

Name: - Prekshita Vasudeo Patil
Registration Number: 20MAI0073

Assignment-5

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import sklearn as sk
import seaborn as sns
import keras as k
import tensorflow as tf
import re
import nltk
from nltk.corpus import stopwords
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.decomposition import LatentDirichletAllocation
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import BernoulliNB, GaussianNB, MultinomialNB
from sklearn.svm import SVC
from sklearn.svm import LinearSVC
from sklearn.linear_model import LogisticRegression
from sklearn.ensemble import RandomForestClassifier
from sklearn.ensemble import AdaBoostClassifier
from sklearn.metrics import accuracy_score, confusion_matrix
import tweepy
import warnings
from textblob import TextBlob
warnings.filterwarnings('ignore')
nltk.download('stopwords')
stop = stopwords.words('english')
# ----- Inserting Data -----
read = pd.read_csv("training.csv", names = ["target", "ids", "date", "flag", "user", "text"])
# ----- Twitter tweets -----
consumerKey = 'ha5WeZWCvDIh0CXScpNpDn7y1'
consumerSecret = '1l3tciA5MgrjU675tHSThv08cCTMHRMkBXxQWjAkiCqF66Qhfw'
accessToken = '963415868-TZqxvLp8lF5ThWsDMSZGU6b6fosMmdsEeGxMZrCU'
accessTokenSecret = '7ge4JM4E6WyVloaMIb4vGijBZd2PRjSPGCleoth7xm4kv'
# ----- Giving Keys -----
auth = tweepy.OAuthHandler(consumerKey, consumerSecret)
auth.set_access_token(accessToken, accessTokenSecret)
api = tweepy.API(auth)
# ----- authentication done! -----
new_df = list(read['text'])
what_to_analyse = int(input('What do you want to see\n1. Sentence\n2. Twitter Tweets\n'))
if what_to_analyse == 1:
    # input for sentence to be analyzed
```

```

Sentence_input = input('Enter a sentence to analyse :- \n')
new_df.append(Sentence_input)
elif what_to_analyse == 2:
    # input for term to be searched and how many tweets to search
    searchTerm = input("Enter Keyword/Tag to search about: ")
    NoOfTerms = int(input("Enter how many tweets to search: "))
    tweets = tweepy.Cursor(api.search,q=searchTerm,lang="en",).items(NoOfTerms)
    for x,i in enumerate(tweets):
        new_df.append(i.text)
else:
    print('invalid input')
new_df = pd.DataFrame(new_df,columns=['text'])
new_df['text'] = new_df['text'].replace(to_replace = '@[\w]+', value = '', regex = True)
new_df['text'] = new_df['text'].replace(to_replace = ('RT'), value = '',regex = True)
# for x,i in enumerate(new_df['text']):
#     if 'rt ' in i:
#         i = i.replace('rt ','')
#         new_df['text'][x] = i
new_df['text'] = new_df['text'].str.replace('((www\.[\s]+)|(https?://[^\s]+))','\0',regex=True)
new_df['text'] = new_df['text'].str.replace("[^a-zA-Z]+", " ")
new_df["text"] = new_df["text"].str.lower()
new_df["text"] = new_df["text"].str.split()
new_df['text'] = new_df['text'].apply(lambda x: [item for item in x if item not in stop])
def rejoin_words(row):
    my_list = row['text']
    joined_words = ( " ".join(my_list))
    return joined_words
new_df['text'] = new_df.apply(rejoin_words, axis=1)
if what_to_analyse == 1:
    sentence_df = new_df.tail(1)
elif what_to_analyse == 2:
    tweeeets = new_df[-NoOfTerms:]
vectorizer=CountVectorizer(max_df=0.95,min_df=2,max_features=10000,ngram_range=(1,3),stop_
words='english')
x=vectorizer.fit_transform(new_df['text'])
new_df_vectorized = vectorizer.fit_transform(new_df['text'])
if what_to_analyse == 1:
    xtrain = new_df_vectorized[:-1]
    xtest = new_df_vectorized[-1]
elif what_to_analyse ==2 :
    xtrain = new_df_vectorized[:-NoOfTerms]
    xtest = new_df_vectorized[-NoOfTerms:]
ytrain = read['target']
bern = BernoulliNB()
bern.fit(xtrain,ytrain)
ypredict = bern.predict(xtest)
count_negative,count_positive =0,0
for i in range(len(ypredict)):

```

```

if ypredict[i] == 0:
    count_negative+=1
else:
    count_positive+=1
neg_per = round(count_negative/len(ypredict)*100,2)
pos_per = round(count_positive/len(ypredict)*100,2)
if what_to_analyse == 1:
    sentence_df.to_csv('Sentence.csv',header=False,index=False,mode='w')
elif what_to_analyse == 2:
    tweeeets.to_csv('Tweets.csv',header=False,index=False,mode='w')
user_input = int(input("How do you want to see pie or bar chart \n1. Pie\n2. Bar\n"))
plt.figure(figsize=(8,8),)
plt.ylim((0,100))
if user_input == 1:
    if neg_per == 0:
        plt.pie([pos_per],data=[pos_per],labels=['Postive tweets'],colors=['green'],autopct='%1.1f%%')
    elif pos_per == 0:
        plt.pie([neg_per],data=[neg_per],labels=['Negative Tweets'],colors=['red'],autopct='%1.1f%%')
    else:
        plt.pie([neg_per,pos_per],data=[neg_per,pos_per],labels=['Negative Tweets','Postive tweets'],colors=['red','green'],autopct='%1.1f%%')
elif user_input == 2:
    plt.bar(['Negative Percentage','Positve Percentage'],[neg_per,pos_per],color=['red','green'])
    for i,x in zip(['Negative Percentage','Positve Percentage'],[neg_per,pos_per]):
        plt.text(i,x+5,x,size=15)
plt.savefig('Answer.jpg')
plt.show()

```

Output

What do you want to see

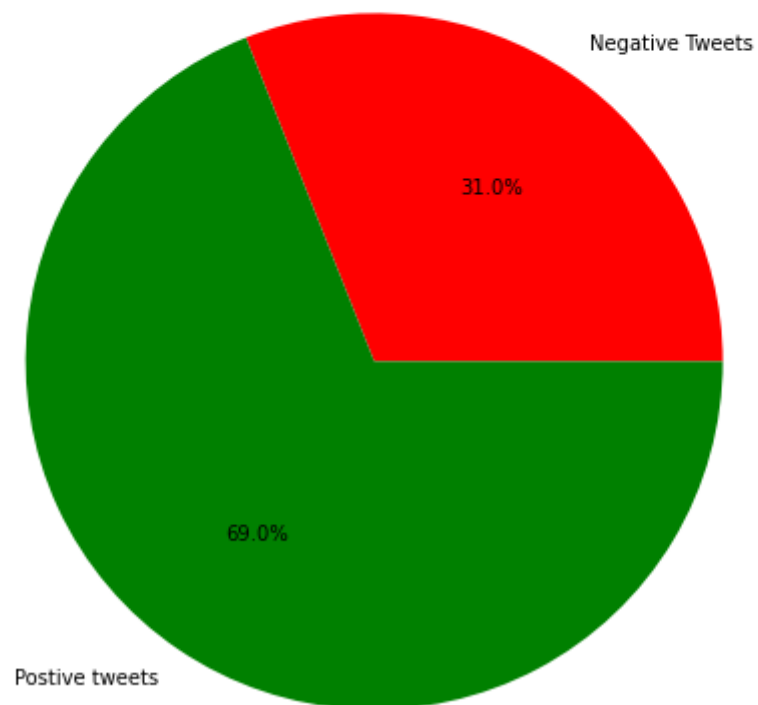
1. Sentence
 2. Twitter Tweets
- 2

Enter Keyword/Tag to search about: India

Enter how many tweets to search: 100

How do you want to see pie or bar chart

1. Pie
 2. Bar
- 1



Tweets(Modified) came from Twitter
owned micromaxinfinity n never received single os update
armys belong bts
biggest football juggling championship india almost ready coming soon watch space
precisely eat india company india
convey thankful even try explain cant winter right want
certain indian media interview head taiwan authorities foreign relations department jaushieh joseph
guys need support gst taking toll small buisness ta
india secularism means hindus communalism means hindus
list fundraisers transgender people india thread
spotify india dynamite plays days
taehyung comfort videos healing thread taehyung always abundance love army love vast ocean
kejriwal govt efforts stop pollution banned electricity generator tree transplantation policy green delhi app
isupportarnabgoswami warriors arnab sir needs support today always stood sushant people
confirming wrote pm modi february insisting talks happen highest level without pre cond
hong kong china taiwan china tibet also china india cannot economic power long
india biggest cricket festival ipl going lot negativity done boycott gang bb meanwhile bb trp
glad realised party destroyed years india legacy built congress govts merely years
unswayed sidhah shukla trending number india huge small twitter accounts made sid
drdo achieved milestone today successful flight test abhyas high speed expandable aerial target
join us consultation building resilient livelihoods long term economic recovery rural india sessions
rate first mention biryani comes th century centuries mughals entered india
derries trending india
main countries new followers came last week nigeria us england greece france germany portu
fantastic artistry india unmatched richness comes depicting social adminis
raga lov countries like pak well established turn india pak within
arnab minutenotice taken constitution india thrown away blatant disregard const
india provide taiwan trillion tons bullshit
mughals originally speakers chagatai language turkic language belonging altai family means
arnab minutenotice maha assembly gives arnab mins appear assembly republic approach court
amid india aggression hong kong crackdown xi targeting taiwan stepping provocative actions threats
india democracy communist authoritarian regime support freedom speech
india secularism means hindus communalism means hindus
dear government give us rights releaseugcjrf releaseugcjrf releaseugcjrf releaseugcjrf
new india offensive offensive
country forefront mobile market chinese mobiles time
another solid achievement bjp government even pakistan afghanistan handled covid better india
pilot decided emergency landing seeing video screen individual disembarked
case sour grapes disparage india parliamentary democracy imagining decline
shameful truth many indians consider dalits muslims tribals human cm amp police say one
hey sahil working towards improving services customers disappoint
acche din getting scarier every day india falls hunger index million mouths fee
first unambiguous mention biryani comes nuskha shahjahani th century made kitchens indi

extremely shameful step opposition benefitting india
audiq officially launched india today quickfire take hits misses latest
maharashtra india shameful fascism maharashtra isupportarnabgoswami arnab minu
pls make gaming phone good processor good looking mobile best begin journey luck
arnab minutenotice taken constitution india thrown away blatant disregard const
apologies delay already escalated complaint concerned team get touch soon
china locus standi comment ladakh arunachal centre india recognized china illegal takeove
congress released first list candidates bihar polls starting october
curious case trp scam witness go naming india today naming republic tv happened
would love send old school textbooks mr jaishankar seems forgotten india rejected nation bu
indian state biggest proselytizer india makes liability hindu india laws agai
major squad game alliances possible india china conflict
day mixed bag teamindia asiannationscup round india indonesia round indi
long serve nation playing india fortune meddle pigs
salamu alaikum muslim student disappeared overnight top university india police cbi find na
india total covid cases amp deaths pakistan total covid cases amp deaths india populati
busy regional agendas india ranks th countries global hunger index
unacceptable constitution india clearly states freedom speech amp freedom press civil r
kind vendetta politics brazen show power dictatorship living anothe
btsxfla love collections available india jungkook version check
hah idiotic words r u indian
still auction delivery inside india rs per painting
friends arnab needs support today ever fighting justiceforssr given nd notice mh
tiss geared towards producing venom filled brainless wonders claiming ambedkarites viciousjehadis claimin
bust another secularist lie wahabi movement first war indian independence anything
checked data imf estimate shows bangladesh cross india real gdp per capita emerging
govt extends current concessional rate discount cargo movement india chabahar port iran
hey nigerians india support police brutality enough enough endsars
another solid achievement bjp government even pakistan afghanistan handled covid better india
ro karnataka invited e tender appointment h amp contractor tajsultanpur kalaburgi
india mood tolerate anti national contents fill pockets separate state india watchin
ias ips highly educated persons came work people one person stoping th passed mla h
salamu alaikum muslim student disappeared overnight top university india police cbi find na
india
rate first mention biryani comes th century centuries mughals entered india
devfest india live join us forget use devfestindia sharing yo
acche din getting scarier every day india falls hunger index million mouths fee
empower dalit adivasi appoint chiefs national commission scheduled castes amp tribes dalitlivesmatter adivasilivesma
india secularism means hindus communalism means hindus
india southern hemisphere proved wrong least
next time buy micromax mobile

apple online store india since morning appleindia
ladakh hair trigger insight talks prevent shots fired loathe surrender ad
unswayed sidhah shukla currently trending india position k tweets
political representation transgender community india sign petition
arnab minutenotice taken constitution india thrown away blatant disregard const
history says wherever india intervened singrauli sardarsarovar tatamundra support
warriors arnab sir needs support today always stood sushant people india bullywood amp tv indus
arnab minutenotice maha assembly gives arnab mins appear assembly republic approach court
friends arnab needs support today ever fighting justiceforssr given nd notice mh
happens covid meets toxic air india find
worldfoodday important consider environmental impact eat depth report inv
maharashtra government blowing away constitution
fresher industrial engineering mysore karnataka india plant customer data project software
pakistani science students deface poster country nobel prize winner physics prof abdul salaam chant allah