**Sentimental Analysis using Python**

A

Major project report

*Submitted in partial fulfillment for the award of*

*Degree of*

**BACHELOR OF TECHNOLOGY**

**IN**

**(ELECTRONIC AND COMMUNICATION ENGINEERING)**

**Session: 2020**

**SUBMITTED TO:**



**JAMIA HAMDARD**

**Hamdard Nagar, New Delhi , Delhi 110062**

SUBMITTED BY:

**MRIDUL KUMAR MEENA**

**2017-333-069(7TH Sem)**

UNDER THE GUIDANCE OF

**ZEESHAN ALI HAQ**

**DEPARTMENT OF**

**COMPUTER SCIENCE AND ENGINEERING**

**Table of Content**

|  |  |
| --- | --- |
| **1** | **CERTIFICATE** |
| **2** | **ABSTRACT** |
| **3** | **ACKNOWLEDGEMENT** |
| **4** | **OBJECTIVE** |
| **5** | **WHAT IS SENTIMENTAL ANALYSIS ?** |
| **6** | **SYSTEM MODEL** |
| **7** | **HOW DOES THE ALGORITHM WORK** |
| **8** | **WHAT IS PYTHON?** |
| **9** | **SYSTEM IMPLEMENTATION** |
| **10** | **HOW DOES IT WORK?** |
| **11** | **BLOCK DIAGRAM** |
| **12** | **SOURCE CODE** |
| **13** | **HARDWARE /SOFTWARE REQUIREMENT** |
| **13** | **DATA COLLECTION** |
| **14** | **RESULT** |
| **15** | **WORD CLOUD** |
| **16** | **CONCLUSION** |
| **17** | **BIBLIOGRAPHY** |

**JAMIA HAMDARD**

**HAMDARD NAGAR, NEW DELHI-110062**

**DEPARTMENT: COMPUTER SCIENCE AND ENGINEERING**

**CERTIFICATE**

## This is to certify that synopsis entitled “SENTIMENTAL ANALYSIS USING PYTHON” submitted to JAMIA HAMDARD by MRIDUL KUMAR MEENA is a partial fulfillment of the requirement for the award of the degree in B. Tech with specialization in ELECTRONIC AND COMMUNICATION.

**CANDIDATE NAME: MRIDUL KUMAR MEENA**

**ENROLLMENT NUMBER: 2017-333-069**

**(Signature and Name of the Supervisor)**

**Zeeshan Ali Haq**

**ABSTRACT**

In this project we will be doing Sentimental Analysis using python. Analysis of Twitter Sentiment using Python can be done through popular Python libraries like Tweepy and TextBlob. Tweepy, the Python client for the official Twitter API supports accessing Twitter via Basic Authentication and the newer method, OAuth. Sentiment analysis enables enterprises to understand consumer sentiments in relation to specific products/services.

**ACKNOWLEDGEMENT**

I would like to express my sincere gratitude and deep appreciation to my esteemed teacher and major project supervisor **Zeeshan Ali** **Haq** who gave me the opportunity to work on this wonderful project titled “**Sentimental Analysis using python**” .

I would also like to extend my gratitude to **Prof. M. Afshar Alam**, Dean and head of the **School of Engineering Science and Technology, Jamia Hamdard** for providing me the necessary infrastructure for the development of this project.

**(Signature and Name of the Applicant)**

**MRIDUL KUMAR MEENA**

**B.TECH (ECE)**

**OBJECTIVE**

The goal is to conceptually design and implement aspect-based sentiment analysis toolkit for extracting sentiment from the Twitter Data.

sentiment analysis on data from Twitter (tweets), to extract user’s opinions and

The main goal is to explore how text analysis techniques can be used to dig into some of the data in a series of posts focusing on different trends of tweets languages, tweets volumes on twitter. Experimental evaluations show that python is efficient and performs better in terms of accuracy. The proposed algorithm is also implemented in python.

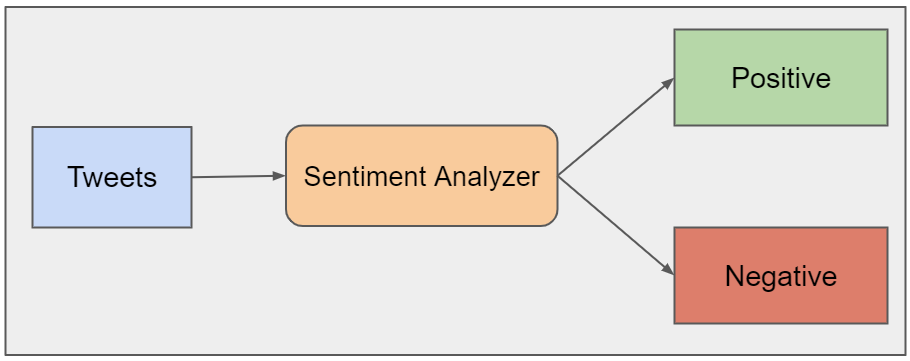
**What is Sentiment Analysis?**

.

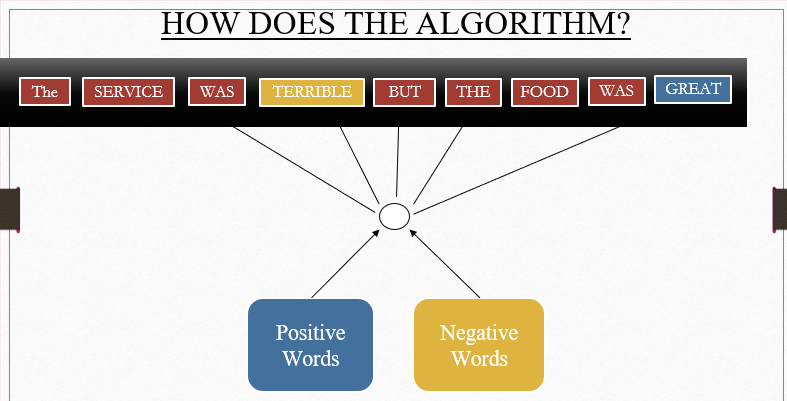
Sentimental analysis is the process of computationally determining the opinion or attitude of the writers as positive, negative or neutral. Data mining is another name for sentimental analysis. In many fields like business, politics and public actions, determining the sentimental analysis is very important. Considering business, it is very useful to understand the customer’s feelings in order to develop their company. Next in politics: It can be even be used to predict the election results. It has commercial applications such as determining the acceptance and effectiveness of a product or service among the consumers etc.

.

**SYSTEM MODEL**



**HOW DOES THE ALGORITHM WORK?**



**WHAT IS PYTHON?**

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

**SYSTEM IMPLEMENTATION**

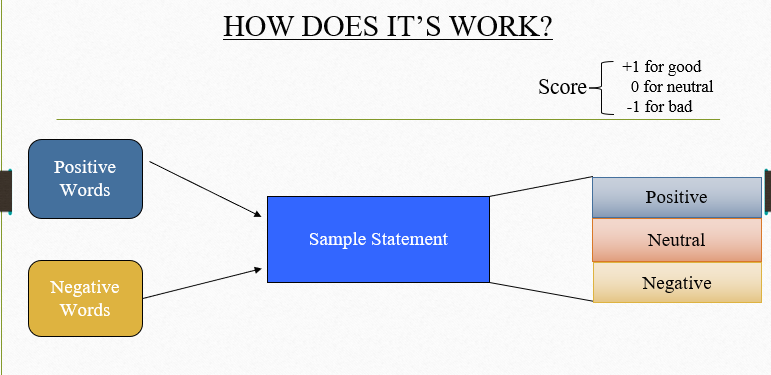
Implementation is the carrying out, execution, or practice of a plan, a method, or any design, idea, model, specification, standard or policy for doing something. As such, implementation is the action that must follow any preliminary thinkingin order for something to actually happen.

Implementation is a process of ensuring that the information system is

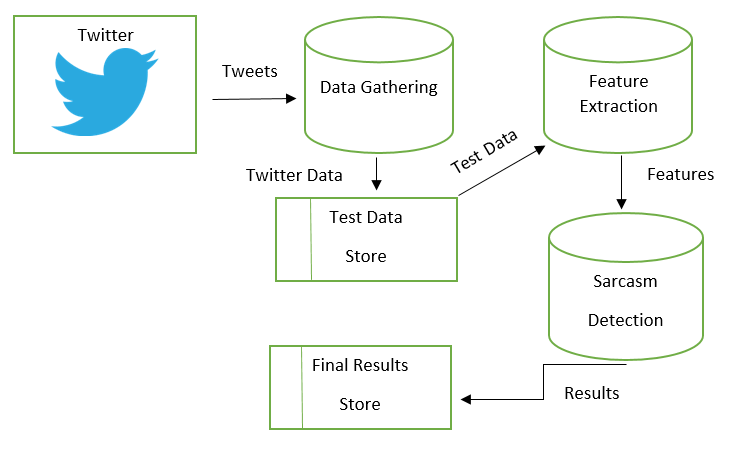
●Constructing a new system from scratch

●Constructing a new system from the existing one.

**HOW DOES IT WORK?**



**BLOCK DIAGRAM**

****

**SOURCE CODE**

import re

import tweepy

from tweepy import OAuthHandler

from textblob import TextBlob

from wordcloud import WordCloud

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

plt.style.use('fivethirtyeight')

#twitter Api credentials

consumerKey = "iLskwWqgXgmKap7w3c1nO2kun"

consumerSecret = 'ZtBRZxXcyQNNjhTFJbPRFJdunmewYnsSdzo69RY0y4nUUMtPSb'

accessToken = '1339257813055377409-koopLQjwQSd15qtmSkugDUwRC61JvN'

accessTokenSecret = 'wkUsMdV6XBwaRy7WwqMqV2JHG9lKoBYvYWxKv6F1qBikK'

#Create the authentication object

authenticate = tweepy.OAuthHandler(consumerKey, consumerSecret)

#set the authentication ojject

authenticate.set\_access\_token(accessToken, accessTokenSecret)

#Create the API object while passing in the auth information

api = tweepy.API(authenticate, wait\_on\_rate\_limit=True)

#Extract 100 tweets from the twitter user

posts = api.user\_timeline(screen\_name = "BillGates", count =100, lang = 'en', tweet\_mode="extended")

#create a dataframe with a coulmn called Tweets

df = pd.DataFrame( [tweet.full\_text for tweet in posts] , columns=['Tweets'])

#Show the first 5 rows of data

df.head()

#Clean the text

#Create a function to clean text

def cleantxt(text):

    text = re.sub(r'@[A-Za-z0-9]+','',text) #Removed @mentions

    text = re.sub(r'#', '', text)   #Removing the '#' symbol

    text = re.sub(r'RT[\s]+', '', text) #Removing RT

    text = re.sub(r'https?:\/\/\S+', '', text)

    return text

#Cleaning the text

df['Tweets']=df['Tweets'].apply(cleantxt)

#show the cleaned text

df

#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.subjectivity

#Create two new columns

df['Subjectivity'] = df['Tweets'].apply(getSubjectivity)

df['Polarity'] = df['Tweets'].apply(getPolarity)

#Show the new dataframe with the new columns

df

#Plot the word cloud

allWords = ' '.join( [twts for twts in df['Tweets']] )

cloud = WordCloud(width = 500, height = 300, random\_state = 21, max\_font\_size = 119).generate(allWords)

plt.imshow(cloud, interpolation = "bilinear")

plt.axis('off')

plt.show()

#create a function to compute the negatve, neutral and positive

def getAnalysis(score):

    if score < 0:

        return 'Negative'

    elif score == 0:

        return 'Neutral'

    else:

        return 'Positive'

df['Analysis'] = df['Polarity'].apply(getAnalysis)

#show the dataframe

df

#Print all of the positive tweets

j=1

sortedDF = df.sort\_values(by=['Polarity'])

for i in range(0,sortedDF.shape[0]):

    if(sortedDF['Analysis'][i] == 'Positive'):

        print(str(j) + ') '+sortedDF['Tweets'][i])

        print()

        j = j+1

#print the negative tweets

j=1

sortedDF = df.sort\_values(by=['Polarity'], ascending = 'False')

for i in range(0, sortedDF.shape[0]):

    if( sortedDF['Analysis'][i] == 'Negative'):

        print(str(j) +') '+ sortedDF['Tweets'][i])

        print()

        j = j+1

#plot the polarity and subjectivity

plt.figure(figsize=(8,6))

for i in range(0,df.shape[0]):

    plt.scatter(df['Polarity'][i], df['Subjectivity'][i], color='Blue')

plt.title('Sentiment Analysis')

plt.xlabel('Polarity')

plt.ylabel('Subjectivity')

plt.show()

#Get the percentage of positive tweets

ptweets = df[df.Analysis == 'Positive']

ptweets = ptweets['Tweets']

round( (ptweets.shape[0] / df.shape[0]) \*100 , 1)

#Get the percentage of negative tweets

ntweets = df[df.Analysis == 'Negative']

ntweets = ntweets['Tweets']

round( (ntweets.shape[0] / df.shape[0]\*100),1)

#show the value counts

df['Analysis'].value\_counts()

#plot and visualize the counts

plt.title('Sentiment Analysis')

plt.xlabel('Sentiment')

plt.ylabel('Counts')

df['Analysis'].value\_counts().plot(kind='bar')

plt.show()

**HARDWARE REQUIREMENTS**

* **Processor**: Intel i5 with clock speed of 2.4GHz or above.
* **RAM**: 4GB or above.
* **Hard disk**: 100GB or above.
* **Input device**: Keyboard or mouse or compatible pointing devices.
* **Display**: XGA (1024\* 768px) or higher resolution monitor.

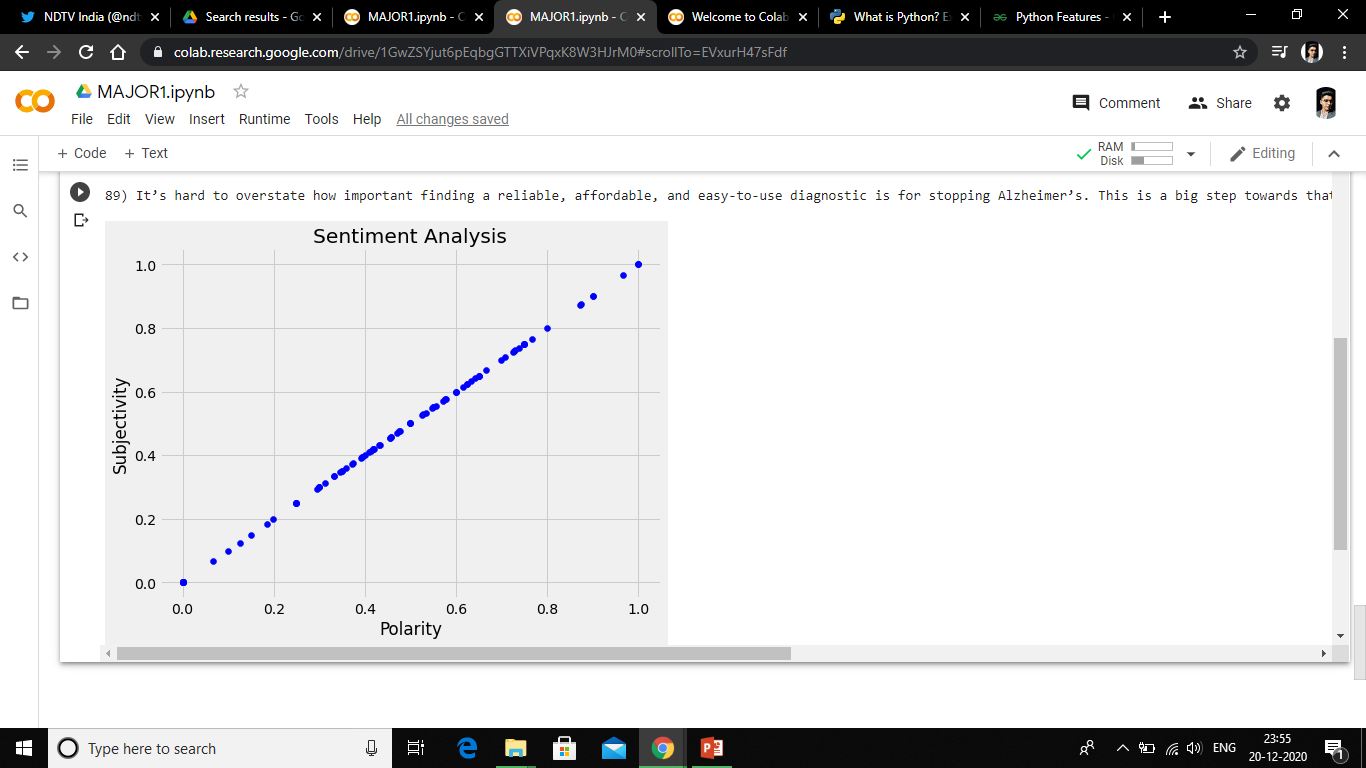
**SOFTWARE REQUIREMENTS**

**Operating system :** Windows 7 or above.

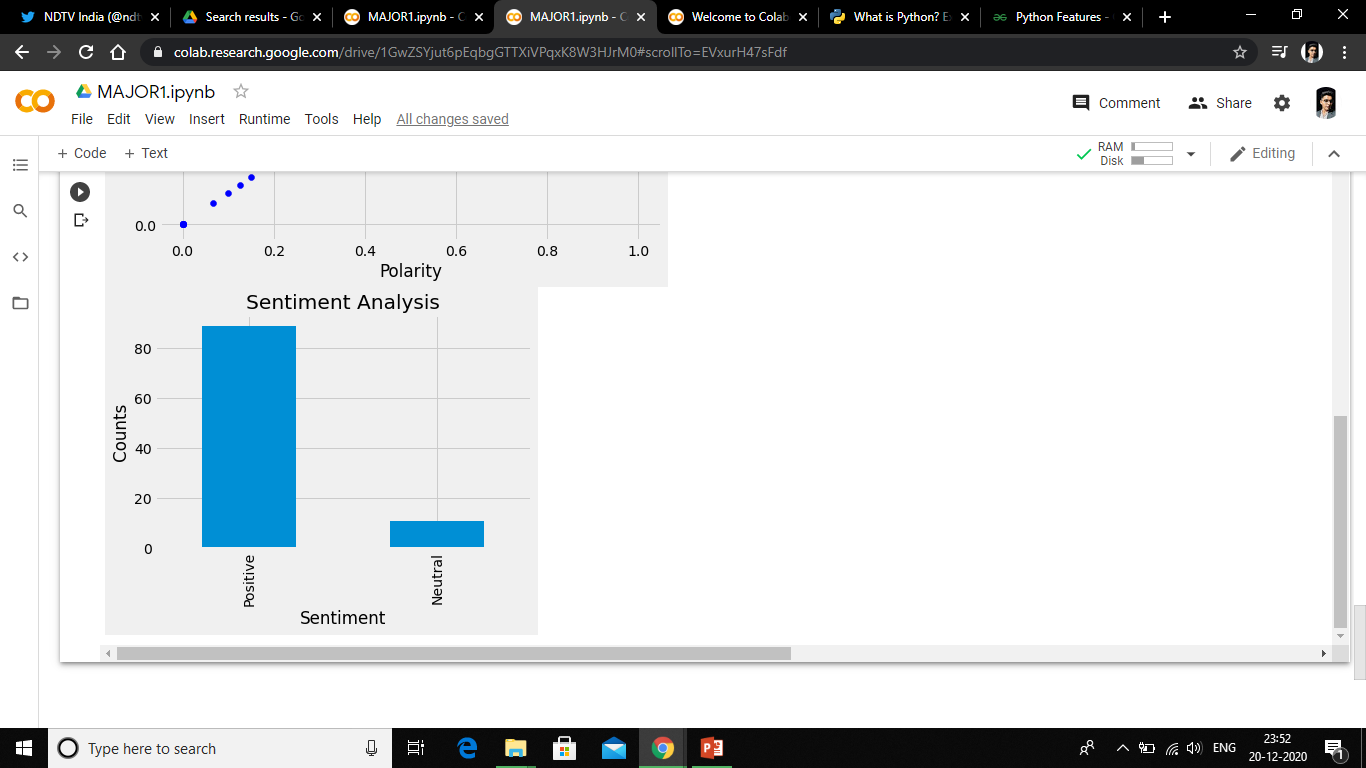
**Programming Language :** PYTHON 3

**Development Environment :** GOOGLE COLAB

**DATA COLLECTION/ TWEET EXTRACTION**



**RESULT**



**WORD CLOUD**

****

**CONCLUSION**

In this project, As can be seen, sentiment analysis enables enterprises to understand consumer sentiments in relation to specific products/services. Moreover, these insights could be used to improve their products and services by gauging consumers’ comments and feedback using sentiment analysis. In the long run, sentiment analysis, if implemented the right way can aid business enterprises in improving the overall consumer experience, enhance brand image and propel business growth.

**BIBLIOGRAPHY**

[1] Efthymios Kouloumpis and Johanna Moore,IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012

[2] S. Batra and D. Rao, ”Entity Based Sentiment Analysis on Twitter”, Stanford University,2010

[3] Saif M.Mohammad and Xiaodan zhu ,Sentiment Analysis on of social media texts:,2014

[4] Manju Venugopalan and Deepa Gupta ,Exploring Sentiment Analysis on Twitter Data, IEEE 2015

[5] Brett Duncan and Yanqing Zhang, Neural Networks for Sentiment Analysis on Twitter.2017