3.3 Python

Python is termed as taken, object-oriented, high-level artificial language with dynamic linguistics. The high-level is formed in information structures, together with dynamic writing and binding, that helps in creating it terribly engaging for fast Application Development. It helps in scripting languages wherever the parts are along. Python is termed as terribly straightforward, straightforward to be told and has straightforward syntax. It helps in reducing the value. Python supports packages which inspires program code apply. The in depth commonplace library is out there in supply and might be freely distributed.

Often, programmers like committal to writing in Python as a result of it provides productivity. Written material testing and debugging cycle is extremely quick. Debugging is extremely straightforward in python. Whenever the interpreter finds a slip it generates the exception. If this doesn't happen then interpreter prints a stack trace. A supply level computer program helps in examination of native and international variables, analysis of capricious expressions, setting breakpoints, stepping through the code a line at a time, and so on. The quickest thanks to rectify a program is by adding few print statements to the supply.

3.4 Modules Used

3.4.1 Tweepy

Python is pleasant language for a wide range of things. Awfully dynamic engineer network makes a few libraries that broaden the language and construct it simpler to utilize fluctuated administrations. one in everything about libraries is tweepy. Tweepy is publicly released, facilitated on GitHub and enables Python to talk with Twitter stage and utilize its API.

Installing tweepy is straightforward and simple and can be installed from GitHub:

```
puke clone https://github.com/tweepy/tweepy.git
python setup.py install
Or exploitation straightforward install:
pip install tweepy
```

Fig 3: Installing tweepy

Either means provides you with the newest version.

3.4.1.1 Authentication Tweepy

Tweepy supports accessing Twitter via Basic Authentication and therefore the newer methodology, OAuth. Twitter has stopped accepted Basic Authentication thus OAuth is currently the sole thanks to use the Twitter API.

Here may be a sample of a way to access the Twitter API exploitation tweepy with OAuth:

```
Type "help", "copyright", "credits" or "license()" for more information.
>>> import tweepy
 # client keys and access tokens, used for OAuth c c
access_token = 'z00Xy9AkHwp8vSTJ04L0'
access_token_secret = 'A1cK98w2NXXaCWMqMW6p'
 # OAuth method, exploitation the keys and tokens
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)
 # Creation of the particular interface, exploitation authentication
api = tweepy.API(auth)

# Sample methodology, accustomed update a standing
api.update_status('Hello Python Central!')
```

Fig 4: Accessing Twitter API using OAuth

The results of this code is:

Ahmet Novalic @elahmo   6s
Hello Python Central!
Expand

Fig 5: Updated status using Twitter API.

The major distinction among Basic and OAuth authentication are the consumer and access keys. You are only required to provide a username and a passkey for accessing using basic authentication, but from 2010 since OAuth became

mandatory by twitter, the method has become a lot more cumbersome. A twitter app must be created at dev.twitter.com.

OAuth could be a bit a lot of difficult at the start than Basic Authentication, since it needs a lot of effort, however the advantages it offers are terribly lucrative:

• Tweets is tailor-made to acquire a string that recognizes the application which was utilized.

• The user password is not provided making it safer.

• It's simpler to deal with the authorizations, for instance a gathering of tokens and keys is created that exclusively allows perusing from the courses of events, in this way just in the event that someone acquires those certifications, he/she won't have the option to compose or send direct messages, limiting the opportunity.

• The app does not answer on a secret word, therefore regardless of whether

the client transforms it or not the app will function.

Subsequent to signing in to the gateway, and navigating to Apps another app can be made which is able to offer the required information for act with Twitter API.

This is a screen that has all of the information required to speak to Twitter network. it's vital to notice that by default, the app has no access to direct messages, thus by attending to the settings and dynamical the suitable choice to "Read, write and direct messages", you'll be able to modify your app to possess access to each Twitter feature.

OAuth settings

Your application's OAuth settings. Keep the "Consumer secret" a secret. This key should never be human-readable in your application.

Access level	Read, write, and direct messages
	About the application permission model
Consumer key	PHG9tkvUpVdCLHuluiQFAA
Consumer secret	dqpNZnLTwteX1YGnQ0VQ3Pv2up6ensEFeaS8MnQDE
Request token URL	https://api.twitter.com/oauth/request_token
Authorize URL	https://api.twitter.com/oauth/authorize
Access token URL	https://api.twitter.com/oauth/access_token
Callback URL	None

Your access token

Use the access token string as your "oauth_token" and the access token secret as your "oauth_token_secret" to sign requests with your own Twitter account. Do not share your oauth_token_secret with anyone.

Access token	38744894-0TBISZIcuDE5Sm1VI6VqZXGVYH9Yjn63e9ZM8v7ei
Access token secret	g6ElhezlPulcrPzM1jDyqqjXMH25EDeJncHaxvQeu0
Access level	Read, write, and direct messages

Recreate my access token

Fig 6: Twitter application settings

3.4.1.2 Twitter API

Tweepy gives entry to the well mentioned Twitter API. With tweepy, it's attainable to induce any object and use any methodology that the official Twitter API offers. For instance, a User object has its documentation at https://dev.twitter.com/docs/platform-objects/users and following those tips, tweepy will get the suitable data.

```
# Creates the user object. The me() methodology returns the user whose authentic
ation keys were used.
user = api.me()
print('Name: ' + user.name)
print('Location: ' + user.location)
print('Friends: ' + str(user.friends_count))
```

Fig 7: Accessing user entities

Main Model categories within the Twitter API are Tweets, Users, Entities and Places. Access to every returns a JSON-formatted response and traversing through data is extremely straightforward in Python.

The above set of statements in figure 7 gives the following result:

Name: Robin Singh

Location: India

Friends: 55

3.4.1.3 Tweepy Streaming API

One of the most usage cases of tweepy is observance for tweets and doing actions once some event happens. Key part of that's the Stream Listener object, that monitors tweets in real time and catches them. Stream Listener has many ways, with on_data() and on_status() being the foremost helpful ones. Here may be a sample program that implements this behavior:

```
>>> class StdOutListener(StreamListener):
    ''' Handles information received from the stream. '''
    def on_status(self, status):
        # Prints the text of the tweet
        print('Tweet text: ' + standing.text)
        # There are several choices within the standing object,
        # hashtags will be terribly simply accessed.
        for hashtag in standing.entries['hashtags']:
            print(hashtag['text'])
come back true

    def on_error(self, status_code):
        print('Got a slip-up with standing code: ' + str(status_code))
come back True # To continue listening
    def on_timeout(self):
        print('Timeout...')
come back True # To continue listening

if __name__ == '__main__':
hearer = StdOutListener()
    auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
    auth.set_access_token(access_token, access_token_secret)

    stream = Stream(auth, listener)
    stream.filter(follow=[38744894], track=['#pythoncentral'])
```

Fig 8: Stream Listener methods

So, this program contains a Stream Listener enforced and therefore the code is about up to use OAuth. The Stream object is made, that uses that hearer as output. Stream, being another necessary object in tweepy conjointly has several

25

ways, during this case filter() is employed with parameters passed. "Follow" may be a list of followers whose tweets are monitored, and "track" may be a list of hashtags which is able to trigger the Stream Listener.

In this example, we've got used my user ID to follow and therefore the #pythoncentral hashtag as a condition. when running the program and tweeting this status:
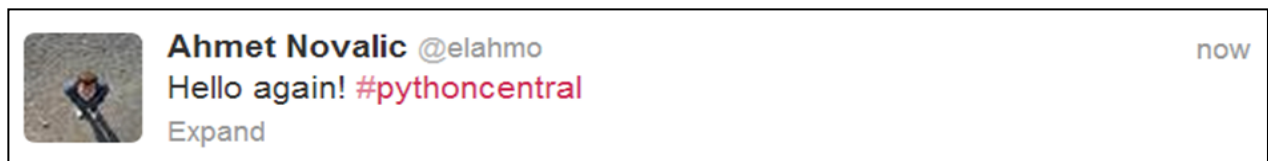

Fig 9: Tweeting a status

The program virtually instantly catches the tweet, and calls the on_status() methodology, that produces the subsequent output within the console:

Fig 10: Result of on_status()

Besides printing the tweet, within the on_status() methodology there are some further things that illustrate the amount of potentialities which will be finished the tweet data:

```python
# There are several choices within the standing object,
# hashtags may be terribly simply accessed.
for hashtag in standing.entities['hashtags']:
print(hashtag['text'])
```

Fig 11: Prints all the hashtag values

This code traverses through entities, picks the "hashtags" one and for every hashtag the tweet contains, it prints its value. This is often simply a sample; a whole list of tweet entities can be found on twitter's website.

To sum up, tweepy may be a nice ASCII text file library that gives entry to the TwitterAPI for Python. Despite the fact that the documentation for tweepy might be somewhat rare and doesn't have a few models, the undeniable reality is that it intensely relies upon the Twitter API, that has wonderful documentation, makes it in all probability the most effective Twitter library for Python, particularly once seeing the Streaming API help. Alternative modules similar to python-twitter give several options also, however the tweepy has most alive association and large carry out to the implementation in the previous years.

3.4.2 TextBlob

TextBlob is a python module and provides a simplistic API to use its methods and carry out NLP tasks. TextBlob aims to provide access to common text-processing operations through a familiar interface. You can treat TextBlob objects as if they were Python strings that learned how to do Natural Language Processing.

A nice thing regarding TextBlob is that it is similar to strings. So, you can use it the same way just as strings. Below, I have performed some of the simple tasks. The below code is used to show that TextBlob is very same to string and the syntax is just to explain things right.

```python
string1 = TextBlob("Analytics")

string1[1:5]    ### extracting 1 to 5 letters
TextBlob("naly")

string1.upper()  ## to upper case the entire text
TextBlob("ANALYTICS")

string2 = TextBlob("Vidhya")

## concat two sentences similar as python

string1 + " " + string2
TextBlob("Analytics Vidhya")
```

Fig12: Similarity between TextBlob and Strings

28

3.4.2.1 Setting up TextBlob

Installation of TextBlob in your system in a simple task, all you need to do is open anaconda prompt (or terminal if using Mac OS or Ubuntu) and enter the following command:

pip install -U textblob

For the uninitiated – practical work in Natural Language Processing typically uses large bodies of linguistic data, or corpora. To download the necessary corpora, you can run the following command:

python -m textblob.download_corpora

3.4.2.2 Pros and Cons

Pros:
1. Since, it is built on the shoulders of NLTK and Pattern, therefore making it simple for beginners by providing an intuitive interface to NLTK.
2. It provides language translation and detection which is powered by Google Translate (not provided with Spacy).

Cons:
1. It is little slower in the comparison to spacy but faster than NLTK. (Spacy > TextBlob > NLTK)
2. It does not provide features like dependency parsing, word vectors etc. which is provided by spacy.

3.4.2 WordCloud

It is a visual representation tool used for describing text data wherein the size of individual word tells about the significance of word. Huge literary information focuses can be featured utilizing a word cloud. Word Clouds are generally utilized for examining information from social networking websites. To generate Word Cloud in python we are required to have matplotlib, pandas and WordCloud.

WordCloud can be a little tricky to install. If you only need it for plotting a basic wordcloud, then "pip install wordcloud'' would be sufficient. However, the latest version with the ability to mask the cloud into any shape of your choice requires a different method of installation as below:

```
git clone https://github.com/amueller/word_cloud.git
cd word_cloud
pip install
```

Fig13: Commands to install Word Cloud

Benefits:
1. Assessing client and agent assessment.
2. Knowing about new SEO words to aim.

Disadvantages:
1. WordClouds are not suitable for every situation.
2. Information ought to be advanced for setting.

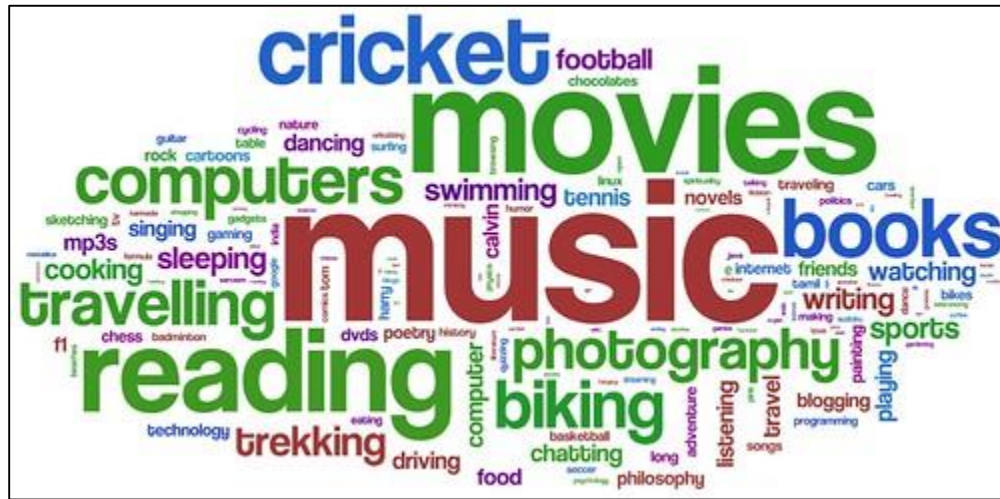The following is an example of word cloud:

Fig14: Word Cloud of hobbies

This word cloud is formed by different hobby words. As you can see the word "music" is the biggest in size hence it has been used most frequently as a hobby followed by movies, reading and so on. The word music is appearing as most frequently but it could change depending upon the dataset that the particular person is using.

3.4.3 Regular Expression(RegEx)

Regular expression is a set of special characters that is used to find a string or a substring that is provided by the user in a pattern or similar kind of variable. I have used the sub method of the re to clean up the tweets of unnecessary information and get only the individuals words. Regular expression is mostly used in Unix environment.

The regular expression modules provide support for Perl like regular expression in python. Whenever an error is present the re module raises an exception re.error.

3.4.3.1 RegEx Functions

The re module offers a set of functions that allows us to search a string for a match:

Table 1: RegEx functions

Function	Description
findall	Returns a list containing all matches
search	Returns a Match object if there is a match anywhere in the string
split	Returns a list where the string has been split at each match
sub	Replaces one or many matches with a string

3.4.3.2 Metacharacters

Metacharacters are characters with a special meaning:

Table 2: Metacharacters

Character	Description	Example
[]	A set of characters	"[a-m]"
\	Signals a special sequence (can also be used to escape special characters)	"\d"
.	Any character (except newline character)	"he..o"
^	Starts with	"^hello"
$	Ends with	"world$"
*	Zero or more occurrences	"aix*"
+	One or more occurrences	"aix+"
{}	Exactly the specified number of occurrences	"al{2}"
|	Either or	"falls|stays"

Chapter 4: Performance Analysis

Here we have discussed about the extraction of tweets and storing the attribute fields of the tweets into a data frame using pandas and numpy module. We have also applied various mathematical functions to count the number of tweets, likes, retweets etc.

4.1 Tweets extraction

Now that we've created a perform to setup the Twitter API, we will use this perform to make an "extractor" object. After this, we'll use tweepy's command:

```
>>> extractor.user_timeline(screen_name, count)
```

Fig 15: Extracting tweet count from user

This extracts from screen_name's user the amount of tweets count.

As it is mentioned within the title, I've chosen @narendramodi as the user to extract knowledge for a posterior analysis.

```
extractor = twitter_setup()
tweets = extractor.user_timeline(screen_name={};NarendraModi{}, count = 200)
print({}number of tweets extracted: .\n{}.format(len(tweets)))

print({}5 recent tweets:\n{})
for tweet in tweets[:5]:
print(tweet.text)
print()
```

Fig 16: Extracting 5 recent tweets

Output:

```
number of tweets extracted: two hundred.
5 recent tweets:
```

Fig 17: Result of tweet count and extraction

The 5 recent tweets are:

Had a beautiful interaction with IFS officer trainees. We tend to mentioned a large vary of problems. Urged the officers to… https://t.co/qtnNEHVSob

Sharing Associate in Nursing animated video that may teach you the ropes of Shalabhasana. #FitIndia #4thYogaDay https://t.co/zHZn9MsUwG

आप सभी के साथ शलभासन का एक एनिमेटेड वीडियो शेयर कर रहा हूं। https://t.co/YG4DC5MrYK

Saddened by the loss of lives thanks to storms in some elements of the country. Condolences to the bereft families. I p… https://t.co/YeJCdfv6he

Veena Abhyankar foreign terrorist organization from Pune wrote a beautiful letter to ME, mentioning that she has been learning paper-quilling f… https://t.co/D5DOJ4Ts5K

4.2 Making a (pandas) Data Frame

We currently have initial data to construct pandas knowledge Frame, so as to control the information in a very straightforward approach.

Python's perform show plots Associate in nursing output in an exceedingly friendly approach, and therefore the head methodology of an information frame permits us to check the primary five components of the information frame (or the first variety of elements that are passed as Associate in argument).

So, exploitation Python's list comprehension:

```
data = pd.DataFrame(data=[tweet.text for tweet in tweets], columns=[&#039;tweets&#039;])
display(data.head(10))
```

Fig 18: Creating and displaying a data frame.

This will produce following result:

Tweets

0 Had a beautiful interaction with IFS officer t...

1 Sharing Associate in Nursing animated video that may teach you ...

2 आप सभी के साथ शलभासन का एक एनिमेटेड वीडियो शेय...

3 Saddened by the loss of lives thanks to storms in...

4 Veena Abhyankar foreign terrorist organization from Pune wrote a beautiful...

5 My Asian nation visit was historic. It gave ME a grea...

6 यस भ्रमण माफ़ैत भारत-नेपाल सम्बन्धमा नयाँ उर्...

7 मेरो नेपाल भ्रमण ऐतिहासिक रह्यो । यस भ्रमणले म...

8 At the programme in Kathmandu, I reiterated In...

9 I convey the folks of Kathmandu for the memora...

The fascinating half from here is that the amount of information contained in an exceedingly single tweet. If we would like to get knowledge like the creation date, or the supply of creation, we will access the information with this attributes. Let's look at an e.g.

```
>>>

# we tend to print information from the primary tweet:

print(tweets[0].id)
print(tweets[0].created_at)
print(tweets[0].source)
print(tweets[0].favorite_count)
print(tweets[0].retweet_count)
print(tweets[0].geo)
print(tweets[0].coordinates)
print(tweets[0].entities)

# we tend to add relevant knowledge
data['len'] = np.array([len(tweet.text) for tweet in tweets])
data['ID'] = np.array([tweet.id for tweet in tweets])
data['Date'] = np.array([tweet.created_at for tweet in tweets])
data['Source'] = np.array([tweet.source for tweet in tweets])
data['Likes'] = np.array([tweet.favorite_count for tweet in tweets])
data['RTs'] = np.array([tweet.retweet_count for tweet in tweets])

display(data.head(10))
```

Fig 19: Accessing the information inside a single tweet

Output:

	tweets	len	ID	Date	Source	Likes	RTs
0	Had a wonderful interaction with IFS officer t...	140	996034840242749440	2018-05-14 14:29:46	Twitter Web Client	7189	1590
1	Sharing an animated video that will teach you ...	118	995877078791802880	2018-05-14 04:02:52	Media Studio	6971	1878
2	आप सभी के साथ शलभासन का एक एनिमेटेड वीडियो शेय...	83	995876437004517376	2018-05-14 04:00:19	Media Studio	12769	2935
3	Saddened by the loss of lives due to storms in...	140	995724191210106883	2018-05-13 17:55:21	Twitter for iPhone	19028	3679
4	Veena Abhyankar Ji from Pune wrote a wonderful...	140	995690625965142017	2018-05-13 15:41:58	Twitter for iPhone	15904	2908
5	My Nepal visit was historic. It gave me a grea...	139	995260100527898624	2018-05-12 11:11:13	Twitter Web Client	13131	2928
6	यस भ्रमण माफ़ॅत भारत-नेपाल सम्बन्धमा नयाँ उर्...	140	995258851174764545	2018-05-12 11:06:15	Twitter Web Client	8622	1880
7	मेरो नेपाल भ्रमण ऐतिहासिक रह्यो । यस भ्रमणले म...	140	995258545066074112	2018-05-12 11:05:02	Twitter Web Client	9643	2031
8	At the programme in Kathmandu, I reiterated In...	140	995257904067297280	2018-05-12 11:02:29	Twitter Web Client	5457	1357
9	I thank the people of Kathmandu for the memora...	140	995257425044299776	2018-05-12 11:00:35	Twitter Web Client	7096	1639

Fig 20: Attributes of a tweets

Now, we try to find the most liked tweet and number of liked on that tweet. Also, we find the number of characters contained in that particular tweet. Similarly, we also do this for the most retweeted tweet.

38

Code:

```
>>> # we tend to extract the mean of length
mean = np.mean(data['len'])
print("the length's average in tweets: ".format(mean))

the length's average in tweets: 129.065

# we tend to extract the tweet with a lot of FAVs and atom RTs:

fav_max = np.max(data['Likes'])
rt_max = np.max(data['RTs'])

fav = data[data.Likes == fav_max].index[0]
rt = data[data.RTs == rt_max].index[0]

# soap FAVs:
print("the tweet with a lot of likes is: \n".format(data['tweets'][fav]))
print("Number of likes: ".format(fav_max))
print(" character.\n".format(data['len'][fav]))

# Max RTs:
print("The tweet with a lot of retweets is: \n".format(data['tweets'][rt]))
print("Number of retweets: ".format(rt_max))
print(" character.\n".format(data['len'][rt]))
```

Fig21: Most liked and most retweeted tweet.

Output:

```
>>> the tweet with a lot of likes is:
वीरता, दृढ़ता, साहस और देशभक्ति के प्रतीक महान योद्धा महाराणा प्रताप को उनकी जयंती पर सादर नमन।
Number of likes: 33452
95 character.
The tweet with a lot of retweets is:
Urging my sisters and brothers of Mysore to pick out giant number
s nowadays. i might notably wish to decision upon… https://t.co/h
QXpnjnoY2
Number of retweets: 8018
139 character.
```

Fig 22: Result of tweet analysis

39

4.3 Cleaning tweets

The method of re library that has been used for cleaning text is sub.

Syntax

```
re.sub(pattern, repl, string, max=0)
```

Fig 23: Syntax of substitute method

The sub method finds for the given pattern in the string that the user has given and if there are any occurrences it is replaced by the substitute pattern. It then returns the newly altered string.

To clean up the tweets of any hash tags, RT, @ mentions and hyperlinks we have proceeded in the way shown below:

```
def cleantext(text):
        text = re.sub('@[A-Za-z0-9]+', '', text) #Removing @mentions
        text = re.sub('#', '', text) # Removing '#' hash tag
        text = re.sub('RT[\s]+', '', text) # Removing RT
        text = re.sub('https?:\/\/\S+', '', text) # Removing hyperlink

        return text
```

Fig 24: cleaning up the tweets

4.4 Calculation of Polarity and Subjectivity

I have used the TextBlob module to find the sentiment associated with the given tweet of the respective user.

The TextBlob module has a sentiment function which return either the polarity or the subjectivity. Polarity determines whether the statement is positive statement, negative statement or neutral statement whereas subjectivity is personal opinion and objectivity is factual information.

Polarity lies in the range of [-1,0,1] whereas subjectivity lies in the range of [0,1]. I have implemented the below given code to successfully determine the polarity and subjectivity.

```python
# Create a function to get the subjectivity
def getSubjectivity(text):
    return TextBlob(text).sentiment.subjectivity

# Create a function to get the polarity
def getPolarity(text):
    return  TextBlob(text).sentiment.polarity

# Create a function to compute negative (-1), neutral (0) and positive (+1) analysis
def getAnalysis(score):
 if score < 0:
  return 'Negative'
 elif score == 0:
  return 'Neutral'
 else:
  return 'Positive'
```

Fig 25: Polarity and subjectivity methods

4.5 Visualizing the tweets

I have used WordCloud library to determine the most prominently used word in the given set of tweets. Using WordCloud helps one to determine the most frequently used word in the given set of text and enables the person to get a better look and understanding of the vocabulary of the person who has written the sentence.

```python
allWords = ' '.join([twts for twts in df['Tweets']])
wordCloud = WordCloud(width=500, height=300, random_state=21, max_font_size=110).generate(allWords)
plt.imshow(wordCloud, interpolation="bilinear")
plt.axis('off')
plt.show()
```

Fig 26: Generating WordCloud.

Herein, I have joined all the tweets under the Tweets column name in the data frame. This forms the set of text which I have used for my WordCloud. Then, I have used the generate method of the WordCloud library and passed the allWords variable to generate a WordCloud.

Chapter 5: Conclusion

Nowadays, sentiment analysis or opinion mining could be a hot topic in machine learning. We have still so much to find regarding the feelings of corpus of texts terribly accurately thanks to the complexity within the English language. In this project we are tending to specialize in sentiments analysis. There is capability of labor within the range of sentiment analysis with slightly accepted background. For eg. we tend to seen that clients as a rule utilize our site for explicit sorts of watchwords which can be partitioned into a couple of particular classes, to be specific: governmental issues/lawmakers, big names, items/brands, sports person, media and music. Subsequently we will attempt to perform separate feeling investigation on tweets that exclusively have a place with 1 of those categories (for example the training data wouldn't be general anyway explicit to 1 of those classifications) and analyze the outcomes we tend to get if we apply general sentiment analysis on that instead.

Twitter's API is vastly helpful in data processing applications, and may offer large insights into the general public opinion if the Twitter API and large information analytics are a few things you've got more interest in. Twitter API can be used in most of the difficult sentiment gathering, involving people, trends, and social graphs that is very different for the human mind to get.

References

[1] M.Rambocas , and J. Gama, "Marketing Research: *The Role of Sentiment Analysis"*. The 5th SNA-KDD Workshop'11. University of Porto, 2013.

[2] A. K. Jose, N. Bhatia, and S. Krishna, "Twitter Sentiment Analysis". *National Institute of Technology Calicut,* 2010.

[3] P. Lai, "Extracting Strong Sentiment Trend from Twitter". Stanford University, 2012.

[4] Y. Zhou, and Y. Fan, " A Sociolinguistic Study of American Slang," *Theory and Practice in Language Studies*, *3*(12), 2209–2213, 2013.
doi:10.4304/tpls.3.12.2209-2213

[5] M. Comesaña, A. P.Soares, M.Perea, A.P. Piñeiro, I. Fraga, and A.Pinheiro, " Author's personal copy Computers in Human Behavior ERP correlates of masked affective priming with emoticons," C*omputers in Human Behavior*, *29*, 588–595, 2013.

[6] A.H.Huang, D.C. Yen, & X. Zhang, "Exploring the effects of emoticons," *Information & Management*, 45(7), 466–473, 2008.

[7] D. Boyd, S. Golder, & G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on twitter," *System Sciences (HICSS), 2010 ….*
Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5428313

[8] T. Carpenter, and T. Way, "Tracking Sentiment Analysis through Twitter,". *ACM computer survey*. Villanova:VillanovaUniversity, 2010.

[9] D. Osimo, and F. Mureddu, "Research Challenge on Opinion Mining and Sentiment Analysis," *Proceeding of the 12th conference of Fruct association*, 2010, United Kingdom.

[10] A. Pak,and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," *Special Issue of International Journal of Computer Application*, France:Universitede Paris-Sud, 2010.

[11] S.Lohmann, M. Burch, H. Schmauder and D. Weiskopf, "Visual Analysis of Microblog Content Using Time-Varying Co-occurrence Highlighting in Tag Clouds," *Annual conference of VISVISUS.*
Germany: University of Stuttgart, 2012.

[12] H. Saif, Y.He, and H. Alani, "Semantic Sentiment Analysis of Twitter,"
Proceeding of the Workshop on Information Extraction and Entity Analytics on Social Media Data. United Kingdom: Knowledge Media Institute, 2011.

[13] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R.Passonneau,
"Sentiment Analysis of Twitter Data," Annual International Conferences. New York:Columbia University, 2012.

[14] J. Zhang, Y. Qu, J. Cody and Y. Wu, " A case study of Microblogging in the Enterprise: *Use, Value, and Related Issues*," Proceeding of the workshop on Web 2.0., 2010.

[15] G. Kalia, "A Research Paper on Social Madia: *An Innovative Educational Too"*, Vol.1, pp. 43-50, Chitkara University, 2013.

[16] Internet World Start, "Usage and Population Statistic", Retrieved 10 15, 2013 from: http://www.internetworldstats.com/stats.htm

[17] A.M. Kaplan, and M, Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," France: *Paris*, 2010.

[18] Q. Tang, B. Gu, and A.B. Whinston, "Content Contribution in Social Media: *The case of YouTube",* 2nd conference of social media. Hawaii: *Maui, 2012.*

[19] M.Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, " Lexicon- Based Methods for Sentiment Analysis," *Association for Computational Linguistics*, 2011.

[20] M. Annett, and G. Kondrak, "A Comparison of Sentiment Analysis Techniques: *Polarizing Movie Blogs," Conference on web search and web data mining (WSDM)*. University of Alberia: Department of Computing Science, 2009.

[21] P. Goncalves, F. Benevenuto, M. Araujo and M. Cha, "Comparing and Combining Sentiment Analysis Methods", 2013.

[22] E. Kouloumpis, T. Wilson, and J. Moore, "Twitter Sentiment Analysis: *The Good the Bad and the OMG!",* (Vol.5). International
AAAI, 2011.

[23] S. Sharma, "Application of Support Vector Machines for Damage detection in Structure," *Journal of Machine Learning Research, 2008.*

[24] A.Sharma, and S. Dey, "Performance Investigation of Feature SelectionMethods and Sentiment Lexicons for Sentiment Analysis," *Association for the advancement of Artificial Intelligence*, 2012.

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT
PLAGIARISM VERIFICATION REPORT

Date: ....

Type of Document (Tick): B.Tech Project Report

Name: __Robin Singh__ __Department: __CSE____ Enrolment No __161240__

Contact No. ___7018940208___ E-mail. ___robin.122497@gmail.com_____

Name of the Supervisor: ___Dr. Hari Singh_____

Title of the Thesis/Dissertation/Project Report/Paper (In Capital letters): _TWITTER SENTIMENT__
_ANALYSIS_____

UNDERTAKING

I undertake that I am aware of the plagiarism related norms/ regulations, if I found guilty of any plagiarism and copyright violations in the above thesis/report even after award of degree, the University reserves the rights to withdraw/revoke my degree/report. Kindly allow me to avail Plagiarism verification report for the document mentioned above.

Complete Thesis/Report Pages Detail:
- Total No. of Pages = 55
- Total No. of Preliminary pages = 9
- Total No. of pages accommodate bibliography/references = 3

Robin Singh
Digitally signed by Robin Singh
Date: 2020.07.14
21:12:31 +05'30'

(Signature of Student)

FOR DEPARTMENT USE

We have checked the thesis/report as per norms and found **Similarity Index** at ....(%). Therefore, we are forwarding the complete thesis/report for final plagiarism check. The plagiarism verification report may be handed over to the candidate.

(Signature of Guide/Supervisor)   **Signature of HOD**

FOR LRC USE

The above document was scanned for plagiarism check. The outcome of the same is reported below:

Copy Received on	Excluded	Similarity Index (%)	Generated Plagiarism Report Details (Title, Abstract & Chapters)	
	• All Preliminary Pages		Word Counts	
Report Generated on	• Bibliography/Images/Quotes		Character Counts	
	• 14 Words String	**Submission ID**	Total Pages Scanned	
			File Size	

Checked by
Name & Signature   **Hari Singh**   Digitally signed by Hari Singh
Date: 2020.07.15 14:13:26 +05'30'   **Librarian**

Please send your complete thesis/report in (PDF) with Title Page, Abstract and Chapters in (Word File) through the supervisor at plagcheck.juit@gmail.com