

tweepyStreamExample

November 1, 2021

1 Tweepy Stream Example

In this example we use the twitter api and tweepy to collect tweets based on a hash tag, cleans the text data, and organizes the data into a data frame with labels.

1.1 Install Modules

Note you will need to use tweepy v 3.7 for this example. Regular Expressions, re, and CountVectorizer are used to text cleaning and vectorization.

```
[1]: import pandas as pd
import tweepy
import matplotlib.pyplot as plt
#pip install tweepy==3.7
from tweepy import OAuthHandler
import json
from tweepy import Stream
#from tweepy.streaming import StreamListener
import re
from sklearn.feature_extraction.text import CountVectorizer
```

1.2 Connect to twitter

You need a Twitter Developer Acct and associated access codes to access the Twitter API. See the developer portal at twitter. For this version, you will need a consumer key, consumer secret, access token, and access secret. I store my passcodes in a local text file.

```
[2]: ## All 4 keys are in my TwitterCodesFile.txt and are comma sep
filename="C://Users//jerem//Google Drive//Online//iCuse//IST736//IST736repo2//
↳codeShare//profB//TwitterCodesFile.txt"
with open(filename, "r") as FILE:
    keys=[i for line in FILE for i in line.split(',')]

#API Key:
consumer_key = keys[0]
#API Secret Key:
consumer_secret =keys[1]
#Access Token:
access_token =keys[2]
```

```

#Access Token Secret:
access_secret =keys[3]

auth = OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_secret)

api = tweepy.API(auth)

```

1.2.1 Stream Listener

Tweepy has a listener class that allows you to collect “streaming” tweets. To set the parameters for this class, it is best to subclass / extend. See below.

[3]:

```

class Listener(tweepy.StreamListener):
    print("In Listener...")
    tweet_number=0
    #__init__ runs as soon as an instance of the class is created
    def __init__(self, max_tweets, hfilename, rawfile, TweetsList, LabelsList, nohashname):
        self.max_tweets=max_tweets
        print(self.max_tweets)
        #on_data() is a function of StreamListener as is on_error and on_status
        def on_data(self, data):
            self.tweet_number+=1
            #print("In on_data", self.tweet_number)
            try:
                #print("In on_data in try")
                with open(hfilename, 'a') as f:
                    with open(rawfile, 'a') as g:
                        tweet=json.loads(data)
                        ## RE: https://realpython.com/python-json/
                        tweet_text=tweet["text"]
                        #print(tweet_text,"\n")
                        TweetsList.append(tweet_text)
                        LabelsList.append(nohashname)
                        #print(TweetsList)
                        f.write(tweet_text) # the text from the tweet
                        json.dump(tweet, g) #write the raw tweet
            except BaseException:
                #print("NOPE")
                pass
            if self.tweet_number>=self.max_tweets:
                #sys.exit('Limit of '+str(self.max_tweets)+' tweets reached.')
                print("Got ", str(self.max_tweets), "tweets.")
                return False

```

```

#method for on_error()
def on_error(self, status):
    print("ERROR")#machi
    print(status)    #401 your keys are not working
    if(status==420):
        print("Error ", status, "rate limited")
        return False
#-----end of class Listener

```

In Listener...

2 Let the fun begin!

We will gather tweets based on 3 different hashtags (which will be used as labels for classification purposes / analyses)

```

[4]: label1 = 'cat'
    label2='dog'

Hashes = [label1, label2]
numtweets= 50  ## you can change this number!
CompleteTweetList=[]
CompleteLabelsList=[]

for hashname in Hashes:

    #####
    ## The commented out lines can be integrated if you want
    ## the USER to enter the hashname and the number of tweets
    ## to get:
    #hashname=input("Enter the hash name, such as #womensrights: ")
    #numtweets=eval(input("How many tweets do you want to get?: "))

    # if(hashname[0]=="#"):
    #     nohashname=hashname[1:] #remove the hash
    # else:
    #     nohashname=hashname
    #     hashname="#" +hashname
    #####

    print("getting...", hashname, "\n")

    nohashname=hashname
    hashname="#" +hashname
    #Create a file for any hash name
    hfilename="file_"+nohashname+".txt"

```

```

## For example, file_football.txt if you used #football
rawfile="file_rawtweets_"+nohashname+".txt"
## For example, file_rawtweets_football.txt
## Notice that the raw file is in json
## The hfilename is just text
TweetsList=[]
LabelsList=[]

##### Get the tweets.....

twitter_stream = Stream(auth, Listener(numtweets,
                                     hfilename, rawfile,
                                     TweetsList, LabelsList,
                                     nohashname))

## https://developer.twitter.com/en/docs/twitter-api/tweets/filtered-stream/
↳ introduction
## https://developer.twitter.com/en/docs/twitter-api/enterprise/
↳ powertrack-api/overview
twitter_stream.filter(track=['#womensrights'])
twitter_stream.filter(track=[hashname], languages=["en"])
##.....

## Save each Tweet in a list
## This will create a LIST OF CONTENT
## that you can use with CountVectorizer, etc.

print(TweetsList)
print(LabelsList)

## Add these to the complete lists
## extend (rather than append) will add the elements
## one by one to the existing list.
CompleteTweetList.extend(TweetsList)
CompleteLabelsList.extend(LabelsList)

##### END OF LOOP #####

```

getting... cat

50

Got 50 tweets.

```

['RT @sachikoko: CODE RED RELISTED TBD TUE 11/2/21 \n TIMID, STRESSED 2YO WHITE
& #TABBY KITTY "TAZ" \n DUMPED BY OWNER 4 LITTERBOX ISSUES\n N...', 'RT
@WonderfulSpace4: Lovely new month dear friends... \n\n#GoodMorningEveryone
#November1st #HappyNewWeek #CoffeeTime #AutumnFalls ...', 'RT @manytle:
New Drop no.11 at Tungmeng Life Collection\nTungmeng travel Queen Sirikit

```

Botanic Garden,\nChiang Mai, Thailand \nAvailable on...', 'RT @_dalim44:
 " " ,Have a meow day everyone next painting will be for Odd I won't
 forget @jaybnow_hr \n#JAYB #Def #Jaebeom #GOT7 #...', "she's cute!!! #pet #cat
 #cute #adorable #cute #kitten #petfinder I'm looking for a home!
<https://t.co/k2Qnfqxyzls> <https://t.co/bdtAUDmUm2>", ' PLEASE HELP
 SAVE ', 'RT @Dizzle_tan: Commission for anonymous of felicia, thanks for the
 support. \n\n#Darkstalkers #oppai #Nsfw #digitalartwork #Capcom #Felicia...',
 "'That's What I'm Tacoin' About!' Premium T-Shirt <https://t.co/MmEmD11ASl> #cute
 #taco #funny #food #tacos #tacobell... <https://t.co/DTv1QOVdn4>", 'RT @MShagalaih:
 Palestine Street Angeles Need the #Help and support of their savior
 @MShagalaih in securing more #Cat food that is suffic...', ' PLEASE HELP
 SAVE ', 'How cute?? 2 tap if you liked \n.\n.\nDouble tap my little heart
 if you love me \n.\n.\nTag your friends who love...
<https://t.co/rUYQotIH6X>', 'Omg Felicia is my favorite darkstalkers woman',
 ' PLEASE HELP SAVE ', 'RT @Dizzle_tan: Commission for @herrow52 of
 felicia, thanks for the support \n\n#Darkstalkers #Oppai #Nsfw #DigitalArtist
 #capcom #Felicia...', ' PLEASE HELP SAVE ', ' PLEASE HELP
 SAVE ', 'RT @manytle: New Drop no.11 at Tungmeng Life Collection\nTungmeng
 travel Queen Sirikit Botanic Garden,\nChiang Mai, Thailand \nAvailable on...',
 'This #Influencer #Tattooed Her #Cat and Faced Zero Repercussions! #care2
<https://t.co/z71ZUjaJdF>', 'RT @manytle: New Drop no.11 at Tungmeng Life
 Collection\nTungmeng travel Queen Sirikit Botanic Garden,\nChiang Mai, Thailand
 \nAvailable on...', 'RT @HoudiniMeow: RETWEET, PLEASE VOTE FOR ME. Thank
 mew all! You can vote 10x a day, refresh the web page to do this after every
 v...', 'RT @tulipteardrops: #cute #siamese #miniart #handmade #cat Siamese cat
 watercolor ACEO original painting #Miniature <https://t.co/jGCVghxci...>', 'RT
 @sachikoko: SWEET 3MO #TABBY #KITTENS "HANSEL" & "GRETEL" \n DUMPED BY
 OWNER\n NEED #ADOPTION OR #RESCUE ASAP\n ID 632997 - Hansel, M\n I...', 'A sweet
 photo of Otis from four years ago. #GoodbyeOtis #memories #cat
<https://t.co/IyBgwf7QnS>', 'RT @greatstuffuneeed: Cat and Basket Vintage Metal
 Lunch Box Signed Pam Jacobsen Hand Painted Design #vintage #cat #metal #lunch
 #box <http...>', 'Round Shaped Pet ID Tags Set\n\$27.72 and FREE Shipping\n\nTag a
 friend who would love this!\n\n #petgroomer #cat #cats...
<https://t.co/jAqKv5LNsX>', 'RT @tulipteardrops: #cute #siamese #miniart #handmade
 #cat Siamese cat watercolor ACEO original painting #Miniature
<https://t.co/jGCVghxci...>', 'RT @sachikoko: SWEET 3MO #TABBY #KITTENS "HANSEL"
 & "GRETEL" \n DUMPED BY OWNER\n NEED #ADOPTION OR #RESCUE ASAP\n ID 632997
 - Hansel, M\n I...', 'Cat all the time ! #cat <https://t.co/bZ2Y9IcMSL>', 'RT
 @necotogoron: before→after\n#cat # \u3000# <https://t.co/PlmZFIC7J9>', 'RT
 @sachikoko: SWEET 1YO WHITE & BLACK KITTY "TALIO" \n NEEDS #ADOPTION OR
 #RESCUE \n ID 632914 Cage PT1 <https://t.co/x8tiPB8b9X>\n #ADOPT #P...', '#Missing
 male silver tabby #cat - Pleasant View Subdivision, #Lynchburg, VA, USA
<https://t.co/zyQ2igT3sl>', 'RT @Snishaa2: Good night Twitter friends... \nHave a
 peaceful night and Sweet\ndreams.. \nMaybe people can learn from these sweet
 animals. ...', 'RT @WonderfulSpace4: Lovely new month dear
 friends... \n\n#GoodMorningEveryone #November1st #HappyNewWeek #CoffeeTime
 #AutumnFalls ...', 'RT @sachikoko: SWEET 1YO WHITE & BLACK KITTY
 "TALIO" \n NEEDS #ADOPTION OR #RESCUE \n ID 632914 Cage PT1

Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @bandidosauces: How about a healthy tasty treat for your puppers? We got you.. marrow bones and liver slow cooked for 18 hours. Your do...', 'Paws for effect! Score great deals on gifts for all your furry friends at <https://t.co/xJtxYOHuoB>.\n.\n.\n.\n#dog #pet... <https://t.co/3fjEDvuRFE>', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @ShadowDogDesign: I will take over the world . . . right after this nap.\n\n#Dogs #Dog #Puppy #BestFriend #DogsAreLove #TakeOverTheWorld h...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @favcorgi_usa: The laziest horror movie villain ever\n#corgi #corgitwitter #dog <https://t.co/dcHs6nYlF8>', 'A boy and his dog... beautiful!', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @dog_coin_: If you are #shiba influencer with significant engagement from #SHIBARMY reach out with pricing fro shilling our 7 % #shibai...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @Kindness626: this is #urgent Kramer is #feeling really #sad no one seems interested in him. He needs a #foster or gwt adopted. Plea...', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @AdoptCCAnimals: #Dog #Remo_CCSTCA_01 I know how to sit!! Especially if you give me a treat <https://t.co/BeuA6oW7DE> <https://t.co/aqUjYMD...>', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'Magical Tours\nEnter The Lake Tahoe Gallery\nArt for sale\n\n<https://t.co/yoYwIBeZG4> \n#Prints #Canvas #Totes #Puzzles...


```
[5]: print(CompleteTweetList)
      print(CompleteLabelsList)
```

```
['RT @sachikoko: CODE RED RELISTED TBD TUE 11/2/21 \nTIMID, STRESSED 2YO WHITE
& #TABBY KITTY "TAZ" \n DUMPED BY OWNER 4 LITTERBOX ISSUES\n N...', 'RT
@WonderfulSpace4: Lovely new month dear friends... \n\n#GoodMorningEveryone
#November1st #HappyNewWeek #CoffeeTime #AutumnFalls ...', 'RT @manytle:
New Drop no.11 at Tungmeng Life Collection\nTungmeng travel Queen Sirikit
Botanic Garden,\nChiang Mai, Thailand \nAvailable on...', 'RT @_dalim44:
" " ,Have a meow day everyone next painting will be for Odd I won\'t
forget @jaybnow_hr \n#JAYB #Def #Jaebeom #GOT7 #...', "she's cute!!! #pet #cat
#cute #adorable #cute #kitten #petfinder I'm looking for a home!
https://t.co/k2Qnfqxyzls https://t.co/bdtAUDmUm2", ' PLEASE HELP
SAVE ', 'RT @Dizzle_tan: Commission for anonymous of felicia, thanks for the
support. \n\n#Darkstalkers #oppai #Nsfw #digitalartwork #Capcom #Felicia...',
"'That's What I'm Tacoin' About!' Premium T-Shirt https://t.co/MmEmD11ASl #cute
#taco #funny #food #tacos #tacobell... https://t.co/DTv1QOVdn4", 'RT @MShagalaih:
Palestine Street Angeles Need the #Help and support of their savior
@MShagalaih in securing more #Cat food that is suffic...', ' PLEASE HELP
SAVE ', 'How cute?? 2 tap if you liked \n.\n.\nDouble tap my little heart
if you love me \n.\n.\nTag your friends who love...
https://t.co/rUYQotIH6X', 'Omg Felicia is my favorite darkstalkers woman',
' PLEASE HELP SAVE ', 'RT @Dizzle_tan: Commission for @herrow52 of
felicia, thanks for the support \n\n#Darkstalkers #Oppai #Nsfw #DigitalArtist
#capcom #Felicia...', ' PLEASE HELP SAVE ', ' PLEASE HELP
SAVE ', 'RT @manytle: New Drop no.11 at Tungmeng Life Collection\nTungmeng
travel Queen Sirikit Botanic Garden,\nChiang Mai, Thailand \nAvailable on...',
'This #Influencer #Tattooed Her #Cat and Faced Zero Repercussions! #care2
https://t.co/z71ZUjaJdF', 'RT @manytle: New Drop no.11 at Tungmeng Life
Collection\nTungmeng travel Queen Sirikit Botanic Garden,\nChiang Mai, Thailand
\nAvailable on...', 'RT @HoudiniMeow: RETWEET, PLEASE VOTE FOR ME. Thank
mew all! You can vote 10x a day, refresh the web page to do this after every
v...', 'RT @tulipteardrops: #cute #siamese #miniart #handmade #cat Siamese cat
watercolor ACEO original painting #Miniature https://t.co/jGCVghxci...', 'RT
@sachikoko: SWEET 3MO #TABBY #KITTENS "HANSEL" & "GRETEL" \n DUMPED BY
OWNER\n NEED #ADOPTION OR #RESCUE ASAP\n ID 632997 - Hansel, M\n I...', 'A sweet
photo of Otis from four years ago. #GoodbyeOtis #memories #cat
https://t.co/IyBgwf7QnS', 'RT @greatstuffneed: Cat and Basket Vintage Metal
Lunch Box Signed Pam Jacobsen Hand Painted Design #vintage #cat #metal #lunch
#box http...', 'Round Shaped Pet ID Tags Set\n$27.72 and FREE Shipping\n\nTag a
friend who would love this!\n\n #petgroomer #cat #cats...
https://t.co/jAqKv5LNsX', 'RT @tulipteardrops: #cute #siamese #miniart #handmade
#cat Siamese cat watercolor ACEO original painting #Miniature
https://t.co/jGCVghxci...', 'RT @sachikoko: SWEET 3MO #TABBY #KITTENS "HANSEL"
& "GRETEL" \n DUMPED BY OWNER\n NEED #ADOPTION OR #RESCUE ASAP\n ID 632997
- Hansel, M\n I...', 'Cat all the time ! #cat https://t.co/bZ2Y9IcMSL', 'RT
@necotogoron: before→after\n#cat # \u3000# https://t.co/PlmZFiC7J9', 'RT
@sachikoko: SWEET 1YO WHITE & BLACK KITTY "TALIO" \n NEEDS #ADOPTION OR
```

#RESCUE \n ID 632914 Cage PT1 <https://t.co/x8tiPB8b9X>\n #ADOPT #P...', '#Missing male silver tabby #cat - Pleasant View Subdivision, #Lynchburg, VA, USA <https://t.co/zyQ2igT3sl>', 'RT @Snishaa2: Good night Twitter friends... \nHave a peaceful night and Sweet\ndreams.. \nMaybe people can learn from these sweet animals. ...', 'RT @WonderfulSpace4: Lovely new month dear friends... \n\n#GoodMorningEveryone #November1st #HappyNewWeek #CoffeeTime #AutumnFalls ...', 'RT @sachikoko: SWEET 1YO WHITE & BLACK KITTY "TALIO" \n NEEDS #ADOPTION OR #RESCUE \n ID 632914 Cage PT1 <https://t.co/x8tiPB8b9X>\n #ADOPT #P...', 'RT @AdoptMePimaAZ: Adoptable #cat #BuzzBandit_PIMAAZ_01 cute markings <https://t.co/w4DHbXAqcS> <https://t.co/wlE90uzs86>', 'RT @AdoptMePimaAZ: Adoptable #Cat #Alex_PIMAAZ_02 Do you have a sunny spot in your home just waiting for a kitten? I might be the one! <http://t.co/7mZnLflyLQ>', 'Ancient Lizard Tales #LOLcat #cats #LOLcats #cat <https://t.co/E1ABJGc57x>', 'RT @AdoptMePimaAZ: Adoptable #Cat #Harvey_PIMAAZ_02 Harvey is just the sweetest guy! Loves pets. <https://t.co/7mZnLflyLQ>. <https://t.co/JPycK3n5Dz> via @YouTube #ClaireLuvcat... <https://t.co/D3ST6SZEUI>', 'RT @manytle: New Drop no.11 at Tungmeng Life Collection\nTungmeng travel Queen Sirikit Botanic Garden,\nChiang Mai, Thailand \nAvailable on...', 'Drink water #CatsOfTwitter #cats #cutecats #cuteanimals #cat <https://t.co/UUYIXEWjK4>', '#Action #Alert! #Stop the #SouthKorean #dog and #cat #meat #cruelty! #dogmeat #catmeat #animalcruelty... <https://t.co/sqhTpxpRQM>', 'RT @AdoptMePimaAZ: Adoptable #Cat #SwissChard_PIMAAZ_01 Enjoyed giving her lots of pets on recent visit <https://t.co/1S1sDpXWg0> <https://t.co/1S1sDpXWg0>', 'RT @AdoptMePimaAZ: Adoptable #Cat #Drobbie_PIMAAZ_01 Had fun playing with him on recent visit! <https://t.co/nWu86eLrVT> <https://t.co/Bjs...>', 'RT @ClothCrow: Majestic cat is majestic #CatsOfTwitter #cat <https://t.co/ZGC2fKWJbG>', 'Get my art printed on awesome products. Support me at Redbubble #RBandME: <https://t.co/A5gMPV4UCW> #findyourthing... <https://t.co/07CEz0os6F>', 'RT @AdoptMePimaAZ: Adoptable #Cat #Alex_PIMAAZ_02 Do you have soft blankets I can make biscuits in? I would love a person to snuggle! <https://t.co/07CEz0os6F>', 'RT @manytle: New Drop no.11 at Tungmeng Life Collection\nTungmeng travel Queen Sirikit Botanic Garden,\nChiang Mai, Thailand \nAvailable on...', 'RT @manytle: New Drop no.11 at Tungmeng Life Collection\nTungmeng travel Queen Sirikit Botanic Garden,\nChiang Mai, Thailand \nAvailable on...', 'A real pickup for you!\n\n<https://t.co/4ZkZLcvjbl> \n #happy #halloween #black #cat #pumpkin #usa #detroit #shopper... <https://t.co/4p0f0stC6J>', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @ShadowDogDesign: So, Monday . . . We meet again.\n\n#Dogs #Dog #BestFriend #DogsAreLove #ExceptOnMondayMorning <https://t.co/w4Tz3u6k8v>', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @Crazyrottweiler1: How is this cute fight???\n\n#rottweiler #rottie #puppy #dog <https://t.co/X02Q0a1lFk>', 'Our first post is a story about finding the right place to be.\n\n#dogs #dog #shortstory #video #music #FirstPost... <https://t.co/cTvPOFYgsC>', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @bandidosauces: How about a healthy tasty treat for your

puppers? We got you.. marrow bones and liver slow cooked for 18 hours. Your do...', 'Paws for effect! Score great deals on gifts for all your furry friends at <https://t.co/xJtxYOHuoB>.\n.\n.\n.\n#dog #pet... <https://t.co/3fjEDvuRFE>', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @ShadowDogDesign: I will take over the world . . . right after this nap.\n\n#Dogs #Dog #Puppy #BestFriend #DogsAreLove #TakeOverTheWorld h...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @favcorgi_usa: The laziest horror movie villain ever\n#corgi #corgitwitter #dog <https://t.co/dcHs6nYlF8>', 'A boy and his dog... beautiful!', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @dog_coin_: If you are #shiba influencer with significant engagement from #SHIBARMY reach out with pricing fro shilling our 7 % #shibai...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @Kindness626: this is #urgent Kramer is #feeling really #sad no one seems interested in him. He needs a #foster or gwt adopted. Plea...', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @AdoptCCAnimals: #Dog #Remo_CCSTCA_01 I know how to sit!! Especially if you give me a treat <https://t.co/BeuA6oW7DE> <https://t.co/aqUjYMDm>', 'RT @SafeShibaMoon: #SafeShibaMoon #Sol\n\nA decentralized, deflationary, solana based #doge #meme token with inbuilt charity support\n\nWith #S...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'RT @RealAkamaruInu: Talk about how much you love Akamaru by using the #akamaruinu hashtag along with the hashtags below.\n\n10 people will wi...', 'Magical Tours\nEnter The Lake Tahoe Gallery\nArt for sale\n\n<https://t.co/yoYwIBeZG4> \n#Prints #Canvas #Totes #Puzzles... <https://t.co/kXREPB3zCo>', 'RT @penny1332: BEGGING FOR COLIN TO HAVE A CHANCE TO LEARN HOW SAFE & COMFORTABLE LIFE CAN BE , HE DOES NOT HOPE BECAUSE HE HAS

3 Basic EDA: WordCloud

Build a wordcloud for each class and identify high frequency distinguishing words.

```
[6]: import wordcloud
from wordcloud import WordCloud, STOPWORDS

ind1 = (pd.DataFrame(CompleteLabelsList) == label1)[0].values.tolist()
ind2 = (pd.DataFrame(CompleteLabelsList) == label2)[0].values.tolist()
#print(ind1)
slice1 = list(filter(lambda i: ind1[i], range(len(ind1))))
slice2 = list(filter(lambda i: ind2[i], range(len(ind2))))

text1 = [CompleteTweetList[i] for i in slice1]
text2 = [CompleteTweetList[i] for i in slice2]

print(text1)

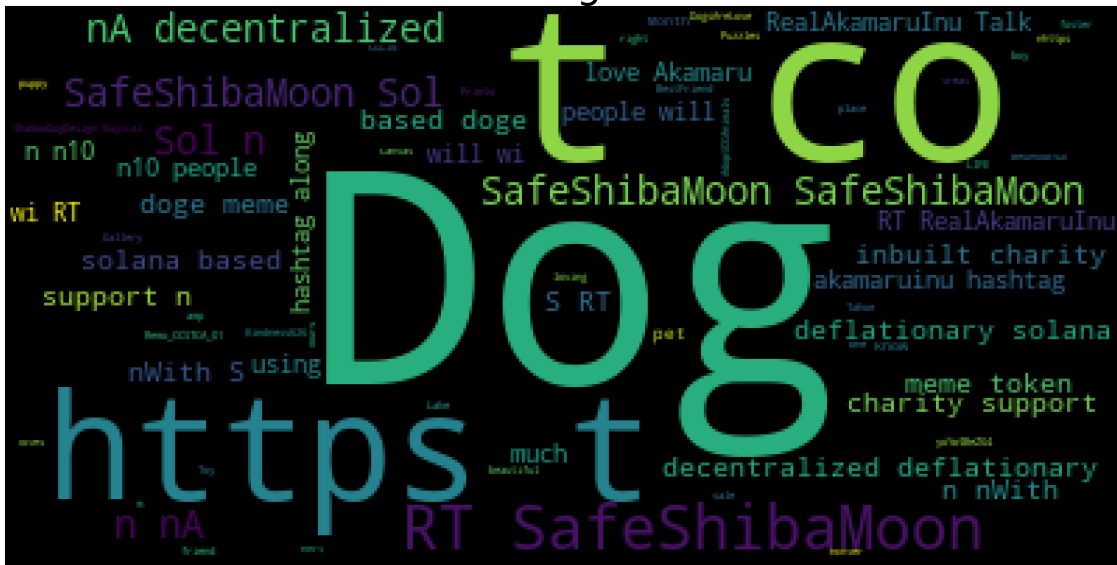
wordcloud = WordCloud().generate(str(text1))
# Open a plot of the generated image.
#figure(figsize = (20,2))
plt.figure(figsize=(50,40))
plt.imshow(wordcloud)
          #, aspect="auto")
plt.title(label1,fontsize=100)
plt.axis("off")
plt.show()

wordcloud = WordCloud().generate(str(text2))
# Open a plot of the generated image.
#figure(figsize = (20,2))
plt.figure(figsize=(50,40))
plt.imshow(wordcloud)
          #, aspect="auto")
plt.title(label2, fontsize=100)
plt.axis("off")
plt.show()
```

```
['RT @sachikoko: CODE RED RELISTED TBD TUE 11/2/21 \n TIMID, STRESSED 2YO WHITE
&amp; #TABBY KITTY "TAZ" \n DUMPED BY OWNER 4 LITTERBOX ISSUES\n N...', 'RT
@WonderfulSpace4: Lovely new month dear friends... \n\n#GoodMorningEveryone
#November1st #HappyNewWeek #CoffeeTime #AutumnFalls ...', 'RT @manytle:
New Drop no.11 at Tungmeng Life Collection\nTungmeng travel Queen Sirikit
Botanic Garden,\nChiang Mai, Thailand \nAvailable on...', 'RT @_dalim44:
" " ,Have a meow day everyone next painting will be for Odd I won\t
```

forget @jaybnow_hr \n#JAYB #Def #Jaebeom #GOT7 #...', "she's cute!!! #pet #cat
 #cute #adorable #cute #kitten #petfinder I'm looking for a home!
<https://t.co/k2Qnfqxyzls> <https://t.co/bdtAUDmUm2>", ' PLEASE HELP
 SAVE ', 'RT @Dizzle_tan: Commission for anonymous of felicia, thanks for the
 support. \n\n#Darkstalkers #oppai #Nsfw #digitalartwork #Capcom #Felicia...',
 "'That's What I'm Tacoin' About!' Premium T-Shirt <https://t.co/MmEmD11ASl> #cute
 #taco #funny #food #tacos #tacobell... <https://t.co/DTv1QOVdn4>", 'RT @MShagalaih:
 Palestine Street Angeles Need the #Help and support of their savior
 @MShagalaih in securing more #Cat food that is suffic...', ' PLEASE HELP
 SAVE ', 'How cute?? 2 tap if you liked \n.\n.\nDouble tap my little heart
 if you love me \n.\n.\nTag your friends who love...
<https://t.co/rUYQotIH6X>', 'Omg Felicia is my favorite darkstalkers woman',
 ' PLEASE HELP SAVE ', 'RT @Dizzle_tan: Commission for @herrow52 of
 felicia, thanks for the support \n\n#Darkstalkers #Oppai #Nsfw #DigitalArtist
 #capcom #Felicia...', ' PLEASE HELP SAVE ', ' PLEASE HELP
 SAVE ', 'RT @manytle: New Drop no.11 at Tungmeng Life Collection\nTungmeng
 travel Queen Sirikit Botanic Garden,\nChiang Mai, Thailand \nAvailable on...',
 'This #Influencer #Tattooed Her #Cat and Faced Zero Repercussions! #care2
<https://t.co/z71ZUjaJdF>', 'RT @manytle: New Drop no.11 at Tungmeng Life
 Collection\nTungmeng travel Queen Sirikit Botanic Garden,\nChiang Mai, Thailand
 \nAvailable on...', 'RT @HoudiniMeow: RETWEET, PLEASE VOTE FOR ME. Thank
 mew all! You can vote 10x a day, refresh the web page to do this after every
 v...', 'RT @tulipteardrops: #cute #siamese #miniart #handmade #cat Siamese cat
 watercolor ACEO original painting #Miniature <https://t.co/jGCVghxci...>', 'RT
 @sachikoko: SWEET 3MO #TABBY #KITTENS "HANSEL" & "GRETEL" \n DUMPED BY
 OWNER\n NEED #ADOPTION OR #RESCUE ASAP\n ID 632997 - Hansel, M\n I...', 'A sweet
 photo of Otis from four years ago. #GoodbyeOtis #memories #cat
<https://t.co/IyBgwf7QnS>', 'RT @greatstuffneed: Cat and Basket Vintage Metal
 Lunch Box Signed Pam Jacobsen Hand Painted Design #vintage #cat #metal #lunch
 #box http...', 'Round Shaped Pet ID Tags Set\n\$27.72 and FREE Shipping\n\nTag a
 friend who would love this!\n\n #petgroomer #cat #cats...
<https://t.co/jAqKv5LNsX>', 'RT @tulipteardrops: #cute #siamese #miniart #handmade
 #cat Siamese cat watercolor ACEO original painting #Miniature
<https://t.co/jGCVghxci...>', 'RT @sachikoko: SWEET 3MO #TABBY #KITTENS "HANSEL"
 & "GRETEL" \n DUMPED BY OWNER\n NEED #ADOPTION OR #RESCUE ASAP\n ID 632997
 - Hansel, M\n I...', 'Cat all the time ! #cat <https://t.co/bZ2Y9IcMSL>', 'RT
 @necotogoron: before→after\n#cat # \u3000# <https://t.co/PlmZFIC7J9>', 'RT
 @sachikoko: SWEET 1YO WHITE & BLACK KITTY "TALIO" \n NEEDS #ADOPTION OR
 #RESCUE \n ID 632914 Cage PT1 <https://t.co/x8tiPB8b9X>\n #ADOPT #P...', '#Missing
 male silver tabby #cat - Pleasant View Subdivision, #Lynchburg, VA, USA
<https://t.co/zyQ2igT3sl>', 'RT @Snishaa2: Good night Twitter friends... \nHave a
 peaceful night and Sweet\ndreams.. \nMaybe people can learn from these sweet
 animals. ...', 'RT @WonderfulSpace4: Lovely new month dear
 friends... \n\n#GoodMorningEveryone #November1st #HappyNewWeek #CoffeeTime
 #AutumnFalls ...', 'RT @sachikoko: SWEET 1YO WHITE & BLACK KITTY
 "TALIO" \n NEEDS #ADOPTION OR #RESCUE \n ID 632914 Cage PT1
<https://t.co/x8tiPB8b9X>\n #ADOPT #P...', 'RT @AdoptMePimaAZ: Adoptable #cat
 #BuzzBandit_PIMAAZ_01 cute markings <https://t.co/w4DHbXAqcS>

dog



3.1 Clean and Tidy the Data

3.1.1 Intermediate File Save

The next goal is to create a csv file where the first column is the label which is the hashtag and the second column is the entire tweet WITH NO newlines, etc.

```
[7]: TW_file="TweetCSV2.csv"
FILE=open(TW_file, "w")
WhatToWrite="LABEL,Tweet\n"
FILE.write(WhatToWrite)
FILE.close()

## This fun python trick allows you to loop
## through two lists at once and together
ZIP_List = zip(CompleteTweetList, CompleteLabelsList)

with open(TW_file, "a") as f:
    for TW, Lab in ZIP_List:
        TW=TW.replace("\n", " ")
        TW=TW.replace(",", " ")
        TW=re.sub(r"^[A-Za-z\-\]", " ", TW)
        TW=re.sub(r"\bhttp\b | \bhttps\b", " ", TW)
        ## Why do we need to do this?
        #print("NEXT\n")
        print(TW)
        print(Lab)
        WhatToWrite=str(Lab+", "+TW+"\n")
```


f.write(WhatToWrite)

RT sachikoko CODE RED RELISTED TBD TUE TIMID STRESSED YO WHITE
amp TABBY KITTY TAZ DUMPED BY OWNER LITTERBOX ISSUES N
cat

RT WonderfulSpace Lovely new month dear friends GoodMorningEveryone
November st HappyNewWeek CoffeeTime AutumnFalls
cat

RT manytle New Drop no at Tungmeng Life Collection Tungmeng travel Queen
Sirikit Botanic Garden Chiang Mai Thailand Available on
cat

RT dalim Have a meow day everyone next painting will be for
Odd I won t forget jaybnow hr JAYB Def Jaebeom GOT
cat

she s cute pet cat cute adorable cute kitten petfinder I m looking for
a home t co k Qnfqxzls t co bdtAUDmUm
cat

PLEASE HELP SAVE

cat
RT Dizzle tan Commission for anonymous of felica thanks for the support
Darkstalkers oppai Nsfw digitalartwork Capcom Felicia
cat

That s What I m Tacoin About Premium T-Shirt t co MmEmD ASl cute taco
funny food tacos tacobell t co DTv QOVdn
cat

RT MShagalaih Palestine Street Angeles Need the Help and support of their
savior MShagalaih in securing more Cat food that is suffic
cat

PLEASE HELP SAVE

cat
How cute tap if you liked Double tap my little heart if you love
me Tag your friends who love t co rUYQotIH X
cat

Omg Felicia is my favorite darkstalkers woman
cat

PLEASE HELP SAVE

cat
RT Dizzle tan Commission for herrow of felicia thanks for the support
Darkstalkers Oppai Nsfw DigitalArtist capcom Felicia
cat

PLEASE HELP SAVE

cat
PLEASE HELP SAVE

cat
RT manytle New Drop no at Tungmeng Life Collection Tungmeng travel Queen
Sirikit Botanic Garden Chiang Mai Thailand Available on
cat

This Influencer Tattooed Her Cat and Faced Zero Repercussions care t co
z ZUJaJdF

cat

RT manytle New Drop no at Tungmeng Life Collection Tungmeng travel Queen
Sirikit Botanic Garden Chiang Mai Thailand Available on

cat

RT HoudiniMeow RETWEET PLEASE VOTE FOR ME Thank mew all You can
vote x a day refresh the web page to do this after every v

cat

RT tulipteardrops cute siamese miniart handmade cat Siamese cat
watercolor ACEO original painting Miniature t co jGCVghxci

cat

RT sachikoko SWEET MO TABBY KITTENS HANSEL amp GRETEL DUMPED BY
OWNER NEED ADOPTION OR RESCUE ASAP ID - Hansel M I

cat

A sweet photo of Otis from four years ago GoodbyeOtis memories cat t co
IyBgwf QnS

cat

RT greatstuffuneeed Cat and Basket Vintage Metal Lunch Box Signed Pam Jacobsen
Hand Painted Design vintage cat metal lunch box

cat

Round Shaped Pet ID Tags Set and FREE Shipping Tag a friend who would
love this petgroomer cat cats t co jAqKv LNsX

cat

RT tulipteardrops cute siamese miniart handmade cat Siamese cat
watercolor ACEO original painting Miniature t co jGCVghxci

cat

RT sachikoko SWEET MO TABBY KITTENS HANSEL amp GRETEL DUMPED BY
OWNER NEED ADOPTION OR RESCUE ASAP ID - Hansel M I

cat

Cat all the time cat t co bZ Y IcMSL

cat

RT necotogoron before after cat t co PlmZFIC J

cat

RT sachikoko SWEET YO WHITE amp BLACK KITTY TALIO NEEDS ADOPTION OR
RESCUE ID Cage PT t co x tiPB b X ADOPT P

cat

Missing male silver tabby cat Pleasant View Subdivision Lynchburg VA USA
t co zyQ igT sl

cat

RT Snishaa Good night Twitter friends Have a peaceful night and Sweet
dreams Maybe people can learn from these sweet animals

cat

RT WonderfulSpace Lovely new month dear friends GoodMorningEveryone
November st HappyNewWeek CoffeeTime AutumnFalls

cat

RT sachikoko SWEET YO WHITE amp BLACK KITTY TALIO NEEDS ADOPTION OR
RESCUE ID Cage PT t co x tiPB b X ADOPT P

cat
RT AdoptMePimaAZ Adoptable cat BuzzBandit PIMAAZ cute markings t co w
DHbXAqcS t co wLE Ouzs
cat
RT AdoptMePimaAZ Adoptable Cat Alex PIMAAZ Do you have a sunny spot in
your home just waiting for a kitten I might be the one
cat
Ancient Lizard Tales LOLcat cats LOLcats cat t co E ABjGc x
cat
RT AdoptMePimaAZ Adoptable Cat Harvey PIMAAZ Harvey is just the
sweetest guy Loves pets t co mZnLflyLQ t co
cat
LaLa s Mean Streak Is Real I m So Sorry ENG SUB t co JPycK n Dz via
YouTube ClaireLuvcat t co D ST SZEUI
cat
RT manytle New Drop no at Tungmeng Life Collection Tungmeng travel Queen
Sirikit Botanic Garden Chiang Mai Thailand Available on
cat
Drink water CatsOfTwitter cats cutecats cuteanimals cat t co UUYIXEWjK
cat
Action Alert Stop the SouthKorean dog and cat meat cruelty
dogmeat catmeat animalcruelty t co sqhTpxpRQM
cat
RT AdoptMePimaAZ Adoptable Cat SwissChard PIMAAZ Enjoyed giving her
lots of pets on recent visit t co S sDpXWgO
cat
RT AdoptMePimaAZ Adoptable Cat Drobbie PIMAAZ Had fun playing with him
on recent visit t co nWu eLrVT t co Bjs
cat
RT ClothCrow Majestic cat is majestic CatsOfTwitter cat t co ZGC fKWJbG
cat
Get my art printed on awesome products Support me at Redbubble RBandME t
co A gMPV UCW findyourthing t co O CEzOos F
cat
RT AdoptMePimaAZ Adoptable Cat Alex PIMAAZ Do you have soft blankets I
can make biscuits in I would love a person to snuggle
cat
RT manytle New Drop no at Tungmeng Life Collection Tungmeng travel Queen
Sirikit Botanic Garden Chiang Mai Thailand Available on
cat
RT manytle New Drop no at Tungmeng Life Collection Tungmeng travel Queen
Sirikit Botanic Garden Chiang Mai Thailand Available on
cat
A real pickup for you t co ZkZLcvjbl happy halloween black cat
pumpkin usa detroit shopper t co pOf stC J
cat
RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana
based doge meme token with inbuilt charity support With S

dog

RT ShadowDogDesign So Monday We meet again Dogs Dog BestFriend
DogsAreLove ExceptOnMondayMorning t co w Tz u k v

dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the
akamaruinu hashtag along with the hashtags below people will wi

dog

RT Crazyrottweiler How is this cute fight rottweiler rottie puppy dog
t co X QOa lFk

dog

Our first post is a story about finding the right place to be dogs dog
shortstory video music FirstPost t co cTvPOFYgsC

dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the
akamaruinu hashtag along with the hashtags below people will wi

dog

RT bandidosauces How about a healthy tasty treat for your puppies We got you
marrow bones and liver slow cooked for hours Your do

dog

Paws for effect Score great deals on gifts for all your furry friends at t
co xJtxYOHuoB dog pet t co fjEDvuRFE

dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana
based doge meme token with inbuilt charity support With S

dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana
based doge meme token with inbuilt charity support With S

dog

RT ShadowDogDesign I will take over the world right after this nap
Dogs Dog Puppy BestFriend DogsAreLove TakeOverTheWorld h

dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the
akamaruinu hashtag along with the hashtags below people will wi

dog

RT favcorgi usa The laziest horror movie villain ever corgi corgitwitter
dog t co dcHs nYlF

dog

A boy and his dog beautiful

dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana
based doge meme token with inbuilt charity support With S

dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana
based doge meme token with inbuilt charity support With S

dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the
akamaruinu hashtag along with the hashtags below people will wi

dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the akamaruinu hashtag along with the hashtags below people will wi
dog

RT dog coin If you are shiba influencer with significant engagement from SHIBARMY reach out with pricing fro shilling our shibai
dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the akamaruinu hashtag along with the hashtags below people will wi
dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the akamaruinu hashtag along with the hashtags below people will wi
dog

RT Kindness this is urgent Kramer is feeling really sad no one seems interested in him He needs a foster or gwt adopted Plea
dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana based doge meme token with inbuilt charity support With S
dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana based doge meme token with inbuilt charity support With S
dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana based doge meme token with inbuilt charity support With S
dog

RT AdoptCCAnimals Dog Remo CCSTCA I know how to sit Especially if you give me a treat t co BeuA oW DE t co aqUjYMd
dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana based doge meme token with inbuilt charity support With S
dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the akamaruinu hashtag along with the hashtags below people will wi
dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the akamaruinu hashtag along with the hashtags below people will wi
dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the akamaruinu hashtag along with the hashtags below people will wi
dog

Magical Tours Enter The Lake Tahoe Gallery Art for sale t co yoYwIBeZG
Prints Canvas Totes Puzzles t co kXREPB zCo
dog

RT penny BEGGING FOR COLIN TO HAVE A CHANCE TO LEARN HOW SAFE amp COMFORTABLE LIFE CAN BE HE DOES NOT HOPE BECAUSE HE HAS NEVER EXPER
dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana based doge meme token with inbuilt charity support With S
dog

RT PhotoSarsfield Magical Tours Enter The Lake Tahoe Gallery Art for sale
t co yoYwIBeZG Prints Canvas Totes Puzzles Swea
dog

The peace of a dog sleeping at home One month being part of our family
sleepingbeauty pet onemothlater t co IrcEBESDlj
dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the
akamaruinu hashtag along with the hashtags below people will wi
dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana
based doge meme token with inbuilt charity support With S
dog

RT AdoptCCAnimals Dog Remo CCSTCA I m happy to go out w a two legged
friend t co BeuA oW DE t co YU wU GO
dog

RT RealAkamaruInu Talk about how much you love Akamaru by using the
akamaruinu hashtag along with the hashtags below people will wi
dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana
based doge meme token with inbuilt charity support With S
dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana
based doge meme token with inbuilt charity support With S
dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana
based doge meme token with inbuilt charity support With S
dog

RT Kindness LisaBrabson enaceccyy MamaPKL LilacMoon TammieHedden
ArizonaStation NMBewitched klyonskalmenson KathiuskaFon
dog

RT ZadenZane Bertie s ever so dignified amp restrained with a gentle tail
wobble and a polite lick of his chops waiting patiently I know
dog

I m not gonna apologize for being an artist schmutzy fineart art dog dogs
rescuedog brindle t co ko RQDdxH
dog

RT AdoptMePimaAZ Adoptable Dog Rio PIMAAZ Beautiful boy loving life in
foster t co kcFXILex W t co OIztnYkwWT
dog

RT cookiedakota Angels of the Planet please get Colin a good loving
place b look this little Guy
dog

RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana
based doge meme token with inbuilt charity support With S
dog

Very Classic Dog Rope Toy Hot selling more than years Highly recommend why
not dog toy pet products pet s t co McZ Y QJPL
dog

Hey Hudson the August Dog Star of the Month wants to know if you have entered the November Dog of the Month contest to xcfQXdwnz dog

3.1.2 Clean, Stem, and Vectorize

Using CountVectorizer, the data is cleaned, stopwords and 'garbage' are removed, and the data is vectorized. The result is a document term matrix.

```
[8]: import nltk
nltk.download('wordnet')
from nltk import word_tokenize
from nltk.stem import WordNetLemmatizer

class LemmaTokenizer:
    def __init__(self):
        self.wnl = WordNetLemmatizer()
    def __call__(self, doc):
        return [self.wnl.lemmatize(t) for t in word_tokenize(doc)]

MyCV_T=CountVectorizer(input='content',
                        stop_words='english',
                        ngram_range=(1,1),
                        tokenizer=LemmaTokenizer(),
                        max_features=100
                        )

### READ our csv file of tweets into two lists
## The first is the label, second the tweet

FileContents=pd.read_csv(TW_file)
print(FileContents["LABEL"])
print(FileContents["Tweet"])

## CONVERT the tweets to a LIST so we can use
## CountVectorizer
TweetsList = FileContents["Tweet"].to_list()
print(TweetsList)

LABEL_List = FileContents["LABEL"].to_list()
print(LABEL_List)

My_DTM_T=MyCV_T.fit_transform(TweetsList)
```

```
[nltk_data] Downloading package wordnet to
[nltk_data] C:\Users\jerem\AppData\Roaming\nltk_data...
```

[nltk_data] Package wordnet is already up-to-date!

0 cat
1 cat
2 cat
3 cat
4 cat

...

95 dog
96 dog
97 dog
98 dog
99 dog

Name: LABEL, Length: 100, dtype: object

0 RT sachikoko CODE RED RELISTED TBD TUE ...
1 RT WonderfulSpace Lovely new month dear fri...
2 RT manytle New Drop no at Tungmeng Life ...
3 RT dalim Have a meow day eve...
4 she s cute pet cat cute adorable cute ...

...

95 RT AdoptMePimaAZ Adoptable Dog Rio PIMAAZ ...
96 RT cookiedakota Angels of the Planet p...
97 RT SafeShibaMoon SafeShibaMoon Sol A dece...
98 Very Classic Dog Rope Toy Hot selling more t...
99 Hey Hudson the August Dog Star of the Month ...

Name: Tweet, Length: 100, dtype: object

['RT sachikoko CODE RED RELISTED TBD TUE TIMID STRESSED YO WHITE
amp TABBY KITTY TAZ DUMPED BY OWNER LITTERBOX ISSUES N ', 'RT
WonderfulSpace Lovely new month dear friends GoodMorningEveryone
November st HappyNewWeek CoffeeTime AutumnFalls ', 'RT manytle
New Drop no at Tungmeng Life Collection Tungmeng travel Queen Sirikit Botanic
Garden Chiang Mai Thailand Available on ', 'RT dalim Have
a meow day everyone next painting will be for Odd I won t forget jayknow hr
JAYB Def Jaebeom GOT ', 'she s cute pet cat cute adorable cute
kitten petfinder I m looking for a home t co k Qnfqxzls t co
bdtAUDmUm ', ' PLEASE HELP SAVE ', 'RT Dizzle tan Commission for
anonymous of felica thanks for the support Darkstalkers oppai Nsfw
digitalartwork Capcom Felicia ', ' That s What I m Tacoin About Premium
T-Shirt t co MmEmD ASl cute taco funny food tacos tacobell t co
DTv QOVdn ', 'RT MShagalaih Palestine Street Angeles Need the Help and
support of their savior MShagalaih in securing more Cat food that is suffic ',
' PLEASE HELP SAVE ', 'How cute tap if you liked Double tap
my little heart if you love me Tag your friends who love t co
rUYQotIH X', 'Omg Felicia is my favorite darkstalkers woman', ' PLEASE HELP
SAVE ', 'RT Dizzle tan Commission for herrow of felicia thanks for
the support Darkstalkers Oppai Nsfw DigitalArtist capcom Felicia ', '
PLEASE HELP SAVE ', ' PLEASE HELP SAVE ', 'RT manytle New Drop
no at Tungmeng Life Collection Tungmeng travel Queen Sirikit Botanic Garden

Chiang Mai Thailand Available on ', 'This Influencer Tattooed Her Cat and Faced Zero Repercussions care t co z ZUJaJdF', 'RT manytle New Drop no at Tungmeng Life Collection Tungmeng travel Queen Sirikit Botanic Garden Chiang Mai Thailand Available on ', 'RT HoudiniMeow RETWEET PLEASE VOTE FOR ME Thank mew all You can vote x a day refresh the web page to do this after every v ', 'RT tulipteardrops cute siamese miniart handmade cat Siamese cat watercolor ACEO original painting Miniature t co jGCVghxci ', 'RT sachikoko SWEET MO TABBY KITTENS HANSEL amp GRETEL DUMPED BY OWNER NEED ADOPTION OR RESCUE ASAP ID - Hansel M I ', 'A sweet photo of Otis from four years ago GoodbyeOtis memories cat t co IyBgwf QnS', 'RT greatstuffneed Cat and Basket Vintage Metal Lunch Box Signed Pam Jacobsen Hand Painted Design vintage cat metal lunch box ', 'Round Shaped Pet ID Tags Set and FREE Shipping Tag a friend who would love this petgroomer cat cats t co jAqKv LNsX', 'RT tulipteardrops cute siamese miniart handmade cat Siamese cat watercolor ACEO original painting Miniature t co jGCVghxci ', 'RT sachikoko SWEET MO TABBY KITTENS HANSEL amp GRETEL DUMPED BY OWNER NEED ADOPTION OR RESCUE ASAP ID - Hansel M I ', 'Cat all the time cat t co bZ Y IcMSL', 'RT necotogoron before after cat t co PlmZFIC J ', 'RT sachikoko SWEET YO WHITE amp BLACK KITTY TALIO NEEDS ADOPTION OR RESCUE ID Cage PT t co x tiPB b X ADOPT P ', 'Missing male silver tabby cat Pleasant View Subdivision Lynchburg VA USA t co zyQ igT sl', 'RT Snishaa Good night Twitter friends Have a peaceful night and Sweet dreams Maybe people can learn from these sweet animals ', 'RT WonderfulSpace Lovely new month dear friends GoodMorningEveryone November st HappyNewWeek CoffeeTime AutumnFalls ', 'RT sachikoko SWEET YO WHITE amp BLACK KITTY TALIO NEEDS ADOPTION OR RESCUE ID Cage PT t co x tiPB b X ADOPT P ', 'RT AdoptMePimaAZ Adoptable cat BuzzBandit PIMAAZ cute markings t co w DHbXAqcS t co wLE OuZs ', 'RT AdoptMePimaAZ Adoptable Cat Alex PIMAAZ Do you have a sunny spot in your home just waiting for a kitten I might be the one ', 'Ancient Lizard Tales LOLcat cats LOLcats cat t co E ABjGc x', 'RT AdoptMePimaAZ Adoptable Cat Harvey PIMAAZ Harvey is just the sweetest guy Loves pets t co mZnLflyLQ t co ', 'LaLa s Mean Streak Is Real I m So Sorry ENG SUB t co JPycK n Dz via YouTube ClaireLuvcat t co D ST SZEUI', 'RT manytle New Drop no at Tungmeng Life Collection Tungmeng travel Queen Sirikit Botanic Garden Chiang Mai Thailand Available on ', 'Drink water CatsOfTwitter cats cutecats cuteanimals cat t co UUYIXEWjK ', 'Action Alert Stop the SouthKorean dog and cat meat cruelty dogmeat catmeat animalcruelty t co sqhTpxpRQM', 'RT AdoptMePimaAZ Adoptable Cat SwissChard PIMAAZ Enjoyed giving her lots of pets on recent visit t co S sDpXWg0 ', 'RT AdoptMePimaAZ Adoptable Cat Drobbie PIMAAZ Had fun playing with him on recent visit t co nWu eLrVT t co Bjs ', 'RT ClothCrow Majestic cat is majestic CatsOfTwitter cat t co ZGC fKWJbG', 'Get my art printed on awesome products Support me at Redbubble RBandME t co A gMPV UCW findyourthing t co O CEZ0os F', 'RT AdoptMePimaAZ Adoptable Cat Alex PIMAAZ Do you have soft blankets I can make biscuits in I would love a person to snuggle ', 'RT manytle New Drop no at Tungmeng Life Collection

Tungmeng travel Queen Sirikit Botanic Garden Chiang Mai Thailand Available
 on ', 'RT manytle New Drop no at Tungmeng Life Collection Tungmeng travel
 Queen Sirikit Botanic Garden Chiang Mai Thailand Available on ', 'A real
 pickup for you t co ZkZLcvjbl happy halloween black cat pumpkin
 usa detroit shopper t co pOf stC J', 'RT SafeShibaMoon SafeShibaMoon
 Sol A decentralized deflationary solana based doge meme token with inbuilt
 charity support With S ', 'RT ShadowDogDesign So Monday We meet again
 Dogs Dog BestFriend DogsAreLove ExceptOnMondayMorning t co w Tz u k v',
 'RT RealAkamaruInu Talk about how much you love Akamaru by using the
 akamaruinu hashtag along with the hashtags below people will wi ', 'RT
 Crazyrottweiler How is this cute fight rottweiler rotti puppy dog t
 co X QOa lFk', 'Our first post is a story about finding the right place to be
 dogs dog shortstory video music FirstPost t co cTvPOFYgsC', 'RT
 RealAkamaruInu Talk about how much you love Akamaru by using the akamaruinu
 hashtag along with the hashtags below people will wi ', 'RT bandidosaucos
 How about a healthy tasty treat for your puppers We got you marrow bones and
 liver slow cooked for hours Your do ', 'Paws for effect Score great deals
 on gifts for all your furry friends at t co xJtxYOHuoB dog pet t
 co fjEDvuRFE', 'RT SafeShibaMoon SafeShibaMoon Sol A decentralized
 deflationary solana based doge meme token with inbuilt charity support With
 S ', 'RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary
 solana based doge meme token with inbuilt charity support With S ', 'RT
 ShadowDogDesign I will take over the world right after this nap Dogs
 Dog Puppy BestFriend DogsAreLove TakeOverTheWorld h ', 'RT RealAkamaruInu
 Talk about how much you love Akamaru by using the akamaruinu hashtag along with
 the hashtags below people will wi ', 'RT favcorgi usa The laziest horror
 movie villain ever corgi corgitwitter dog t co dcHs nYlF ', 'A boy and his
 dog beautiful ', 'RT SafeShibaMoon SafeShibaMoon Sol A decentralized
 deflationary solana based doge meme token with inbuilt charity support With
 S ', 'RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary
 solana based doge meme token with inbuilt charity support With S ', 'RT
 RealAkamaruInu Talk about how much you love Akamaru by using the akamaruinu
 hashtag along with the hashtags below people will wi ', 'RT RealAkamaruInu
 Talk about how much you love Akamaru by using the akamaruinu hashtag along with
 the hashtags below people will wi ', 'RT dog coin If you are shiba
 influencer with significant engagement from SHIBARMY reach out with pricing fro
 shilling our shibai ', 'RT RealAkamaruInu Talk about how much you love
 Akamaru by using the akamaruinu hashtag along with the hashtags below
 people will wi ', 'RT RealAkamaruInu Talk about how much you love Akamaru by
 using the akamaruinu hashtag along with the hashtags below people will wi
 ', 'RT Kindness this is urgent Kramer is feeling really sad no one
 seems interested in him He needs a foster or gwt adopted Plea ', 'RT
 SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary solana based
 doge meme token with inbuilt charity support With S ', 'RT SafeShibaMoon
 SafeShibaMoon Sol A decentralized deflationary solana based doge meme
 token with inbuilt charity support With S ', 'RT SafeShibaMoon
 SafeShibaMoon Sol A decentralized deflationary solana based doge meme
 token with inbuilt charity support With S ', 'RT AdoptCCAnimals Dog Remo

CCSTCA I know how to sit Especially if you give me a treat t co BeuA oW
DE t co aqUjYMd ', 'RT SafeShibaMoon SafeShibaMoon Sol A decentralized
deflationary solana based doge meme token with inbuilt charity support With
S ', 'RT RealAkamaruInu Talk about how much you love Akamaru by using the
akamaruinu hashtag along with the hashtags below people will wi ', 'RT
RealAkamaruInu Talk about how much you love Akamaru by using the akamaruinu
hashtag along with the hashtags below people will wi ', 'RT RealAkamaruInu
Talk about how much you love Akamaru by using the akamaruinu hashtag along with
the hashtags below people will wi ', 'Magical Tours Enter The Lake Tahoe
Gallery Art for sale t co yoWIBeZG Prints Canvas Totes Puzzles t
co kXREPB zCo', 'RT penny BEGGING FOR COLIN TO HAVE A CHANCE TO LEARN HOW
SAFE amp COMFORTABLE LIFE CAN BE HE DOES NOT HOPE BECAUSE HE HAS NEVER EXPER
', 'RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary
solana based doge meme token with inbuilt charity support With S ', 'RT
PhotoSarsfield Magical Tours Enter The Lake Tahoe Gallery Art for sale t co
yoWIBeZG Prints Canvas Totes Puzzles Swea ', 'The peace of a dog
sleeping at home One month being part of our family sleepingbeauty pet
onemothlater t co IrcEBESDlj', 'RT RealAkamaruInu Talk about how much you
love Akamaru by using the akamaruinu hashtag along with the hashtags below
people will wi ', 'RT SafeShibaMoon SafeShibaMoon Sol A decentralized
deflationary solana based doge meme token with inbuilt charity support With
S ', 'RT AdoptCCAnimals Dog Remo CCSTCA I m happy to go out w a two
legged friend t co BeuA oW DE t co YU wU GO', 'RT RealAkamaruInu Talk
about how much you love Akamaru by using the akamaruinu hashtag along with the
hashtags below people will wi ', 'RT SafeShibaMoon SafeShibaMoon Sol A
decentralized deflationary solana based doge meme token with inbuilt charity
support With S ', 'RT SafeShibaMoon SafeShibaMoon Sol A decentralized
deflationary solana based doge meme token with inbuilt charity support With
S ', 'RT SafeShibaMoon SafeShibaMoon Sol A decentralized deflationary
solana based doge meme token with inbuilt charity support With S ', 'RT
Kindness LisaBrabson enaceccy MamaPKL LilacMoon TammieHedden
ArizonaStation NMBewitched klyonskalmenson KathiuskaFon ', 'RT ZadenZane
Bertie s ever so dignified amp restrained with a gentle tail wobble and a
polite lick of his chops waiting patiently I know ', 'I m not gonna apologize
for being an artist schmutzy fineart art dog dogs rescuedog brindle
t co ko RQDdxH', 'RT AdoptMePimaAZ Adoptable Dog Rio PIMAAZ Beautiful
boy loving life in foster t co kcFXILex W t co OIztnYkwWT', 'RT
cookiedakota Angels of the Planet please get Colin a good loving place
b look this little Guy ', 'RT SafeShibaMoon SafeShibaMoon Sol A
decentralized deflationary solana based doge meme token with inbuilt charity
support With S ', 'Very Classic Dog Rope Toy Hot selling more than years
Highly recommend why not dog toy pet products pet s t co McZ Y QJPL',
'Hey Hudson the August Dog Star of the Month wants to know if you have
entered the November Dog of the Month cont t co xcfQXdwnz']
['cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat',
'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat',
'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat',
'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat',

```
'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog']
```

```
C:\Users\jerem\anaconda3\lib\site-
packages\sklearn\feature_extraction\text.py:383: UserWarning: Your stop_words
may be inconsistent with your preprocessing. Tokenizing the stop words generated
tokens ['ha', 'le', 'u', 'wa'] not in stop_words.
  warnings.warn('Your stop_words may be inconsistent with '
```

3.1.3 Reformating and Tidying the Data

We convert the DTM to a Dataframe and add labels!

```
[9]: ## First - use your CountVectorizer to get all the column names
ColNames=MyCV_T.get_feature_names()
print("The vocab is: ", ColNames, "\n\n")

## NEXT - Use pandas to create data frames
My_DF_T=pd.DataFrame(My_DTM_T.toarray(),columns=ColNames)
```

```
The vocab is: ['- ', 'adoptable', 'adoption', 'adoptmepimaaz', 'akamaru',
'akamaruinu', 'amp', 'art', 'asap', 'available', 'b', 'based', 'beua', 'black',
'botanic', 'cat', 'charity', 'chiang', 'collection', 'cute', 'darkstalkers',
'decentralized', 'deflationary', 'dog', 'doge', 'drop', 'dumped', 'enter',
'felicia', 'friend', 'gallery', 'garden', 'gretel', 'hansel', 'hashtag',
'hashtags', 'help', 'home', 'id', 'inbuilt', 'j', 'kitten', 'kitty', 'know',
'life', 'love', 'lunch', 'm', 'magical', 'mai', 'manytle', 'meme', 'metal',
'mo', 'month', 'need', 'new', 'november', 'ow', 'owner', 'painting', 'people',
'pet', 'pimaaz', 'queen', 'realakamaruinu', 'rescue', 'rt', 's', 'sachikoko',
'safeshibamoon', 'sale', 'save', 'siamese', 'sirikit', 'sol', 'solana', 'st',
'support', 'sweet', 't', 'tabby', 'tag', 'tahoe', 'talk', 'thailand', 'token',
'tour', 'travel', 'tungmeng', 'usa', 'using', 'vintage', 'w', 'white', 'wi',
'x', 'y', 'year', 'yo']
```

3.1.4 Inspect and further cleaning

Twitter data is notoriously messy. So lets inspect and further clean. Lets remove numeric digits, links and more. Use Regular Expressions to identify word classes to be removed (EG Links) .

```
[10]: ## Let's look!
print(My_DF_T)
print(LABEL_List)
```

	-	adoptable	adoption	adoptmepimaaz	akamaru	akamaruinu	amp	art	\
0	0	0	0	0	0	0	1	0	
1	0	0	0	0	0	0	0	0	
2	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	0	0	
..	
95	0	1	0	1	0	0	0	0	
96	0	0	0	0	0	0	0	0	
97	0	0	0	0	0	0	0	0	
98	0	0	0	0	0	0	0	0	
99	0	0	0	0	0	0	0	0	

	asap	available	...	usa	using	vintage	w	white	wi	x	y	year	yo
0	0	0	...	0	0	0	0	1	0	0	0	0	1
1	0	0	...	0	0	0	0	0	0	0	0	0	0
2	0	1	...	0	0	0	0	0	0	0	0	0	0
3	0	0	...	0	0	0	0	0	0	0	0	0	0
4	0	0	...	0	0	0	0	0	0	0	0	0	0
..
95	0	0	...	0	0	0	1	0	0	0	0	0	0
96	0	0	...	0	0	0	0	0	0	0	0	0	0
97	0	0	...	0	0	0	0	0	0	0	0	0	0
98	0	0	...	0	0	0	0	0	0	0	1	1	0
99	0	0	...	0	0	0	0	0	0	0	0	0	0

[100 rows x 100 columns]

```
['cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat',
'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat',
'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat',
'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'cat',
'cat', 'cat', 'cat', 'cat', 'cat', 'cat', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog', 'dog',
'dog']
```

```
[11]: ## Some clean-up
droplist2 = [i for i in My_DF_T.columns if re.search(r'[^A-Za-z\~]', i)]
print(droplist2)
```

```
[]
```

```
[12]: My_DF_T.drop(droplist2,axis=1,inplace=True)
print(My_DF_T)
```

	-	adoptable	adoption	adoptmepimaaz	akamaru	akamaruinu	amp	art	\
0	0	0	0	0	0	0	1	0	

1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0
..
95	0	1	0	1	0	0	0	0	0
96	0	0	0	0	0	0	0	0	0
97	0	0	0	0	0	0	0	0	0
98	0	0	0	0	0	0	0	0	0
99	0	0	0	0	0	0	0	0	0

	asap	available	...	usa	using	vintage	w	white	wi	x	y	year	yo
0	0	0	...	0	0	0	0	1	0	0	0	0	1
1	0	0	...	0	0	0	0	0	0	0	0	0	0
2	0	1	...	0	0	0	0	0	0	0	0	0	0
3	0	0	...	0	0	0	0	0	0	0	0	0	0
4	0	0	...	0	0	0	0	0	0	0	0	0	0
..
95	0	0	...	0	0	0	1	0	0	0	0	0	0
96	0	0	...	0	0	0	0	0	0	0	0	0	0
97	0	0	...	0	0	0	0	0	0	0	0	0	0
98	0	0	...	0	0	0	0	0	0	1	1	0	0
99	0	0	...	0	0	0	0	0	0	0	0	0	0

[100 rows x 100 columns]

```
[13]: ## Note: The \bWORD\b matches the exact WORD
droplist3 = [i for i in My_DF_T.columns
              if re.search(r'\bhttp\b | \bhttps\b | \brt\b', i)]
print(droplist3)

My_DF_T.drop(droplist3,axis=1,inplace=True)

print(My_DF_T)
```

	-	adoptable	adoption	adoptmepimaaz	akamaru	akamaruinu	amp	art	\
0	0	0	0	0	0	0	1	0	
1	0	0	0	0	0	0	0	0	
2	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	0	0	
..
95	0	1	0	1	0	0	0	0	
96	0	0	0	0	0	0	0	0	
97	0	0	0	0	0	0	0	0	
98	0	0	0	0	0	0	0	0	

```

99  0          0          0          0          0          0          0          0          0
      asap  available  ...  usa  using  vintage  w  white  wi  x  y  year  yo
0      0          0  ...  0      0          0  0      1  0  0  0      0  1
1      0          0  ...  0      0          0  0      0  0  0  0      0  0
2      0          1  ...  0      0          0  0      0  0  0  0      0  0
3      0          0  ...  0      0          0  0      0  0  0  0      0  0
4      0          0  ...  0      0          0  0      0  0  0  0      0  0
..  ...          ...  ...  ...          ...  ...  ...  ...  ...  ...  ...
95     0          0  ...  0      0          0  1      0  0  0  0      0  0
96     0          0  ...  0      0          0  0      0  0  0  0      0  0
97     0          0  ...  0      0          0  0      0  0  0  0      0  0
98     0          0  ...  0      0          0  0      0  0  1  1      1  0
99     0          0  ...  0      0          0  0      0  0  0  0      0  0

```

[100 rows x 100 columns]

3.2 Classification: Simple Example with MNB

```
[14]: My_DF_T['label'] = LABEL_List
```

```

class1_DF = My_DF_T[(My_DF_T["label"] == label1)]
class2_DF = My_DF_T[(My_DF_T["label"] == label2)]

```

```
[ ]:
```

```

[15]: ## Create the testing set - grab a sample from the training set.
      ## Be careful. Notice that right now, our train set is sorted by label.
      ## If your train set is large enough, you can take a random sample.
      from sklearn.model_selection import train_test_split

      class1_train, class1_test = train_test_split(class1_DF, test_size=0.3)
      class2_train, class2_test = train_test_split(class2_DF, test_size=0.3)

      TrainDF = pd.concat([class1_train, class2_train])
      TestDF = pd.concat([class1_test, class2_test])

      trainLabels = TrainDF['label']
      testLabels = TestDF['label']

      TrainDF.drop(labels='label',axis=1,inplace=True)
      TestDF.drop(labels='label',axis=1,inplace=True)

      print(TrainDF)

```

```
- adoptable adoption adoptmepimaaz akamaru akamaruinu amp art \
```

18	0		0		0		0		0	0	0	
7	0		0		0		0		0	0	0	
43	0		1		0		1		0	0	0	
27	0		0		0		0		0	0	0	
42	0		1		0		1		0	0	0	
..	
66	0		0		0		0		1	1	0	0
95	0		1		0		1		0	0	0	0
77	0		0		0		0		1	1	0	0
58	0		0		0		0		0	0	0	0
79	0		0		0		0		1	1	0	0

	asap	available	...	usa	using	vintage	w	white	wi	x	y	year	yo
18	0		1	...	0	0	0	0	0	0	0	0	0
7	0		0	...	0	0	0	0	0	0	0	0	0
43	0		0	...	0	0	0	0	0	0	0	0	0
27	0		0	...	0	0	0	0	0	0	1	0	0
42	0		0	...	0	0	0	0	0	0	0	0	0
..
66	0		0	...	0	1	0	0	0	1	0	0	0
95	0		0	...	0	0	0	1	0	0	0	0	0
77	0		0	...	0	1	0	0	0	1	0	0	0
58	0		0	...	0	0	0	0	0	0	0	0	0
79	0		0	...	0	1	0	0	0	1	0	0	0

[70 rows x 100 columns]

```
[16]: #####
##### Predict Using Naive Bayes #####
#####
from sklearn.naive_bayes import MultinomialNB
#https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.
  ↳MultinomialNB.html#sklearn.naive_bayes.MultinomialNB.fit

MyModelNB= MultinomialNB()
MyModelNB.fit(TrainDF, trainLabels)
Prediction = MyModelNB.predict(TestDF)
print("The prediction from NB is:")
print(Prediction)
print("The actual labels are:")
print(testLabels)
```

The prediction from NB is:

```
['cat' 'cat' 'cat' 'cat' 'cat' 'cat' 'cat' 'cat' 'cat' 'cat' 'cat' 'cat' 'cat'
'cat' 'cat' 'cat' 'dog' 'dog' 'dog' 'dog' 'cat' 'dog' 'dog' 'dog' 'dog'
'cat' 'dog' 'dog' 'dog' 'dog' 'dog']
```

The actual labels are:


```

32    cat
16    cat
33    cat
0     cat
17    cat
37    cat
22    cat
11    cat
20    cat
5     cat
39    cat
24    cat
47    cat
10    cat
1     cat
88    dog
69    dog
50    dog
60    dog
53    dog
97    dog
75    dog
98    dog
56    dog
57    dog
52    dog
65    dog
61    dog
94    dog
70    dog
Name: label, dtype: object

```

```
[ ]:
```

```

[17]: ## How Accurate was the Model ... confusion matrix
      from sklearn.metrics import confusion_matrix
      ## The confusion matrix is square and is labels X labels
      ## We have two labels, so ours will be 2X2
      #The matrix shows
      ## rows are the true labels
      ## columns are predicted
      ## it is alphabetical
      ## The numbers are how many
      ## actual = (TestLabels == 'neg').tolist()
      ## predict = (Prediction == 'neg').tolist()
      y_true = testLabels
      y_predict = (Prediction).tolist()

```

```

labels = [label1, label2]
cm = confusion_matrix(y_true, y_predict, labels)
print(cm)
### prediction probabilities
## columns are the labels in alphabetical order
## The decimal in the matrix are the prob of being
## that label
#print(np.round(MyModelNB.predict_proba(TestDF),2))
#import matplotlib as plt
import matplotlib.pyplot as plt
fig = plt.figure()
ax = fig.add_subplot(111)
cax = ax.matshow(cm)
plt.title('Confusion matrix of the classifier')
fig.colorbar(cax)

ax.set_xticklabels([''] + labels)
ax.set_yticklabels([''] + labels)

plt.xlabel('Predicted')
plt.ylabel('True')
plt.show()

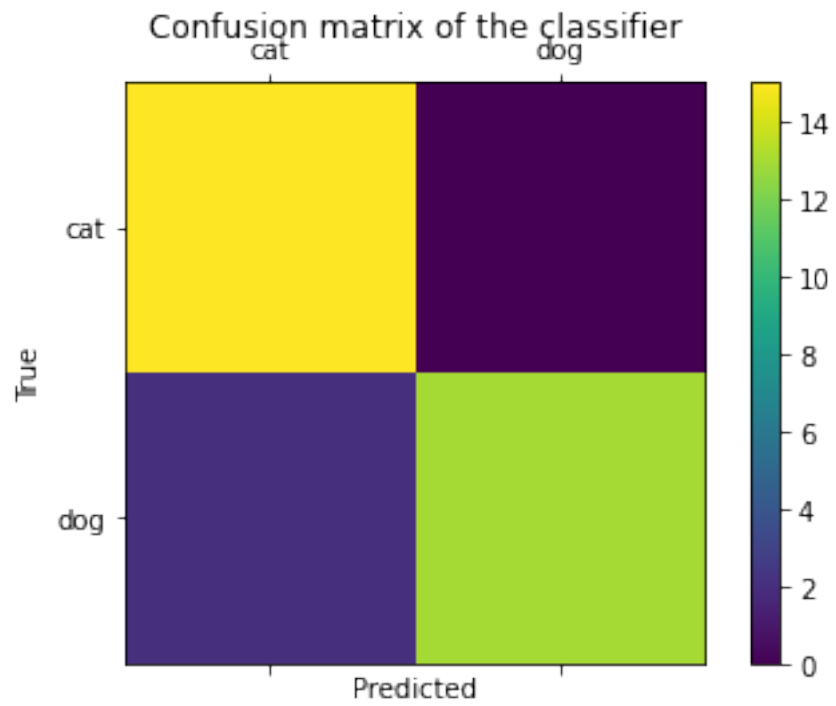
```

C:\Users\jerem\anaconda3\lib\site-packages\sklearn\utils\validation.py:68:
FutureWarning: Pass labels=['cat', 'dog'] as keyword args. From version 0.25
passing these as positional arguments will result in an error
warnings.warn("Pass {} as keyword args. From version 0.25 "

```

[[15  0]
 [ 2 13]]

```



[]:

[]: