----------Module 1(Data Collection)----------

Collection data from twitter for legitimate users

For collecting the data from twitter I am using Tweepy module

For that We need Counsumer_KEY, Counsumer_secret_KEY, Access_token, Access_Token

That all I can get from twitter app.devloper where we need to sign in and make an account and then twitter generate the keys

After that a simple program in python can extract data from twitter in given limit by twitter.

In [1]:
```
import tweepy
import pandas as pd
import time
import numpy as np
import matplotlib.pyplot as plt
from tweepy import Stream
from tweepy.streaming import StreamListener
import numpy as np
import matplotlib.pyplot as plt
from sklearn import svm
from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix
from sklearn.utils.multiclass import unique_labels
```

Connection Authentication

In [3]:
```
consumer_key  = 'd9Ksoz6Wb1jD0mqbW8rjaSNb7'
consumer_secret = 'pHXnVSJeLbOxaYlbOR7BWFdDNhZSF6IzegZV87qUSUqy6Qe8qG'
access_token = '3648603434-dGRu1nHet22tdoYeqaAGoN8MyZrNw9oXZQvGZUD'
access_token_secret = 'PZ8pcQBCb5zVPLRQNVQZc3Yzi0rz1wPef6O7RO7gzcvOf'

auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)

api = tweepy.API(auth, wait_on_rate_limit=True)
print(api)
```

```
<tweepy.api.API object at 0x000002050123E9A0>
```

Collecting Data list of Friends of a given screen_name

Save data in txt file

In [3]:
```
'''print('Name of the Friends of user')
friends = []
sn=['PoliceRajasthan','Uppolice','MumbaiPolice','PunjabPoliceInd','KolkataPolice','DelhiPolice','BlrCityPolice','noidapolice'
print(len(sn))
for name in sn:
    for friend in tweepy.Cursor(api.friends, screen_name = name).items(20):
        try:
            friends.append(friend.screen_name)
            print(friend.screen_name)
            time.sleep()
        except Exception as e:
                        pass

print("Writing to file Luser...")
with open("C:/Users/asha/Desktop/Final Year Project Phase 2/Phase 2/Dataset/luser.txt", "a") as f:
     for item in friends:
        f.write("%s\n" % item)
print("Writed Succesfully completed...")


#output
Name of the Friends of user
19
shomarredpally
ShoMarket
shonarayanguda
shomahankali
shokulsumpura
shokarkhana
shokanchanbagh
shocharminar
shochaderghat
shochilkalguda
shogopalapuram
shobegumpet
shoafzalgunj
shoamberpet
Prabhamudhiraj
rambanoth
shocgt
bhanu02468
shobowenpally
KTRTRS
narendramodi
WomenCid
Writing to file Luser...
Writed Succesfully completed...'''
```

Out[3]:
```
'print(\'Name of the Friends of user\')\nfriends = []\nsn=[\'PoliceRajasthan\',\'Uppolice\',\'MumbaiPolice\',\'PunjabPoliceI
nd\',\'KolkataPolice\',\'DelhiPolice\',\'BlrCityPolice\',\'noidapolice\',\'igrangemeerut\',\'noidatraffic\',\'adgzonemeerut
\',\'meerutpolice\',\'bulandshahrpol\',\'saharanpurpol\',\'shamlipolice\',\'hapurpolice\',\'baghpatpolice\',\'chennaipolice_
\',\'hydcitypolice\']\nprint(len(sn))\nfor name in sn:\n    for friend in tweepy.Cursor(api.friends, screen_name = name).ite
ms(20):\n        try:\n            friends.append(friend.screen_name)\n            print(friend.screen_name)\n            ti
me.sleep()\n        except Exception as e:\n                        pass\n\nprint("Writing to file Luser...")\nwith open
("C:/Users/asha/Desktop/Final Year Project Phase 2/Phase 2/Dataset/luser.txt", "a") as f:\n     for item in friends:\n
f.write("%s\n" % item)\nprint("Writed Succesfully completed...")\n\n\n#output\nName of the Friends of user\n19\nshomarredpal
ly\nShoMarket\nshonarayanguda\nshomahankali\nshokulsumpura\nshokarkhana\nshokanchanbagh\nshocharminar\nshochaderghat\nshochi
lkalguda\nshogopalapuram\nshobegumpet\nshoafzalgunj\nshoamberpet\nPrabhamudhiraj\nrambanoth\nshocgt\nbhanu02468\nshobowenpal
ly\nKTRTRS\nnarendramodi\nsmittal_ips\nshahalibanda\nshomusheerabad\nHCSC_Hyd\nvishwa_raghu\nChowraastaM\nNGopishetty\nravit
ollywood\nKhammamCp\nPuneCityPolice\nSmitaSabharwal\npassportsevamea\nOfficeOfVKS\nTelanganaHealth\nTsspdclCorporat\nsho_mog
halpura\nCPMumbaiPolice\nOnlineBlooDonor\nspmahabubabad\nmhbdpolice\nshomadannapet\nPIBFactCheck\nghmc_adw\nTSEduDept\nMAHal
eem20\nspvikarabad\nCYBTRAFFIC\nipstelangana\nTelanganaCOPs\nMORTHRoadSafety\nTSPSCofficial\nTSMAUDOnline\nTSCSOffice\nTSCon
sumers\nBlrCityPolice\nSpKothagudem\nlrvr1974\nMLA54327644\nWomenCid\nWriting to file Luser...\nWrited Succesfully complete
d...'
```

Now collect 30 tweet from each user in luser file that I extracted from twitter

In [ ]:
```
Total_Data = []
fo = open("C:/Users/asha/Desktop/Final Year Project Phase 2/Phase 2/Dataset/luser.txt", "r")
f = fo.readlines()
fo.close()
dataset = map(lambda s: s.strip(),f)
try:
    for datavar in dataset:
        data = api.get_user(datavar)
        counter = 0
        for status in tweepy.Cursor(api.user_timeline, id = datavar).items(30):
            try:
                counter= counter+1
                Total_Data.append(status)
                time.sleep()
            except Exception as e:
                pass
except Exception as e:
    pass
print(len(Total_Data))
```

Now from each tweet extract useful atributes that I need later for classification

In [5]:
```
import urllib.parse
import pandas as pd

def process_http(string):
    url_count = 0
    for i in string.split():
        s, n, p, pa, q, f = urllib.parse.urlparse(i)
        if s and n:
            url_count += 1
    return url_count

def process_hashtag(string):
    hashtag_count = 0
    for i in string.split():
        s, n, p, pa, q, f = urllib.parse.urlparse(i)
        if i[:1] == '#':
            hashtag_count += 1
    return hashtag_count

def process_mention(string):
    mention_count=0
    for i in string.split():
        s, n, p, pa, q, f = urllib.parse.urlparse(i)
        if i[:1] == '@':
            mention_count += 1
    return mention_count

def process_data(Total_Data):
    TwittID = [tweet.id for tweet in Total_Data]
    # Making the dataset in pandas frame
    Data = pd.DataFrame(TwittID, columns = ['TwittID'])
    # processing the data in Tweet level

    Data["TextData"] = [tweet.text for tweet in Total_Data]
    Data["TweetCreatedAt"] = [tweet.created_at for tweet in Total_Data]
    Data["RetweetCount"] = [tweet.retweet_count for tweet in Total_Data]
    Data["TweetFavouriteCount"] = [tweet.favorite_count for tweet in Total_Data]
    Data["TweetSource"] = [tweet.source for tweet in Total_Data]

    # processing the data in User Graph level

    Data["UserID"] = [tweet.author.id for tweet in Total_Data]
    Data["UserScreenName"] = [tweet.author.screen_name for tweet in Total_Data]
    Data["UserName"] = [tweet.author.name for tweet in Total_Data]
    Data["UserCreatedAt"] = [tweet.author.created_at for tweet in Total_Data]
    Data["UserDescription"] = [tweet.author.description for tweet in Total_Data]
    Data["UserDescriptionLength"] = [len(tweet.author.description) for tweet in Total_Data]
    Data["UserFollowersCount"] = [tweet.author.followers_count for tweet in Total_Data]
    Data["UserFriendsCount"] = [tweet.author.friends_count for tweet in Total_Data]
    Data["UserLocation"] = [tweet.author.location for tweet in Total_Data]

    # Data["url"] = [tweet.author.url for in Total_Data]
    # Data["User_mention"] = [user_mentions.author.screen_name for tweet in Total_Data]
    # Data["HashTag"] = [hashtag.text for tweet in Total_Data]

    Data["HttpCount"] = [process_http(tweet.text) for tweet in Total_Data]
    Data["HashtagCount"] = [process_hashtag(tweet.text) for tweet in Total_Data]
    Data["MentionCount"] = [process_mention(tweet.text) for tweet in Total_Data]
    Data["TweetCount"] = [tweet.author.statuses_count for tweet in Total_Data]
    return Data
Data = process_data(Total_Data)
Data.shape
```

Out[5]: (330, 19)

In [6]:
```
Data.tail(4)
```

Out[6]:

	TwittID	TextData	TweetCreatedAt	RetweetCount	TweetFavouriteCount	TweetSource	UserID	UserScreenName	UserName	User
326	1191764131088437248	RT @shailylPSspeaks: We have seen this. We hav...	2019-11-05 17:08:05	13	0	Twitter for Android	371293553	pihuprasad	IPS YAMUNA	2
327	1191763983344140288	RT @jatinnarwal1: Its not about which professi...	2019-11-05 17:07:30	896	0	Twitter for Android	371293553	pihuprasad	IPS YAMUNA	2
328	1191763742935019522	RT @Hii_VT: Let's learn how to burn policemen ...	2019-11-05 17:06:32	12	0	Twitter for Android	371293553	pihuprasad	IPS YAMUNA	2
329	1191763250133659648	RT @IPS_KTK_Assn: We strongly condemn incident...	2019-11-05 17:04:35	31	0	Twitter for Android	371293553	pihuprasad	IPS YAMUNA	2

Save data in csv_files in Leg_user.csv

In [7]:
```
import sys
# Saving data with item space separating
Data.to_csv('C:/Users/asha/Desktop/Final Year Project Phase 2/Phase 2/Dataset/Leg_user.csv', sep=',' , encoding='utf8')
```

extracting Spam data from twitter by searching @spam and find out the user for reported that hypothesis is that there is highly chances is that that user be fake

We can later analyse by text any volgor word and find our later first like legitimate user we have to collect data for spammer too

In [12]:
```
"""# printing all the friends names of the user

friends = []
class listener(StreamListener):
    def on_data(self, data):
        try:
            tweet = data.split(',"screen_name":"')[1].split('","location')[0]
            print(tweet)
            friends.append(tweet)
            return True
        except BaseException as e:
            print('failed on data' + str(e))
            time.sleep(5)
    def on_error(self, status):
        print(status)

twitterStream = Stream(auth, listener())
try:
    for x in range(1,10):
        twitterStream.filter(track=["cougar"])
except KeyboardInterrupt:
    print("Key board interuption")
with open("C:/Users/asha/Desktop/Final Year Project Phase 2/Phase 2/Dataset/suser.txt", "w") as f:
     for item in friends:
        f.write("%s\n" % item)

output:
josef_morasch
hyenafluff
MattLov14455762
HerFabulousWays
ShatteredReaper
HornyScotsman98
kfkffmdm
kfkffmdm
MdmMcCoy
bradley33885052
pollo0014
tribink1
MattLov14455762
kyte_dave
uDerThehOod
Stefff32720432
kfkffmdm
eli_boggs
_vaeXvae
reno7817
reno7817
Emir20665707
reno7817
pollo0014
carloshdez21
Sara19275758
CTG_tsutsumi
pollo0014
JDublim
ctg_machida
Roberto441220
OkomeeRice
profbadoor
pollo0014
SlaveTraderJoes
Amstron66093131
Masterxii1_4
Sexypro30863666
Onlyfan13594367
1st1966
huggyair
RosePro47247034
Stefff32720432
pollo0014
cougarmilfclub
JDublim
masterbooblover
Alex1little
Dutch_DPP
urbandictionary
nicolas64091667
luiginked
nicolas64091667
mommatiff3
mommatiff3
streitzfrank
pollo0014
beatsbygibbs
noel6132
lionloins
LisaCochonne
pollo0014
JOHNMcILHONE2
BLOODSADx
PauldingRadio
```

justasksam3
OGKING04987433
winekunTM
LuisEnri4400004
WhenSus8
alexand19024887
pollo0014
BbcRon_
PipPromos
bassshake
Para_Dox207
winekunTM
LopedeVegan269
piranhajams
tytheestallio
KevinBr96780686
CougargamingE
flvwrboy
LouisePixie1
alexabrines
pollo0014
Markanthony315
tytheestallio
cougar_an
naughtysussex
Christi07528537
Gsanchez7821
pollo0014
doodlebug247
ps236236
ps236236
hornynigga1521
saveurdesiles28
1041zion
oozytube
geezporn
saveurdesiles28
joohwangblink
EDG_Group
browniej45
saveurdesiles28
joohwangblink
browniej45
browniej45
browniej45
browniej45
browniej45
browniej45
saveurdesiles28
browniej45
browniej45
browniej45
RootNationUA
browniej45
browniej45
browniej45
justasksam3
browniej45
browniej45
browniej45
browniej45
browniej45
browniej45
browniej45
browniej45
browniej45
browniej45
browniej45
browniej45
browniej45
CurvesDelicious
browniej45
browniej45
browniej45
pollo0014
TrylineUK
BaidinNoer
boobhunters1
browniej45
33Metalson
AzurisDraws
pollo0014
GregMooney2
justasksam3
mllgigi
Nasser61597438
Itsimim
SamJenk60032204
pollo0014
Hah52059025
BbcRon_
thibodeaux_nora
PPPvideos1

scully29jamie
Hah52059025
turtletesticles
pollo0014
ninacelibataire
adoptunecougar
opium65
pornlegendsclub
cougarmilfclub
Tightrope_Postm
Baddies74810961
tazats1
TigerShoutout
SH63927626
1041zion
sexxebuddah69
pollo0014
slol59
InsaMarquez
oozytube
geezporn
truckpornxxx
Ben00182369
olddrive69
Paradise000x
himansu92239715
cmu_Eleasar
mjohns2444
saveurdesiles28
wvmilfhunter
pollo0014
mjohns2444
mjohns2444
oozytube
geezporn
Daniel_S013
KinkyCuriousHo1
Daniel_S013
PervertDiaries
WifeisBitchPorn
oozytube
geezporn
pollo0014
DOpararaeKWS
Aaronnowak14
oozytube
geezporn
nate221217
Recaster20
oozytube
geezporn
pollo0014
fast_girlswild
SilverSeductres
donne_mature
BbcRon_
TheGILFNextDoor
james68402115
ktdenise
TheFallenLady_
ovelhaJorge
pollo0014
Qazi04968557
WiccaNymph
mac2002
Socrate90341168
johnlamanna3
mpwCfbFldSQOZTA
ZGLiraxD16
WAVMedia
Hornybo68150984
pollo0014
IrisScarpinelli
Infr3quentAddik
Aaronnowak14
unuucla
Srinu50041745
Srinu50041745
louisejacksonuk
Andy_Lifeguard
Ym78200
meghan_aleo
Asuro_Fluid
Uhmmm26387032
Asuro_Fluid
acupof_TAY
ballc21
cougarmilfclub
pollo0014
sindra38503651
stevenhead38
blindromance_
Canijo74474417

```
pollo0014
rdvlibertins
pollo0014
HErottico
JamesALogan1
BenoitJoris
Gio58652487
18Beastman
JacobStockGuru
BriW74
kamdiv49
Emir20665707
Emir20665707
Belzouille1
AndeeSC2
KPromention
PassionCougar
promotion_feet
SexyPro18754447
dreinschisspur
hellsivle
pollo0014
carolbaker10023
jay818jay
Alendil_fr
ESPINARjos2
StadiumPrint
cougar_an
MurphyDuke_
MurphyDuke_
MurphyDuke_
funinlillian
jav_superfan"""
```

Out[12]:

```
'# printing all the friends names of the user \n\nfriends = []\nclass listener(StreamListener):\n    def on_data(self, dat
a):\n        try:\n            tweet = data.split(\',"screen_name":"\')[1].split(\'","location\')[0]\n            print(twee
t)\n            friends.append(tweet)\n            return True\n        except BaseException as e:\n            print(\'fail
ed on data\' + str(e))\n            time.sleep(5)\n    def on_error(self, status):\n        print(status)\n\ntwitterStream =
Stream(auth, listener())\ntry:\n    for x in range(1,10):\n        twitterStream.filter(track=["cougar"])\nexcept KeyboardIn
terrupt:\n    print("Key board interuption")\nwith open("C:/Users/asha/Desktop/Final Year Project Phase 2/Phase 2/Dataset/su
ser.txt", "w") as f:\n     for item in friends:\n        f.write("%s\n" % item)\n        \noutput:\njosef_morasch\nhyenafluf
f\nMattLov14455762\nHerFabulousWays\nShatteredReaper\nHornyScotsman98\nkfkffmdm\nkfkffmdm\nMdmMcCoy\nbradley33885052\npollo0
014\ntribink1\nMattLov14455762\nkyte_dave\nuDerThehOod\nSteffff32720432\nkfkffmdm\neli_boggs\n_vaeXvae\nreno7817\nreno7817\n
Emir20665707\nreno7817\npollo0014\ncarloshdez21\nSara19275758\nCTG_tsutsumi\npollo0014\nJDublim\nctg_machida\nRoberto441220
\nOkomeeRice\nprofbadoor\npollo0014\nSlaveTraderJoes\nAmstron66093131\nMasterxii1_4\nSexypro30863666\nOnlyfan13594367\n1st19
66\nhuggyair\nRosePro47247034\nSteffff32720432\npollo0014\ncougarmilfclub\nJDublim\nmasterbooblover\nAlex1little\nDutch_DPP
\nurbandictionary\nnicolas64091667\nluiginked\nnicolas64091667\nmommatiff3\nmommatiff3\nstreitzfrank\npollo0014\nbeatsbygibb
s\nnoel6132\nlionloins\nLisaCochonne\npollo0014\nJOHNMcILHONE2\nBLOODSADx\nPauldingRadio\njustasksam3\nOGKING04987433\nwinek
unTM\nLuisEnri4400004\nWhenSus8\nalexand19024887\npollo0014\nBbcRon_\nPipPromos\nbassshake\nPara_Dox207\nwinekunTM\nLopedeVe
gan269\npiranhajams\ntytheestallio\nKevinBr96780686\nCougargamingE\nflvwrboy\nLouisePixie1\nalexabrines\npollo0014\nMarkanth
ony315\ntytheestallio\ncougar_an\nnaughtysussex\nChristi07528537\nGsanchez7821\npollo0014\ndoodlebug247\nps236236\nps236236
\nhornynigga1521\nsaveurdesiles28\n1041zion\noozytube\ngeezporn\nsaveurdesiles28\njoohwangblink\nEDG_Group\nbrowniej45\nsave
urdesiles28\njoohwangblink\nbrowniej45\nbrowniej45\nbrowniej45\nbrowniej45\nbrowniej45\nbrowniej45\nsaveurdesiles28\nbrownie
j45\nbrowniej45\nbrowniej45\nRootNationUA\nbrowniej45\nbrowniej45\nbrowniej45\njustasksam3\nbrowniej45\nbrowniej45\nbrowniej
45\nbrowniej45\nbrowniej45\nbrowniej45\nbrowniej45\nbrowniej45\nbrowniej45\nbrowniej45\nbrowniej45\nbrowniej45\nbrowniej45\n
CurvesDelicious\nbrowniej45\nbrowniej45\nbrowniej45\npollo0014\nTrylineUK\nBaidinNoer\nboobhunters1\nbrowniej45\n33Metalson
\nAzurisDraws\npollo0014\nGregMooney2\njustasksam3\nmllgigi\nNasser61597438\nItsimim\nSamJenk60032204\npollo0014\nHah5205902
5\nBbcRon_\nthibodeaux_nora\nPPPvideos1\nscully29jamie\nHah52059025\nturtletesticles\npollo0014\nninacelibataire\nadoptuneco
ugar\nopium65\npornlegendsclub\ncougarmilfclub\nTightrope_Postm\nBaddies74810961\ntazats1\nTigerShoutout\nSH63927626\n1041zi
on\nsexxebuddah69\npollo0014\nslol59\nInsaMarquez\noozytube\ngeezporn\ntruckpornxxx\nBen00182369\nolddrive69\nParadise000x\n
himansu92239715\ncmu_Eleasar\nmjohns2444\nsaveurdesiles28\nwvmilfhunter\npollo0014\nmjohns2444\nmjohns2444\noozytube\ngeezpo
rn\nDaniel_S013\nKinkyCuriousHo1\nDaniel_S013\nPervertDiaries\nWifeisBitchPorn\noozytube\ngeezporn\npollo0014\nDOpararaeKWS
\nAaronnowak14\noozytube\ngeezporn\nnate221217\nRecaster20\noozytube\ngeezporn\npollo0014\nfast_girlswild\nSilverSeductres\n
donne_mature\nBbcRon_\nTheGILFNextDoor\njames68402115\nktdenise\nTheFallenLady_\novelhaJorge\npollo0014\nQazi04968557\nWicca
Nymph\nmac2002\nSocrate90341168\njohnlamanna3\nmpwCfbFldSQOZTA\nZGLiraxD16\nWAVMedia\nHornybo68150984\npollo0014\nIrisScarpi
nelli\nInfr3quentAddik\nAaronnowak14\nunuucla\nSrinu50041745\nSrinu50041745\nlouisejacksonuk\nAndy_Lifeguard\nYm78200\nmegha
n_aleo\nAsuro_Fluid\nUhmmm26387032\nAsuro_Fluid\nacupof_TAY\nballc21\ncougarmilfclub\npollo0014\nsindra38503651\nstevenhead3
8\nblindromance_\nCanijo74474417\npollo0014\nrdvlibertins\npollo0014\nHErottico\nJamesALogan1\nBenoitJoris\nGio58652487\n18B
eastman\nJacobStockGuru\nBriW74\nkamdiv49\nEmir20665707\nEmir20665707\nBelzouille1\nAndeeSC2\nKPromention\nPassionCougar\npr
omotion_feet\nSexyPro18754447\ndreinschisspur\nhellsivle\npollo0014\ncarolbaker10023\njay818jay\nAlendil_fr\nESPINARjos2\nSt
adiumPrint\ncougar_an\nMurphyDuke_\nMurphyDuke_\nMurphyDuke_\nfuninlillian\njav_superfan'
```

In [4]:
```
Total_Data = []
fo = open("C:/Users/asha/Desktop/Final Year Project Phase 2/Phase 2/Dataset/suser.txt", "r")
f = fo.readlines()
fo.close()
dataset = map(lambda s: s.strip(),f)
try:
    for datavar in dataset:
        data = api.get_user(datavar)
        counter = 0
        for status in tweepy.Cursor(api.user_timeline, id = datavar).items(10):
            try:
                counter= counter+1
                Total_Data.append(status)
                print(Total_data)
                time.sleep()
            except Exception as e:
                pass
except Exception as e:
    pass
print(len(Total_Data))
```

```
353
```

Now from tweet extract useful atributes

In [20]:
```
import urllib.parse
import pandas as pd

def process_http(string):
    url_count = 0
    for i in string.split():
        s, n, p, pa, q, f = urllib.parse.urlparse(i)
        if s and n:
            url_count += 1
    return url_count

def process_hashtag(string):
    hashtag_count = 0
    for i in string.split():
        s, n, p, pa, q, f = urllib.parse.urlparse(i)
        if i[:1] == '#':
            hashtag_count += 1
    return hashtag_count

def process_mention(string):
    mention_count=0
    for i in string.split():
        s, n, p, pa, q, f = urllib.parse.urlparse(i)
        if i[:1] == '@':
            mention_count += 1
    return mention_count

def process_data(Total_Data):
    TwittID = [tweet.id for tweet in Total_Data]
    # Making the dataset in pandas frame
    Data = pd.DataFrame(TwittID, columns = ['TwittID'])
    # processing the data in Tweet level

    Data["TextData"] = [tweet.text for tweet in Total_Data]
    Data["TweetCreatedAt"] = [tweet.created_at for tweet in Total_Data]
    Data["RetweetCount"] = [tweet.retweet_count for tweet in Total_Data]
    Data["TweetFavouriteCount"] = [tweet.favorite_count for tweet in Total_Data]
    Data["TweetSource"] = [tweet.source for tweet in Total_Data]

    # processing the data in User Graph level

    Data["UserID"] = [tweet.author.id for tweet in Total_Data]
    Data["UserScreenName"] = [tweet.author.screen_name for tweet in Total_Data]
    Data["UserName"] = [tweet.author.name for tweet in Total_Data]
    Data["UserCreatedAt"] = [tweet.author.created_at for tweet in Total_Data]
    Data["UserDescription"] = [tweet.author.description for tweet in Total_Data]
    Data["UserDescriptionLength"] = [len(tweet.author.description) for tweet in Total_Data]
    Data["UserFollowersCount"] = [tweet.author.followers_count for tweet in Total_Data]
    Data["UserFriendsCount"] = [tweet.author.friends_count for tweet in Total_Data]
    Data["UserLocation"] = [tweet.author.location for tweet in Total_Data]

    # Data["url"] = [tweet.author.url for in Total_Data]
    # Data["User_mention"] = [user_mentions.author.screen_name for tweet in Total_Data]
    # Data["HashTag"] = [hashtag.text for tweet in Total_Data]

    #Data["HttpCount"] = [process_http(tweet.text) for tweet in Total_Data]
    #Data["HashtagCount"] = [process_hashtag(tweet.text) for tweet in Total_Data]
    #Data["MentionCount"] = [process_mention(tweet.text) for tweet in Total_Data]
    #Data["TweetCount"] = [tweet.author.statuses_count for tweet in Total_Data]
    return Data
Data = process_data(Total_Data)
Data.head()
```

Out[20]:

	TwittID	TextData	TweetCreatedAt	RetweetCount	TweetFavouriteCount	TweetSource	UserID	UserScreenName	Use
0	1396764493443698695	Okay Babe https://t.co/8ft3uj3UWc	2021-05-24 09:46:20	0	0	Twitter Web App	1084687742032924674	josef_morasch	M
1	1396763990424948736	Yes Babe https://t.co/HtlTsLdbDo	2021-05-24 09:44:20	0	0	Twitter Web App	1084687742032924674	josef_morasch	M
2	1396763678498869251	RT @kscpl88_91: Hope everyone had a great #Sun...	2021-05-24 09:43:06	11	0	Twitter Web App	1084687742032924674	josef_morasch	M
3	1396763521283760129	Undress your panty then its perfect https://t....	2021-05-24 09:42:28	0	0	Twitter Web App	1084687742032924674	josef_morasch	M
4	1396763261480144898	You are to fare https://t.co/dYhRDiTuEa	2021-05-24 09:41:26	0	0	Twitter Web App	1084687742032924674	josef_morasch	M

To save spam user in suser.csv

In [23]:
```
# Saving data with item space separating
Data.to_csv("C:/Users/asha/Desktop/Final Year Project Phase 2/Phase 2/Dataset/Spam_user.csv", sep=',' , encoding='utf8')
```

In [ ]: