**Project Report On**

**“Twitter Sentiment Analysis”**

**SUBMITTED TO SAVITRIBAI PHULE PUNE UNIVERSITY**

**BACHELOR OF ENGINEERING**

**(Computer Engineering)**

**By**

**1.Lala Patle**

**2.Sourabh Shriwas**

**3.Ridima Shukla**

**4.Shivani**

**UNDER THE GUIDANCE OF**

**Miss Shivani Dhotre**

****

**Department of Computer Engineering**

**Dr.D.Y.Patil Institute of Technology Pimpri, Pune-18**

**2018-2019**

****

**Dr.D.Y.Patil Institute of Technology Pimpri, Pune-18**

**Department of Computer Engineering**

**CERTIFICATE**

This is to certify that the project entitled

“Twitter Sentiment Analysis”

Submitted By

Lala K Patle TCOC53

Sourabh B Shriwas TCOC54

Ridima Shukla TCOC52

is a bonafide work carried out by them under the supervision of **Miss Shivani Dhotre**  and it is approved for the partial fulfillment of the requirement of SavitribaiPhule Pune University for the award of the Degree of Bachelor of Engineering(Computer Engineering )

.

Prof. Bhagyashree Bhoyar Dr. Pramod. D. Patil

Internal Guide H.O.D

Dept. of Computer Engg. Dept. of Computer Engg.

Dr.S. N. Mali

Principal

Dr. D. Y. Patil Institute of Technology, Pimpri, Pune

**Signature of Internal Examiner Signature of External Examiner**

**Content**

|  |  |  |
| --- | --- | --- |
| **Sr.No** | **Name** | **Page No** |
| **1** | **Introduction** | **5** |
| **2** | **Objective** | **6** |
| **3** | **Implementation** | **7** |
| **4** | **System requirement** | **10** |
| **5** | **Testing** | **11** |
| **6** | **Advantages & Disadvantage** | **12** |
| **7** | **Future scope** | **14** |
| **8** | **Conclusion** | **15** |
| **9** | **References** | **16** |

**List of table**

|  |  |  |
| --- | --- | --- |
| **Sr.No** | **Name** | **Page No** |
| **1** | **Different results of topic** | **11** |

1. **Introduction**

Sentiment analysis is the task of finding the opinions and affinity of people towards specific topics of interest. Be it a product or a movie, opinions of people matter, and it affects the decision-making process of people. The first thing a person does when he or she wants to buy a product online, is to see the kind of reviews and opinions that people have written. Social media such as Facebook, blogs, twitter have become a place where people post their opinions on certain topics. The sentiment of the tweets of a particular subject has multiple usage, including stock market analysis of a company, movie reviews, in psychology to analyze the mood of people that has a variety of applications, and so on. Sentiments of tweets can be categorized into many categories like positive, negative, neutral, extremely positive, extremely negative, and so on

Users often discuss on current affairs and share their personals views on various subjects via tweets. Out of all the popular social media's like Facebook, Google+, Myspace and Twitter, we choose Twitter because of the following reasons:

* Twitter contains an enormous number of text posts and it grows every day. The collected corpus can be arbitrarily large.
* Twitters audience varies from regular users to celebrities, company representatives, politicians, and even country presidents. Therefore, it is possible to collect text posts of users from different social and interest groups.
* Tweets are small in length, thus less ambiguous
* Tweets are unbiased in nature

Using this social media we built models for classifying "tweets" into positive, negative and neutral classes.

1. **Objective**

The objective of this project is to show how sentimental analysis can help improve the user experience over a social network or system interface.

We can show and analysis some of the following aspects in our project:-

* Classify the searched tweets into positive ,negative and neutral tweets
* Show the pie chart of overall percent of following the types of tweets
* Give a line graph of retweets ,since the retweets are the source from which we know how popular a particular thread is.
* Show the average polarity and subjectivity of a particular topic to know how much a topic is positive or negative and how much a topic is real or fake.

The learning algorithm will learn what our emotions are from statistical data then determine the mood. After that it will change our social interactions accordingly on our social network sites or other interfaces like desktop or system services or web-pages. Suppose you are bored or sad ,in the case of social networks one thing the computer could do is to be more suggestive of things that lighten your mood and change interactions like backgrounds color's ,icons services. The site could automatically try suggesting interactions with people and applications that would help improve the mood, while hiding others that might make it worse. The project aims to implement these in the social network community as well as services and interfaces of our systems, while making our lives better and our experience richer and efficient.

1. **Implementation**

Implementation for this particular project is done with help of modularity. This function are basic blocks which contains particular task . we can understand the overall implementation by understanding this functions:-

**Authentication**

consumer\_key =**'U6gwDbDqNNGFHQy7v9VG7i6iP'** consumer\_secret =**'fmlBWzZbaugotaALNHikv7SHiLgN00StAHUIZ89qAXdGfVy5DP'** access\_token = **''** access\_token\_secret = **''** auth = tweepy.OAuthHandler(consumer\_key,consumer\_secret)  
 auth.set\_access\_token(access\_token,access\_token\_secret)  
 api = tweepy.API(auth)

this code snippet is used for connecting the program to twitters api. To get the tweets from twitter we need a twitter’s developers account . then we create a app in developers options and get our consumer key and access key respectively ,

The tweepy.OAuthandler() function is used to get our authentication .then we use tweepy.API() function to create and api variable to use through out the program.

**analyse( )**

In this function the major task is done , which is getting the tweets & analyzing them. For retrieving tweets we use api.search() function which provides us the tweets but have a maximum limit of 100 tweets . To get more tweets we can simply apply a logic as show below in the code snippet.

**while** len(public\_tweets) < int(b):  
 count = int(b) - len(public\_tweets)  
 **try**:  
 new\_tweets = api.search(q=a, count=count, max\_id=str(last\_id - 1))  
 **if not** new\_tweets:  
 **break** public\_tweets.extend(new\_tweets)  
 last\_id = new\_tweets[-1].id  
 **except** tweepy.TweepError **as** e:  
 **break**

other task of this function is to perform classification of tweets in three categories i.e positive ,negative and neutral tweets. By knowing polarity of particular text we can classilfy them .The range of polarity for classification is as following positive(+1.0) > neutral(0.0) > negative(-1.0) using a for loop we can classify all the tweets that we have retrieved. Below code snippet demonstrates the logic for classification .

**for** tweet **in** public\_tweets:

a=deEmojify(tweet.text)  
 *#print(tweet.favorite\_count)* retweet.append(tweet.retweet\_count)  
 favorite.append(tweet.favorite\_count)  
 *#print(tweet.retweet\_count)* analysis=TextBlob(a)  
 *#print(analysis.sentiment)*

avg\_pol=avg\_pol+analysis.sentiment.polarity  
  
 avg\_sub=avg\_sub+analysis.sentiment.subjectivity  
 **if** analysis.sentiment.polarity > 0:  
 *#print('positive')* lb1.insert(END,a)  
 pt = pt +1  
 **elif** analysis.sentiment.polarity == 0:  
 *#print('neutral')* lb3.insert(END, a)  
 n=n+1  
 **else**:  
 *#print('negative')* lb2.insert(END,a)  
 nt=nt+1

but for knowing the polarity of particular text we need to use another dependency know as textblob this is a NLP based sentiment and phrase analyser . for sentiments two terms are needed to calculated, they are polarity and subjectivity ,using textblob we can easily get those value for particular text using functions as

*analysis.sentiment.polarity*

*analysis.sentiment.subjectivity*

here the analysis is the object of textblob for a particular text. this functions does a lot more operations ,it can be simply summarized as :-

* To clear the interface
* To retrieve the tweets
* To classify the tweets
* To get the percent of three different types of tweets
* To append the results to interface

**deEmojify(input)**

This function is used for the basic purpose i.e to make the tweets clean from emojis and other special characters, we pass the tweets and then get a return text which is text free.

**def** deEmojify(inputString):  
 returnString = **""  
 for** character **in** inputString:  
 **try**:  
 character.encode(**"ascii"**)  
 returnString += character  
 **except** UnicodeEncodeError:  
 returnString += **''  
 return** returnString

**Graphs:**

For displaying the graphs we have two functions graph and graph1, graph1() is used for plotting a retweets graphs for each tweets.

**def** graph1():  
 plt.clf()  
 plt.plot(retweet)  
 plt.ylabel(**'retweets'**)  
 plt.xlabel(**'no of tweets'**)  
 plt.show()

and the graph function is used to plot the pie chart for the types of tweets both this graphs are implemented using matplotlib library the graph() function is as following

**def** graph():  
 plt.clf()  
 labels = **'Positive'**, **'Neutral'**, **'Negative'** colors = [**'blue'**, **'yellow'**, **'red'**]  
 explode = (0.1, 0, 0)  
 plt.pie(sizes, explode=explode, labels=labels, colors=colors,autopct=**'%1.1f%%'**, shadow=**True**, startangle=140)  
 plt.axis(**'equal'**)  
 plt.show()

The interface of the program is built with the help of GUI library of python Tkinter . The intergartion of the frontend and backend was done with specific command and functions**.**

1. **System Requirement**

* Active internet connection : as we are retrieving real time tweets we need an internet connection.
* Twitter account : as we are using and twitters api ,we need to have a developers account in twitter.
* Twitters app keys: we also need an application that we can use for its consumer and access keys.
* Some of the python dependencies such as :

1. Tweepy : For API access
2. Textblol : For sentiment analyse
3. Tkinter : For GUI
4. Matplotlib : For showing graphs

* Hardware : laptop or a pc .

1. **Testing**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sr no** | **Topic** | **Average polarity** | **Average subjectivity** | **% of positive tweets** | **% of negative**  **tweets** | **% of neutral tweets** |
| 1 | Trump | 0.07313812229 | 0.274842442279 | 41.0% | 13.0% | 46.0% |
| 2 | Iphone Xs max | 0.14427317821 | 0.203823953823 | 34.0% | 1.0% | 65.0% |
| 3 | Pixel 3 xl | 0.12739745670 | 0.227079725829 | 31.0% | 7.0% | 62.0% |
| 4 | Narendra Modi | 0.07845454545 | 0.163108585858 | 21.0% | 2.0% | 77.0% |

Fig 5.1 :- Different results of topic

This results may vary every time when we search for a particular topic as we consider the real time tweets .

1. **Advantages & Disadvantages**

**Advantages**

* Sentiment analysis is a useful tool for any organization or group for which public sentiment or attitude towards them is important for their success - whichever way that success is defined
* On social media, blogs, and online forums millions of people are busily discussing and reviewing businesses, companies, and organizations. And those opinions are being ‘listened to’ and analyzed.
* Those being discussed are making use of this enormous amount of data by using computer programs that don’t just locate all mentions of their products, services, or business, but also determine the emotions and attitudes behind the words being used.
* The results from sentiment analysis help businesses understand the conversations and discussions taking place about them, and helps them react and take action accordingly.
* They can quickly identify any negative sentiments being expressed, and turn poor customer experiences into very good ones.

They can create better products and services, and they can formulate the marketing messages they send out according to the sentiments being expressed by their target audience or customers.

* By listening to and analyzing comments on Facebook and Twitter, local government departments can gauge public sentiment towards their department and the services they provide, and use the results to improve services such as parking and leisure facilities, local policing, and the condition of roads.
* Universities can use sentiment analysis to analyze student feedback and comments garnered either from their own surveys, or from online sources such as social media. They can then use the results to identify and address any areas of student dissatisfaction, as well as identify and build on those areas where students are expressing positive sentiments.

**Disadvantages**

* Sentiment analysis tools can identify and analyze many pieces of text automatically and quickly but computer programs have problems recognizing things like sarcasm and irony, negations, jokes, and exaggerations - the sorts of things a person would have little trouble identifying. And failing to recognize these can skew the results. 'Disappointed' may be classified as a negative word for the purposes of sentiment analysis, but within the phrase “I wasn't disappointed", it should be classified as positive.
* We would find it easy to recognize as sarcasm the statement "I'm really loving the enormous pool at my hotel!", if this statement is accompanied by a photo of a tiny swimming pool; whereas an automated sentiment analysis tool probably would not, and would most likely classify it as an example of positive sentiment with short sentences and pieces of text, for example like those you find on Twitter especially, and sometimes on Facebook, there might not be enough context for a reliable sentiment analysis. However, in general, Twitter has a reputation for being a good source of information for sentiment analysis, and with the new increased word count for tweets it's likely it will become even more useful .So, automated sentiment analysis tools do a really great job of analyzing text for opinion and attitude, but they're not perfect.

1. **Future Scope**

The task of sentiment analysis, especially in the domain of micro-blogging, is still in the developing stage and far from complete . sentiment analysis using twitter and also other social platform can be beneficial for future. It may have a such applications in future

* **Business:**In marketing field companies may use it to develop their strategies, to understand customers’ feelings towards products or brand, how people respond to their campaigns or product launches and why consumers don’t buy some products.
* **Politics:**In political field, it is may be used to keep track of political view, to detect consistency and inconsistency between statements and actions at the government level. It can be used to predict election results as well!
* **Public Actions:**Sentiment analysis also is used to monitor and analyse social phenomena, for the spotting of potentially dangerous situations and determining the general mood of the blogosphere.
* **Smartphones** : smartphones are said to be smart but what if this smartphones are able to know the sentiments of users and give them a whole new level of experience ,not only smart phones but others devices such as laptops, smartwatches and also use it

1. **Conclusion**

Hence we have implemented a simple twitter sentiment analyzer using python dependencies such as textblob and tweepy , in this project we have successfully classified the tweets into positive ,negative and neutral . We have find the average polarity and subjectivity of the particular thread. The two graphs i.e retweet and overall graphs are showed for pictorial analysis Sentiment analysis can also be carried out using other platforms such as facebook and some forums sites .

1. **Reference**

**Youtube :** [**https://www.youtube.com/watch?v=o\_OZdbCzHUA**](https://www.youtube.com/watch?v=o_OZdbCzHUA)