GraphicEra
(Deemed to be University)
Accredited by NAAC with Grade A

MINI PROJECT REPORT

2021-22

PROJECT NAME : **SENTIMENT ANALYSIS USING PYTHON**

NAME : **Vansh Rastogi**

UNIVERSITY ROLL NO : **2017118**

STUDENT ID : **20021070**

COURSE : **B.Tech**

BRANCH : **Computer Science And Engineering**

DATE : **26-02-2022**

INTRODUCTION

Motivation

Being interested to use Python language, this independent project was the best opportunity to give me time to learn the Python language and several libraries which can be used with python. Actually Python and R are the languages which are in demand in the new technologies used world wide. Cloud computing, Artificial Intelligence, Machine learning, Web and game development uses python .

Problem Statement

Sentiment Analysis using Python for Twitter. Actually twitter is a microblogging website on which discussions on different topics are going on. Some in this we picked some topic and extract the comments and thoughts of different people on the particular topic. After doing the analysis we can differentiate between the positive and negative comments or thoughts and also we will be able to tell that how much percent of thoughts are positive , negative and neutral.

What is Sentiment Analysis ?

Sentiment Analysis is a process of determining whether a piece of information is positive, negative or neutral. This process can also be called as opinion mining ,i.e., determining the opinion or feeling of a person.

Why Sentiment Analysis ?

1. **Politics:** In the field of politics, it can be used to predict the results of elections and to detect the consistency and inconsistency between statements and actions of the government.
2. **Marketing:** Marketing companies take help of sentiment analysis to develop their strategies, and to understand the feelings of customers regarding the brand and its products and whether the customers want to buy those products or not.
3. **Social Media:** In the field of social media it helps to monitor how interested the audience is in any trend that come along.
4. **Actions of Public:** Sentiment analysis is used to analyse social situations, and for tracking of potentially dangerous situation. It can also be used to analyse mood of public involved in a protest or campaign.

TOOLS USED

In this project I have used natural language processing(NLP) and the python editor used for writing and testing the code is Anaconda's Spyder.

NLP is a field which focuses on making the human language usable by computer program codes. NLTK or natural language toolkit is a python package which can be used for NLP.

THE PROJECT

The project ‘Sentiment Analysis Using Python’ is a real life based project which works on the data extracted from Twitter. In this project firstly we need a python editor which in case of mine is Anaconda’s spyder. After installing the editor we need to install some library and packages which are very important in this project. We need a to install pip package using anaconda and after that we need to install the libraries tweepy and textblob. After these libraries are installed the last package to install is the NLTK corpora. Tweepy is a python client for twitter API which is an open-sourced library. It gives an interface to you to access the API. TextBlob is also an open-source python library for textual data processing. It provides an API for performing natural language processing tasks like noun phrase extraction, part of speech tagging, translation, etc. Corpora is a structured and large set of texts.

In this project, I try to classify tweets from Twitter into ‘positive’ and ‘negative’ sentiments. Twitter is a blogging website where people can share feelings quickly and spontaneously by sending tweets. Also I have tried to find how many percent tweets are positive, negative and neutral. To find the sentiment of the tweets we have to clean the data we are collecting from Twitter. After that we check that the sentiment is positive, negative or neutral.

DATA

To gather the data I have created a Twitter developer account and created a new app/project. After creating the app/project I can access the keys and access tokens which we can enter into our code and access the data directly from Twitter. The data I gathered from Twitter is related to Donald Trump.

DETAILS TO ACCESS DATA FROM TWITTER

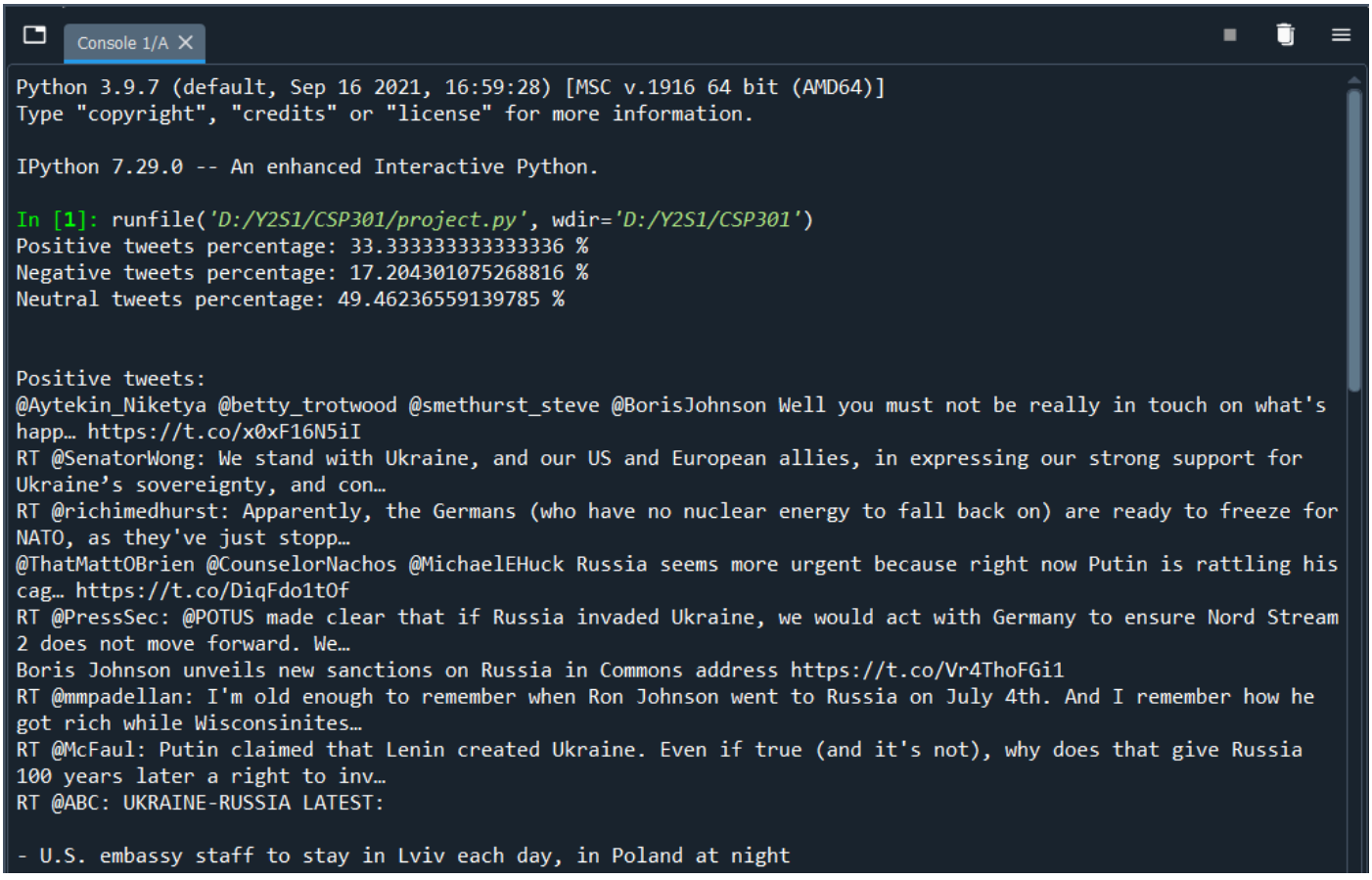
API key -'dmV2vT6CQx4WgrMAQntqu0oAs'

API Secret key-'yX2JF380y40xh1OvjsIn7CiDk2Wo6j3imIbI6fkO5DqVd9dyB1'

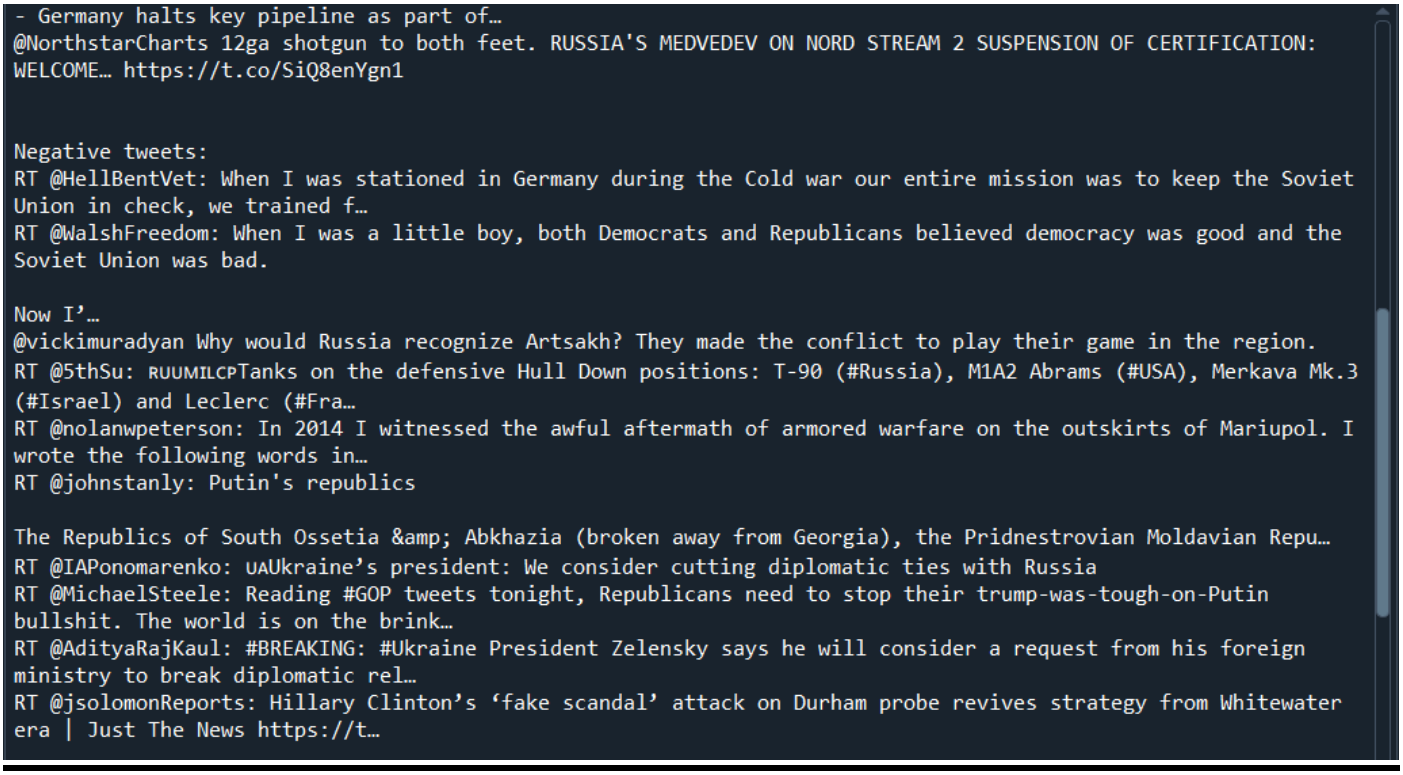
Access token-'1475777230030606339i8XkY6JdEHrdY6DnQRiEhxCWR6V3fB'

Access token secret -'iM0iaVorgT1oXkeRn6ZyhjzHOI8eY9ao2ouveV9eCmBO8'

SCREENSHOTS OF OUTPUTS



```
Console 1/A

Python 3.9.7 (default, Sep 16 2021, 16:59:28) [MSC v.1916 64 bit (AMD64)]
Type "copyright", "credits" or "license" for more information.

IPython 7.29.0 -- An enhanced Interactive Python.

In [1]: runfile('D:/Y2S1/CSP301/project.py', wdir='D:/Y2S1/CSP301')
Positive tweets percentage: 33.333333333333336 %
Negative tweets percentage: 17.204301075268816 %
Neutral tweets percentage: 49.46236559139785 %


Positive tweets:
@Aytekin_Niketya @betty_trotwood @smethurst_steve @BorisJohnson Well you must not be really in touch on what's
happ… https://t.co/x0xF16N5iI
RT @SenatorWong: We stand with Ukraine, and our US and European allies, in expressing our strong support for
Ukraine’s sovereignty, and con…
RT @richimedhurst: Apparently, the Germans (who have no nuclear energy to fall back on) are ready to freeze for
NATO, as they've just stopp…
@ThatMattOBrien @CounselorNachos @MichaelEHuck Russia seems more urgent because right now Putin is rattling his
cag… https://t.co/DiqFdo1tOf
RT @PressSec: @POTUS made clear that if Russia invaded Ukraine, we would act with Germany to ensure Nord Stream
2 does not move forward. We…
Boris Johnson unveils new sanctions on Russia in Commons address https://t.co/Vr4ThoFGi1
RT @mmpadellan: I'm old enough to remember when Ron Johnson went to Russia on July 4th. And I remember how he
got rich while Wisconsinites…
RT @McFaul: Putin claimed that Lenin created Ukraine. Even if true (and it's not), why does that give Russia
100 years later a right to inv…
RT @ABC: UKRAINE-RUSSIA LATEST:

- U.S. embassy staff to stay in Lviv each day, in Poland at night
```



```
- Germany halts key pipeline as part of…
@NorthstarCharts 12ga shotgun to both feet. RUSSIA'S MEDVEDEV ON NORD STREAM 2 SUSPENSION OF CERTIFICATION:
WELCOME… https://t.co/SiQ8enYgn1


Negative tweets:
RT @HellBentVet: When I was stationed in Germany during the Cold war our entire mission was to keep the Soviet
Union in check, we trained f…
RT @WalshFreedom: When I was a little boy, both Democrats and Republicans believed democracy was good and the
Soviet Union was bad.

Now I’…
@vickimuradyan Why would Russia recognize Artsakh? They made the conflict to play their game in the region.
RT @5thSu: ʀuumılcpTanks on the defensive Hull Down positions: T-90 (#Russia), M1A2 Abrams (#USA), Merkava Mk.3
(#Israel) and Leclerc (#Fra…
RT @nolanwpeterson: In 2014 I witnessed the awful aftermath of armored warfare on the outskirts of Mariupol. I
wrote the following words in…
RT @johnstanly: Putin's republics

The Republics of South Ossetia &amp;amp; Abkhazia (broken away from Georgia), the Pridnestrovian Moldavian Repu…
RT @IAPonomarenko: ᴜAUkraine’s president: We consider cutting diplomatic ties with Russia
RT @MichaelSteele: Reading #GOP tweets tonight, Republicans need to stop their trump-was-tough-on-Putin
bullshit. The world is on the brink…
RT @AdityaRajKaul: #BREAKING: #Ukraine President Zelensky says he will consider a request from his foreign
ministry to break diplomatic rel…
RT @jsolomonReports: Hillary Clinton’s ‘fake scandal’ attack on Durham probe revives strategy from Whitewater
era | Just The News https://t…
```

CONCLUSION

These days, sentiment analysis or opinion mining is a common topic. We are still very far to detect the sentiments of the texts very accurately and efficiently because of the complexity in English languages and changes occurring in the English language or any other language used in different countries.

In this project, I have tried to categorize the tweets into positive or positive category taking Naive Bayes as base. Also I have tried to find that how much percent are the positive tweets, negative tweets, and neutral tweets.

REFERENCES

1. https://www.datacamp.com/community/tutorials/simplifying-sentiment-analysis-python
2. https://textblob.readthedocs.io/en/dev/quickstart.html#sentiment-analysis
3. https://docs.tweepy.org/en/stable/authentication.html