

Spark_Twitter_Analysis

December 19, 2021

Michael Woo & Swapnil Basu

Imports

```
[1]: import numpy as np
import pandas as pd
from pyspark.sql import SparkSession
from google_drive_downloader import GoogleDriveDownloader as gdd
from pyspark.ml.feature import RegexTokenizer, StopWordsRemover, CountVectorizer
from pyspark.ml.classification import LogisticRegression
import nltk
from nltk.corpus import stopwords
import string
from pyspark.ml import Pipeline
from pyspark.ml.feature import OneHotEncoder, StringIndexer, VectorAssembler
from pyspark.sql.functions import col
from pyspark.ml.evaluation import MulticlassClassificationEvaluator
```

Download File from link given in Canvas

This will be stored into your local...Do not add into git, file is too large to be pushed onto git master branch

So, we download locally

```
[2]: gdd.download_file_from_google_drive(file_id='OBO4GJPshIjmPRnZManQwWEdTZjg',
                                         dest_path='/Users/mwoo/Downloads/
                                         ↳trainingandtestdata.zip',
                                         unzip=True)

[3]: # gdd.download_file_from_google_drive(file_id='OBO4GJPshIjmPRnZManQwWEdTZjg',
#                                           dest_path='/Users/swapnilbasu/Downloads/
#                                           ↳trainingandtestdata.zip',
#                                           unzip=True)
```

Create spark session object (Data Processing)

```
[4]: spark=SparkSession.builder.appName('classification_tweet').getOrCreate()
```

Load in data

```
[5]: training_data = spark.read.csv("/Users/mwoo/Downloads/training.1600000.
    ↳processed.noemoticon.csv",header=False)
```

Renaming columns

```
[6]: training_data.columns
```

```
[6]: ['_c0', '_c1', '_c2', '_c3', '_c4', '_c5']
```

```
[7]: training_data = training_data.
    ↳toDF("target", 'id', 'date', 'query', 'user_name', 'text')
```

```
[8]: training_data.columns
```

```
[8]: ['target', 'id', 'date', 'query', 'user_name', 'text']
```

Exploratory

```
[9]: training_data.describe()
```

```
[9]: DataFrame[summary: string, target: string, id: string, date: string, query:
    string, user_name: string, text: string]
```

Selecting the target value and text

These are the values that we believe are important in predicting whether a tweet was negative or positive

```
[10]: df = training_data.select('text', 'target')
```

```
[11]: df.show(5)
```

```
+-----+-----+
|          text|target|
+-----+-----+
|@switchfoot http:...|    0|
|is upset that he ...|    0|
|@Kenichan I dived...|    0|
|my whole body fee...|    0|
|@nationwideclass ...|    0|
+-----+-----+
only showing top 5 rows
```

```
[12]: df.printSchema()
```

```
root
 |-- text: string (nullable = true)
```

```
|-- target: string (nullable = true)
```

We can see below that its an even split between positive and negative tweets

0: negative 4: positive

```
[13]: df.groupBy("target").count().orderBy(col("count").desc()).show()
```

```
+-----+-----+
|target| count|
+-----+-----+
|      0|800000|
|      4|800000|
+-----+-----+
```

Model Pipeline

Regular Expression Tokenizer

Seperates the texts into words

```
[14]: regexTokenizer = RegexTokenizer(inputCol="text", outputCol="words",
    ↪pattern="\\W")
```

Stop Words Download from NLTK

Downloads Stop words

```
[15]: nltk.download('stopwords')
```

```
[nltk_data] Downloading package stopwords to /Users/mwoo/nltk_data...
```

```
[nltk_data] Package stopwords is already up-to-date!
```

```
[15]: True
```

Stop Words Remover

Remove unnecessary words

```
[16]: sp = set(string.punctuation)
stop_words = set(stopwords.words('english'))
extra_words = {"http", "https", "amp", "rt", "t", "c", "the"}
for i in extra_words:
    stop_words.add(i)
stop_words = list(stop_words)
```

```
[17]: stopwordsRemover = StopWordsRemover(inputCol="words", outputCol="filtered").
    ↪setStopWords(stop_words)
```

Bag of words count

This is a type of feature engineering

```
[18]: countVectors = CountVectorizer(inputCol="filtered", outputCol="features",  
    ↪ vocabSize=10000, minDF=5)
```

StringIndexer

Indexing target values

```
[19]: label_stringIdx = StringIndexer(inputCol = "target", outputCol = "label")
```

```
[20]: pipeline = Pipeline(stages=[regexTokenizer, stopwordsRemover, countVectors,  
    ↪ label_stringIdx])  
pipelineFit = pipeline.fit(df)  
dataset = pipelineFit.transform(df)  
dataset.show(5)
```

```
+-----+-----+-----+-----+-----+  
-----+-----+  
|          text|target|          words|          filtered|  
features|label|  
+-----+-----+-----+-----+-----+  
-----+-----+  
|@switchfoot http:...|    0|[switchfoot, http...|[switchfoot,  
twit...|(10000,[1,10,16,6...| 0.0|  
|is upset that he ...|    0|[is, upset, that,...|[upset, update,  
f...|(10000,[6,70,172,...| 0.0|  
|@Kenichan I dived...|    0|[kenichan, i, div...|[kenichan,  
dived,...|(10000,[4,213,251...| 0.0|  
|my whole body fee...|    0|[my, whole, body,...|[whole, body,  
fee...|(10000,[3,325,374...| 0.0|  
|@nationwideclass ...|  
0|[nationwideclass,...|[nationwideclass,...|(10000,[20,486],[...| 0.0|  
+-----+-----+-----+-----+-----+  
-----+-----+  
only showing top 5 rows
```

Selecting data from the previous dataframe

```
[21]: dataset = dataset.select('text', 'features', 'label')
```

Set seed for reproducibility

This set is use for testing purposes 70/30 split

```
[22]: (trainingData, testData) = dataset.randomSplit([0.7, 0.3], seed = 100)  
print("Training Dataset Count: " + str(trainingData.count()))  
print("Test Dataset Count: " + str(testData.count()))
```

Training Dataset Count: 1120280

Test Dataset Count: 479720

This will be used to fully train our classification model

```
[23]: model_df = dataset.select('features','label')
      print("Full Training Dataset Count: " + str(model_df.count()))
```

Full Training Dataset Count: 1600000

Testing our model through the split data above

```
[24]: lr = LogisticRegression(maxIter=20, regParam=0.3, elasticNetParam=0).
      ↪fit(trainingData)
      predictions = lr.transform(testData)
      predictions.filter(predictions['prediction'] == 0) \
        .select("text","probability","label","prediction") \
        .orderBy("probability", ascending=False) \
        .show(n = 10, truncate = 30)
      predictions = lr.transform(testData)
      predictions.show(10)
```

text	probability	label	prediction
@KoolioHoolio see i didnt e...	[0.9983509482201712,0.00164...	1.0	0.0
you suck you suck you suck ...	[0.9956474979112804,0.00435...	0.0	0.0
super pissed that another t...	[0.9952937171301914,0.00470...	0.0	0.0
Things I'm feeling now: ang...	[0.9942597194931533,0.00574...	0.0	0.0
so sad, me equal sad, no so...	[0.9926117815441087,0.00738...	0.0	0.0
is feeling sad and stressed...	[0.9921148433314205,0.00788...	0.0	0.0
today i kinda feel sick of ...	[0.9918217293827665,0.00817...	0.0	0.0
Been sick with sore throat ...	[0.9900972897854806,0.00990...	0.0	0.0
Throat is killing me, runny...	[0.9884381180833721,0.01156...	0.0	0.0
Ugh my nose is stuffy, my t...	[0.9883616611740987,0.01163...	0.0	0.0

only showing top 10 rows

text	features	label	rawPrediction
i really2 ...	(10000, [3,5739], [...]		
0.0	[0.20568669397803...	[0.55124114575112...	0.0
jb isnt show...	(10000, [856,1150,...]		
0.0	[0.46716578122309...	[0.61471271582069...	0.0
ok thats it ...	(10000, [98,200,31...		

```

0.0|[-0.3317250925555...|0.41782094086268...|      1.0|
|   Not feeling i...|(10000,[6,54,104,...|
0.0|[1.83966456055162...|0.86290903076261...|      0.0|
|   on the comput...|(10000,[11,295,36...|
0.0|[0.16440935574308...|0.54101000400038...|      0.0|
|   BoRinG   ): wh...|(10000,[92,175,29...|
0.0|[0.65685508146551...|0.65855357269296...|      0.0|
|   Boston Globe c...|(10000,[66,72,222...|
0.0|[0.43332614758788...|0.60666764143247...|      0.0|
|   My phone can u...|(10000,[4,39,124,...|
0.0|[0.35321689180723...|0.58739744882565...|      0.0|
|   hoping to see ...|(10000,[20,171,17...|
0.0|[-0.1492856553285...|0.46274774477281...|      1.0|
|   i'm so cold th...|(10000,[35,223,70...|
0.0|[0.39772605268104...|0.59814119743731...|      0.0|
+-----+-----+-----+-----+-----+
-----+-----+
only showing top 10 rows

```

```
[25]: predictions.printSchema()
```

```

root
 |-- text: string (nullable = true)
 |-- features: vector (nullable = true)
 |-- label: double (nullable = false)
 |-- rawPrediction: vector (nullable = true)
 |-- probability: vector (nullable = true)
 |-- prediction: double (nullable = false)

```

```
[26]: from pyspark.ml.evaluation import BinaryClassificationEvaluator
bce = BinaryClassificationEvaluator()
print("Test Area Under ROC: " + str(bce.evaluate(predictions, {bce.metricName:
↪ "areaUnderROC"})))
```

Test Area Under ROC: 0.8472088126136625

```
[27]: df_selection = predictions.select("label",'prediction').toPandas()
```

```
[28]: from sklearn.metrics import classification_report
true = np.array(df_selection['label'])
pred = np.array(df_selection['prediction'])
print(classification_report(true,pred))
```

	precision	recall	f1-score	support
0.0	0.79	0.75	0.77	239942

	1.0	0.76	0.80	0.78	239778
accuracy				0.77	479720
macro avg	0.77	0.77	0.77	0.77	479720
weighted avg	0.77	0.77	0.77	0.77	479720

We can say that this model is able to distinguish whether a tweet is positive or negative is fair

Retrain our model using the fully dataset

```
[29]: lr = LogisticRegression(maxIter=20, regParam=0.3, elasticNetParam=0).
      ↪ fit(model_df)
```

Twitter Authentication

```
[30]: import tweepy
      from tweepy import OAuthHandler
      from tweepy import Stream

      ACCESS_TOKEN = "1458842253779161088-QFe06udaAdHR4VARxaDza1w4LUl0oE"
      ACCESS_TOKEN_SECRET = "tC7IJDbl5T97Zvu3kE8sdGnmZWC2qx0rkd0v90YkdzIV0"
      API_KEY = "KLP5ct26qaVoOKjAgP804j4y5"
      API_KEY_SECRET = "AbxH3913WIPG0FHIwvVRomul92RWvu0dxRo2ecXR6H0Qgibo29"

      auth = tweepy.OAuthHandler(API_KEY, API_KEY_SECRET)
      auth.set_access_token(ACCESS_TOKEN, ACCESS_TOKEN_SECRET)

      api = tweepy.API(auth)
```

Twitter Streaming Tweets

We can set the limit of tweet samples to 100 and where the streaming only streams tweets that are in english

```
[31]: tweet_list = list()
      # Subclass Stream to print IDs of Tweets received
      class IDPrinter(tweepy.Stream):

          def on_status(self, status):
              tweet_list.append(status.text)
              #print(tweet_list)
              #print(status.text)
              if len(tweet_list) == 100:
                  Stream.disconnect(self)
      # Initialize instance of the subclass
      printer = IDPrinter(
          API_KEY, API_KEY_SECRET,
          ACCESS_TOKEN, ACCESS_TOKEN_SECRET)
```

```
)

printer.sample(languages=['en'])
```

Stream connection closed by Twitter

Create new dataframe from tweet stream

```
[32]: df_2 = pd.DataFrame(np.array(tweet_list))
      df_2.columns = ['text']
      df_2 = spark.createDataFrame(df_2)
      df_2.show()
```

```
+-----+
|          text|
+-----+
|RT @NotOwenMeany:...|
|RT @CryptoTownEU:...|
|RT @MCU_Source: B...|
|@composerchris Th...|
|i actually want t...|
|RT @emailmanROCKS...|
|RT @secretyarrow:...|
|RT @UKFeetToes: B...|
|RT @djrothkopf: M...|
|RT @itsfkntxna: y...|
|RT @Ahsanhoonmein...|
|@kaslina Yah...no...|
|RT @ohits_lauren...|
|RT @zhanxfxiah: *...|
|@CSimmaWX @RColbu...|
|RT @lucifrian: ta...|
|get senjuro in th...|
|@Yeonderella99 Po...|
|RT @Matt_Stepp817...|
|Certain actions b...|
+-----+
only showing top 20 rows
```

Transforms the data received from stream

```
[33]: dataset_1 = pipelineFit.transform(df_2)
      dataset_1.show()
```

```
+-----+-----+-----+-----+
----+
|          text|          words|          filtered|
features|
```



```

+-----+-----+-----+-----+
----+
|RT @NotOwenMeany:...|[rt, notowenmeany...|[notowenmeany,
dc...|(10000,[910,1208,...|
|RT @CryptoTownEU:...|[rt, cryptotowneu...|[cryptotowneu,
ai...|(10000,[40,263,34...|
|RT @MCU_Source: B...|[rt, mcu_source, ...|[mcu_source,
brea...|(10000,[90,1122,1...|
|@composerchris Th...|[composerchris, t...|[composerchris,
l...|(10000,[883,1612,...|
|i actually want t...|[i, actually, wan...|[actually, want,
...|(10000,[26,187,73...|
|RT @emailmanROCKS...|[rt, emailmanrock...|[emailmanrocks,
r...|(10000,[38,145,59...|
|RT @secretarrow:...|[rt, secretarrow...|[secretarrow,
lo...|(10000,[996,1122,...|
|RT @UKFeetToes: B...|[rt, ukfeettoes, ...|[ukfeettoes,
biki...|(10000,[8,699,717...|
|RT @djrothkopf: M...|[rt, djrothkopf, ...|[djrothkopf,
manc...|(10000,[4,6,66,93...|
|RT @itsfkntxna: y...|[rt, itsfkntxna, ...|[itsfkntxna,
allo...|(10000,[69,71,175...|
|RT @Ahsanhoonmein...|[rt, ahsanhoonmei...|[ahsanhoonmein,
s...|(10000,[597,1122,...|
|@kaslina Yah...no...|[kaslina, yah, no...|[kaslina, yah,
do...|(10000,[1124,1922...|
|RT @ohits_laurenn...|[rt, ohits_lauren...|[ohits_laurenn,
a...|(10000,[820,1204,...|
|RT @zhanxfxiah: *...|[rt, zhanxfxiah, ...|[zhanxfxiah,
masa...|(10000,[24,53,58,...|
|@CSimmaWX @RColbu...|[csimmawx, rcolbu...|[csimmawx,
rcolbu...|(10000,[1,127,334...|
|RT @lucifrian: ta...|[rt, lucifrian, t...|[lucifrian, tae,
...|(10000,[126,250,2...|
|get senjuro in th...|[get, senjuro, in...|[get, senjuro,
po...|(10000,[2,90,304,...|
|@Yeonderella99 Po...|[yeonderella99, p...|[yeonderella99,
p...|(10000,[1022,6360...|
|RT @Matt_Stepp817...|[rt, matt_stepp81...|[matt_stepp817,
g...|(10000,[0,44,77,1...|
|Certain actions b...|[certain, actions...|[certain,
actions...|(10000,[3,6,27,34...|
+-----+-----+-----+-----+
----+
only showing top 20 rows

```

Selecting text and Features

```
[34]: dataset_test = dataset_1.select('text', 'features')
```

Having the pre-trained model predict on tweets

```
[35]: model_predictions = lr.transform(dataset_test)
```

```
[36]: model_predictions.show()
```

```
+-----+-----+-----+-----+
----+-----+
|          text|          features|          rawPrediction|
probability|prediction|
+-----+-----+-----+-----+
----+-----+
|RT @NotOwenMeany:...|(10000,[910,1208,...|[-0.4205481364755...|[0.3963855933338
4...|          1.0|
|RT @CryptoTownEU:...|(10000,[40,263,34...|[-0.24266754444595...|[0.4396290814197
0...|          1.0|
|RT @MCU_Source:
B...|(10000,[90,1122,1...|[-0.1111318162142...|[0.47224560461601...|          1.0|
|@composerchris
Th...|(10000,[883,1612,...|[0.33328114913442...|[0.58255751614137...|          0.0|
|i actually want
t...|(10000,[26,187,73...|[0.41763114414529...|[0.60291626474052...|          0.0|
|RT @emailmanROCKS...|(10000,[38,145,59...|[-1.3859129547787...|[0.2000610319972
3...|          1.0|
|RT @secretarrow:...|(10000,[996,1122,...|[-0.1662206607879...|[0.4585402495590
8...|          1.0|
|RT @UKFeetToes:
B...|(10000,[8,699,717...|[-0.4050399582378...|[0.40010204030557...|          1.0|
|RT @djrothkopf:
M...|(10000,[4,6,66,93...|[0.45148455172166...|[0.61099214139275...|          0.0|
|RT @itsfkntxna:
y...|(10000,[69,71,175...|[0.81005702034367...|[0.69212165485828...|          0.0|
|RT @Ahsanhoonmein...|(10000,[597,1122,...|[-0.5938158745110...|[0.3557597952772
0...|          1.0|
|@kaslina
Yah...no...|(10000,[1124,1922...|[-0.2401970645173...|[0.44023778785195...|
1.0|
|RT @ohits_laurenn...|(10000,[820,1204,...|[0.30813505754837...|[0.5764299846857
7...|          0.0|
|RT @zhanxfxiah:
*...|(10000,[24,53,58,...|[-0.8388083176098...|[0.30178582542580...|          1.0|
|@CSimmaWX
@RColbu...|(10000,[1,127,334...|[-0.1438396953481...|[0.46410194862923...|
1.0|
|RT @lucifrian:
ta...|(10000,[126,250,2...|[-0.3218737575057...|[0.42021916678569...|          1.0|
```

```
|get senjuro in
th...|(10000,[2,90,304,...]|[-0.4225067515623...|[0.39591706238032...|      1.0|
|@Yeonderella99
Po...|(10000,[1022,6360...]|[-0.3696149877333...|[0.40863405737944...|      1.0|
|RT @Matt_Stepp817...|(10000,[0,44,77,1...]|[-1.1734634905797...|[0.2362295127502
0...|      1.0|
|Certain actions
b...|(10000,[3,6,27,34...|[0.16488850878777...|[0.54112898407193...|      0.0|
+-----+-----+-----+-----+
----+-----+
only showing top 20 rows
```

Storing Results and Text in Database

Create connection to mongoDB, database, and table

```
[37]: import pymongo
myclient = pymongo.MongoClient("mongodb://localhost:27017/")
mydb = myclient["my_database"]
mycol = mydb["predictions"]
```

Converting spark dataframe to pandas dataframe

```
[41]: mp_df = model_predictions.toPandas()
mp_df.head()
```

```
[41]:                                     text \
0  RT @NotOwenMeany: DC Christmas Light Hunters: ...
1  RT @CryptoTownEU:  Airdrop: BAoE Global\n Va...
2  RT @MCU_Source: BREAKING: #Deadpool's first ap...
3  @composerchris They should lose their tax exem...
4                                     i actually want to die.

                                     features \
0  (0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, ...
1  (0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, ...
2  (0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, ...
3  (0.0, 0.0, 0.0, 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, 0.0, 0.0, 0.0, ...

                                     rawPrediction \
0  [-0.42054813647552053, 0.42054813647552053]
1  [-0.24266754445953495, 0.24266754445953495]
2  [-0.11113181621423963, 0.11113181621423963]
3  [0.3332811491344227, -0.3332811491344227]
4  [0.417631144145294, -0.417631144145294]

                                     probability prediction
```

0	[0.39638559333384943, 0.6036144066661506]	1.0
1	[0.43962908141970913, 0.5603709185802909]	1.0
2	[0.47224560461601417, 0.5277543953839858]	1.0
3	[0.5825575161413753, 0.4174424838586247]	0.0
4	[0.6029162647405205, 0.3970837352594795]	0.0

Inserting into MongoDB

```
[42]: for index, row in mp_df.iterrows():
      #print(row['text'], row['prediction'])
      mydict = { "text": row['text'], "prediction": row['prediction'] }
      mycol.insert_one(mydict)
```

RT @NotOwenMeany: DC Christmas Light Hunters: The area north of American University (roughly Van Ness to Davenport and 43rd to 49th streets... 1.0

RT @CryptoTownEU: Airdrop: BAoE Global
Value: 10 \$BAoE
Referral: 3 \$BAoE
Partnership: Groo INTERNATIONAL, Meta Ultra Holdings, SEM... 1.0

RT @MCU_Source: BREAKING: #Deadpool's first appearance at the #MCU will be in #DoctorStrangeInTheMultiverseOfMadness! <https://t.co/KZYAVcAC...> 1.0

@composerchris They should lose their tax exempt status 0.0

i actually want to die. 0.0

RT @emailmanROCKS: The Rules of Creative Writing by children's author Kurt Chambers <https://t.co/lrQszZIOxL> ... This is a great guide to help... 1.0

RT @secretyarrow: how's it lookin daddy? <https://t.co/EdUjUVIuSk> 1.0

RT @UKFeetToes: Bikini Body & White Toes. Love those legs!
<https://t.co/xFYnY7AGSe> 1.0

RT @djrothkopf: Manchin did not have to say anything today. Did not have to go on Fox. Did not have to issue a statement. He could've kept... 0.0

RT @itsfkntxna: you are allowed to criticize your own community and i hate to tell you this but other people are allowed to criticize your... 0.0

RT @Ahsanhoonmein: Spread the word. <https://t.co/jmBAjtbOI5> 1.0

@kaslina Yah...no doubt! 1.0

RT @ohits_laurenn: i be asking for a sign then i be ignoring the sign 0.0

RT @zhanxfxiah: *masa sees yuki watching*

Masa: "Yuki definitely won't come/join though"

HE KNEW HIM ALL TOO WELL
<https://t.co/oDsx3a0...> 1.0

@CSimmaWX @RColburnnews @AlexCounteeWX @_WeatherStove @PatrickLloydTV Hopefully you went with Wilson JR @CSimmaWX he's having a huge day 1.0

RT @lucifrian: tae: hyung look at my video with facial hair
joon: ?
joon: tae
joon: tae where
joon: tae

joon: tae where isit
joon: wheresth... 1.0
get senjuro in the poll get him first place :(1.0
@Yeonderella99 Poji = Visual of the group wbk 1.0
RT @Matt_Stepp817: Good chance both 4A state champions end up in the same district next year...that ought to be fun #txhsfb @dctf 1.0
Certain actions by NFL officials, like the spot on that last Titans' play today, make you think things might not ex... <https://t.co/V42NPWT3UL> 0.0
Oh just take a look at this! <https://t.co/Rjbc48uSEj> 1.0
@larymary60 @jeremycorbyn Me too. 1.0
RT @uhhchrxs: @MekaApe PUBLIC SALES IS LIVE GO CONNECT YOUR WALLET NOW

THERE WILL NOT BE ENOUGH FOR EVERYONE!

WELCOME TO MEKA APE FA... 1.0
RT @STUDIO_0613: taehyung they're just jealous bts is so popular & has a lot of fans & is doing well
ignore them & quietly enter the legal... 1.0
RT @Michael_Jaayy: Tariq really his daddy's son #PowerBook2 #PowerGhost <https://t.co/AdhGdjBNup> 0.0
My partner is literally the best he stayed over my place just to clean my room so I can be cozy after I got back fr... <https://t.co/Yg9dn01nKk> 1.0
@OneLedgerTech @binance @PancakeSwap PAF DAO PRESALE WHITELISTING
AUDIT & KYC COMPLETED
PRIVATE SALE AND AIRDR... <https://t.co/DFlonhqchB> 1.0
@cryptopalis Hey, saw you've been looking at some NFT games like Wolf Game and WnD. I just found @PeopleGameNFT, th... <https://t.co/1L6abtVzLb> 1.0
@GlexServices @manackz3 Done 1.0
@M_J_C13 I understand why he's calling it. But it's different when lock was practicing it last year as the starter... <https://t.co/QL5IEsPjXD> 0.0
@tori_mcelhaney @AtlantaFalcons Wait he was in the game? Bum ass 1.0
Bike lanes and creativity. Here's the governor of São Paulo, Brazil on the Pinheiros River bike lane, inaugurating... <https://t.co/eCx8uKGUGI> 1.0
If I'm too much, go find less 0.0
RT @jaeedimples: thank you our #nct127 we love you <https://t.co/LEKnM7kjhM> 1.0
RT @TFGmykL: Just realized my mutant was up for sale and someone bought it 2 days ago

Looks like one of you guys is gonna win my 3rd muta... 0.0
@bloodstoppin when catholics kneel and open their mouth so the body of a man can enter... 0.0
RT @grade1to6: Next-Gen Math & English Worksheets from <https://t.co/CljkGA3gX4>. Subscribe (\$1.25 / Rs 100 a year) 6,000 Plus Worksheets for... 1.0
RT @D_Kumii: I like the way mangaka maintain their privacy <https://t.co/C1dT0omihJ> 1.0
@keeptalkinglol i guess i could do worse than Black Kimchi 0.0
RT @glamgamerbytch: Who in their right mind doesn't like cheetah print I need to

get a special cheetah print outfit for the @LingerieFC w... 1.0
RT @misslefty_: PS. Goodnight Tito DJ

2G2BT TeaserReveal

#AnditoTayoParaSaKathNiel <https://t.co/JKG8u0cTiJ> 1.0

@AngelRiveraXXX i do. ur so fkn cute 1.0

thanks bro. i don't feel bad enough about not making enough money at work 0.0

@CryptoTownEU I hope this project develops rapidly in the future, happy to be able to take part in this project... <https://t.co/RaHubVer6k> 1.0

@CaniacZach @james_hockey61 @Canes @Chicago_Wolves He doesn't look better than Jarvis. You can think what you want,... <https://t.co/M4jpSM4kxX> 1.0

Join me in my Space! <https://t.co/xSuR0xBMNQ> 1.0

@guardian ...people hve died from the vaccines, blood clots, no matter how small the risk. People must be allowed t... <https://t.co/V6eoRu68dw> 0.0

@kiramadeira I am aaaangry!!!! Been using them for over 10 years and this is how they repay me the survey I took... <https://t.co/YoUdKu9MXq> 1.0

British tennis is in very good hands! @EmmaRaducanu 1.0

Is she dead ? <https://t.co/10Fpt39zQQ> 0.0

@PAKA_FLOCKA Yea I think imma pass on this Taylor Jenkins rng lineup from the looks of it 1.0

RT @wofai_xo: That kiss after you eat her out >>>>>>>>
https://t.co/NMBAxqtAXs 1.0

Steak your claim on this Legendary career opportunity with our #Plattsburgh, NY team: Host <https://t.co/u73yAztk39> #Host 1.0

RT @SmokinHotTX: It has been speculated that I never wear clothes...I actually do own them and wear them upon occasion. LOL. But I think I... 1.0

@CBSNews 7 Signs of weak immune system

How to know if you have weak immune system to avoid any complications

<https://t.co/DWo0VX30RU> 0.0

RT @russiamey: do you ever just go:

```
men          ran haitani
          0.0
```

```
@kvmhobi u look so pretty ! 1.0
```

@LambertBrownJa Lambert Brown is Narsisistic. PERIOD. Please give up the stipend
1.0

```
{
  "round": 1484223,
  "signature":
  "b6ea4cd2162c0bc81e7d69114f6631438be1cf5fa1ed89359be94139fc22bd64b8e6d32471143..."
}
```

grown ass woman. 1.0

RT @fantasia_daily: When I'm singing When I See U around my friends and they don't join in. <https://t.co/0kQ4p49QGH> 1.0

That 3rd down scramble was so good from Burrow

Loops, stunts. Stayed patient, let his OL pick it up. Picked his sp...
<https://t.co/tA0azcD8Qc> 1.0
 RT @ashleysimpo: Now is a good time to reassess who gets to take up space in
 your life and your mind.

 That includes who you follow on soci... 1.0
 RT @cottontail0112: present for you <https://t.co/UY8v3Cyr27> 1.0
 Don't forget to take your supplements. They'll make you even stronger! 1.0
 @yoongisducky g'night, sleep well ! 1.0
 RT @bitchimcierra: ima be single forever niggas don't excite me nm 0.0
 @ollybarratt Does he watch MOTD? 1.0
 @MouseOfReason You don't think your followers see your behavior? Why do you
 think I don't just block you? lol.

 Yo... <https://t.co/qxvgLyCpyY> 1.0
 Actually nuts 1.0
 @DeonFialkov7775 Try @caltechack on @Instagram his a good hacker he help me
 recover my wallet account. When I met... <https://t.co/Ge4SQ6RbjU> 1.0
 WHAT DO YOU MEAN NOT CUTE <https://t.co/qUi6fmoYRY> 1.0
 @CitrusWavez no. who does that. 1.0
 @xxynchim happy monday, have a nice day and be happy for today dela ! 1.0
 RT @blacksunflower: Grateful for life, love, breath and of course coffee. In all
 things strive for balance, less ego and more love. When yo... 1.0
 @nurirast Wow really sexy 1.0
 RT @BenstarDEV: sloooooow curve <https://t.co/II6992Pnsy> 1.0
 RT @GravelInstitute: Socialist candidate Gabriel Boric has won Chile's
 presidential election, beating a far-right candidate who supports Tr... 1.0
 RT @vpungs: If I'm honest I don't think I've processed many emotions this year.
 Traumatic things have happened and I feel like there's been... 0.0
 RT @alley_junk: @kamenism also people who cry abt China never really care abt
 the specific event beyond the sensationalized 2 week news cyc... 0.0
 RT @_Teajahlashae: Somebody said "I can't wait to live, I just been surviving "
 and I felt that 0.0
 @drajm @vanbadham Or a very large and excited red setter. 1.0
 bring the dead chinchilla chat member. 0.0
 @eagle4congress @TrishT1966 @NikkiFried @ChristinaPushaw Hers is real 1.0
 RT @wayydaminn: Be your best friend. Be your gym partner. Be self motivated. Be
 your partner in crime. Be everything for yourself rather th... 1.0
 @ace_maddie They don't know what to do when someone follows the rules. Matter
 of fact, they probably don't know th... <https://t.co/VgrZGUpaXv> 1.0
 @CryptoSenorita 2021 wrapped in a GIF 1.0
 RT @avalanche_edin: Just bought 800 LPs and 4,500 CDs from an old customer.
 Everything from Low to techno with a massive helping of Scottis... 0.0
 @DRiveraTX @creatureNFT @dannycoleee Hahahah my apologies but I couldn't leave
 him there!! 0.0
 @SteveSkojec I'm sneaking up, kamikaze style. You will receive a 2 hour warning.
 I'll bring the cocktail fixins. 1.0
 RT @FaazNoushad: @ClippedHussar Me (who've also avoided catching COVID every

wave.) <https://t.co/RK3BLaBwbZ> 1.0
 RT @caro_painter: A little #behindthescenes of our lives @caenhillcc
 Live everyday on Facebook, Youtube and TikTok updating you about th... 1.0
 If you used my Social and got cash or a car and have kids I don't feel sorry for
 you. Watch what happens. 0.0
 RT @dobsalvatore: the way they look at eo <https://t.co/CWv40ME8h2> 1.0
 @sswidamh @erecalie r u 1000000% over them? or is there a small part of you that
 still would 0.0
 @CarlonCarpenter Playing out from the back. Largely situational, can be
 something you do more often than other teams, High risk/high reward 1.0
 Aaron breaking the record tonight 0.0
 RT @chartdata: .@BTS_twt's "Permission to Dance" has now surpassed 400 million
 views on YouTube. 1.0
 CD Media | 99% In Democratic Party Referendum Vote To Oust Albanian Opposition
 Leader Basha <https://t.co/1Ex8Bk6z0V> 1.0
 i'm sorry but sometimes i hate my soft ass friends bc once we argue bxches jus
 irking ... then we got my hard friends... <https://t.co/7UcNQN Uttw> 0.0

Definitions

0.0 means that the tweet was negative

1.0 means that the tweet was positive

Retrieving Positive Tweets

```
[47]: myquery = { "prediction": 1.0 }
      mydoc = mycol.find(myquery)
      for x in mydoc:
          print(x)
```

```
{'_id': ObjectId('61bfc4d9b8bc24cd819b369c'), 'text': 'RT @NotOwenMeany: DC
Christmas Light Hunters: The area north of American University (roughly Van Ness
to Davenport and 43rd to 49th streets...', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b369d'), 'text': 'RT @CryptoTownEU:
Airdrop: BAoE Global\n Value: 10 $BAoE\n Referral: 3 $BAoE\n Partnership:
Groo INTERNATIONAL, Meta Ultra Holdings, SEM...', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b369e'), 'text': 'RT @MCU_Source: BREAKING:
#Deadpool's first appearance at the #MCU will be in
#DoctorStrangeInTheMultiverseOfMadness! https://t.co/KZYAVcAC...', 'prediction':
1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36a1'), 'text': "RT @emailmanROCKS: The
Rules of Creative Writing by children's author Kurt Chambers
https://t.co/lrQszZlOxL ... This is a great guide to help...", 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36a2'), 'text': 'RT @secretyarrow: how's
it lookin daddy? https://t.co/EdUjUVIuSk', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36a3'), 'text': 'RT @UKFeetToes: Bikini
Body & White Toes. Love those legs! https://t.co/xFYnY7AGSe', 'prediction':
1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36a6'), 'text': 'RT @Ahsanhoonmein: Spread
```


the word. <https://t.co/jmBAjtb0I5>', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36a7'), 'text': '@kaslina Yah...no doubt!', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36a9'), 'text': 'RT @zhanxfxiah: *masa sees yuki watching*\n\nMasa: "Yuki definitely won\'t come/join though" \n\nHE KNEW HIM ALL TOO WELL \nhttps://t.co/oDsx3a0...', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36aa'), 'text': '@CSimmaWX @RColburnnews @AlexCounteeWX @WeatherStove @PatrickLloydTV Hopefully you went with Wilson JR @CSimmaWX he's having a huge day', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ab'), 'text': 'RT @lucifrian: tae: hyung look at my video with facial hair\nnjoon: ?\nnjoon: tae\nnjoon: tae where\nnjoon: tae\nnjoon: tae where isit\nnjoon: wheresth...', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ac'), 'text': 'get senjuro in the poll get him first place :(', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ad'), 'text': '@Yeonderella99 Poji = Visual of the group wbk ', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ae'), 'text': 'RT @Matt_Stepp817: Good chance both 4A state champions end up in the same district next year...that ought to be fun #txhsfb @dctf', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36b0'), 'text': 'Oh just take a look at this! <https://t.co/Rjbc48uSEj>', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36b1'), 'text': '@larymary60 @jeremycorbyn Me too.', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36b2'), 'text': 'RT @uhhchrxs: @MekaApe PUBLIC SALES IS LIVE GO CONNECT YOUR WALLET NOW\n\nTHERE WILL NOT BE ENOUGH FOR EVERYONE! \n\nWELCOME TO MEKA APE FA...', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36b3'), 'text': "RT @STUDIO_0613: taehyung they're just jealous bts is so popular & has a lot of fans & is doing well\nnignore them & quietly enter the legal...", 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36b5'), 'text': 'My partner is literally the best he stayed over my place just to clean my room so I can be cozy after I got back fr... <https://t.co/Yg9dn01nKk>', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36b6'), 'text': '@OneLedgerTech @binance @PancakeSwap PAF DAO PRESALE WHITELISTING\n\nAUDIT & KYC COMPLETED\n\nPRIVATE SALE AND AIRDR... <https://t.co/DFlonhqchB>', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36b7'), 'text': '@cryptopolis Hey, saw you've been looking at some NFT games like Wolf Game and WnD. I just found @PeopleGameNFT, th... <https://t.co/1L6abtVzLb>', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36b8'), 'text': '@GlexServices @manackz3 Done', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ba'), 'text': '@tori_mcelhaney @AtlantaFalcons Wait he was in the game? Bum ass', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36bb'), 'text': "Bike lanes and creativity. Here's the governor of São Paulo, Brazil on the Pinheiros River bike lane, inaugurating... <https://t.co/eCx8uKGUgI>", 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36bd'), 'text': 'RT @jaeedimples: thank you our #nct127 we love you <https://t.co/LEKnM7kjhM>', 'prediction': 1.0}

[illegible]

{'_id': ObjectId('61bfc4d9b8bc24cd819b36d8'), 'text': 'RT @fantasia_daily: When I\'m singing When I See U around my friends and they don\'t join in. https://t.co/0kQ4p49QGH', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36d9'), 'text': 'That 3rd down scramble was so good from Burrow\n\nLoops, stunts. Stayed patient, let his OL pick it up. Picked his sp... https://t.co/tA0azcD8Qc', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36da'), 'text': 'RT @ashleysimpo: Now is a good time to reassess who gets to take up space in your life and your mind.\n\nThat includes who you follow on soci...', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36db'), 'text': 'RT @cottontail0112: present for you https://t.co/UY8v3Cyr27', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36dc'), 'text': '"Don\'t forget to take your supplements. They\'ll make you even stronger!"', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36dd'), 'text': '"@yoongisducky g\'night, sleep well ! "', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36df'), 'text': '@ollybarratt Does he watch MOTD?', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36e0'), 'text': '@MouseOfReason You don\'t think your followers see your behavior? Why do you think I don\'t just block you? lol. \n\nYo... https://t.co/qxvgLyCpyY", 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36e1'), 'text': 'Actually nuts', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36e2'), 'text': '@DeonFialkov7775 Try @caltechack on @Instagram his a good hacker he help me recover my wallet account. When I met... https://t.co/Ge4SQ6RbjU', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36e3'), 'text': 'WHAT DO YOU MEAN NOT CUTE https://t.co/qUi6fmoYRY', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36e4'), 'text': '@CitrusWavez no. who does that.', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36e5'), 'text': '@xxynchim happy monday, have a nice day and be happy for today dela ! ', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36e6'), 'text': 'RT @blacksunflower: Grateful for life, love, breath and of course coffee. In all things strive for balance, less ego and more love. When yo...', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36e7'), 'text': '@nurirast Wow really sexy ', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36e8'), 'text': 'RT @BenstarDEV: sloooooooooow curve https://t.co/II6992Pnsy', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36e9'), 'text': 'RT @GravelInstitute: Socialist candidate Gabriel Boric has won Chile\'s presidential election, beating a far-right candidate who supports Tr...', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ed'), 'text': '@drajm @vanbadham Or a very large and excited red setter.', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ef'), 'text': '@eagle4congress @TrishT1966 @NikkiFried @ChristinaPushaw Hers is real \u200d', 'prediction': 1.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36f0'), 'text': 'RT @wayydaminn: Be your best friend. Be your gym partner. Be self motivated. Be your partner in crime.

```

Be everything for yourself rather th...', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36f1'), 'text': '@ace_maddie They don't
know what to do when someone follows the rules. Matter of fact, they probably
don't know th... https://t.co/VgrZGUpaXv', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36f2'), 'text': '@CryptoSenorita 2021
wrapped in a GIF', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36f5'), 'text': '@SteveSkojec I'm sneaking
up, kamikaze style. You will receive a 2 hour warning. I'll bring the cocktail
fixins.', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36f6'), 'text': 'RT @FaazNoushad:
@ClippedHussar Me (who've also avoided catching COVID every wave.)
https://t.co/RK3BLaBwbZ', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36f7'), 'text': 'RT @caroPainter: A
little #behindthescenes of our lives @caenhillcc \nLive everyday on Facebook,
Youtube and TikTok updating you about th...', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36f9'), 'text': 'RT @dobsalvatore: the way
they look at eo https://t.co/CWv40ME8h2', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36fb'), 'text': '@CarltonCarpenter Playing
out from the back. Largely situational, can be something you do more often than
other teams, High risk/high reward', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36fd'), 'text': 'RT @chartdata:
.@BTS_twt\'s "Permission to Dance" has now surpassed 400 million views on
YouTube.', 'prediction': 1.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36fe'), 'text': 'CD Media | 99% In
Democratic Party Referendum Vote To Oust Albanian Opposition Leader Basha
https://t.co/1Ex8Bk6z0V', 'prediction': 1.0}

```

Retrieving Negative Tweets

```

[48]: myquery = { "prediction": 0.0 }
      mydoc = mycol.find(myquery)
      for x in mydoc:
          print(x)

```

```

{'_id': ObjectId('61bfc4d9b8bc24cd819b369f'), 'text': '@composerchris They
should lose their tax exempt status', 'prediction': 0.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36a0'), 'text': 'i actually want to die.',
'prediction': 0.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36a4'), 'text': 'RT @djrothkopf: Manchin
did not have to say anything today. Did not have to go on Fox. Did not have to
issue a statement. He could've kept...', 'prediction': 0.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36a5'), 'text': 'RT @itsfkntxna: you are
allowed to criticize your own community and i hate to tell you this but other
people are allowed to criticize your...', 'prediction': 0.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36a8'), 'text': 'RT @ohits_laurenn: i be
asking for a sign then i be ignoring the sign \u200d', 'prediction': 0.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36af'), 'text': 'Certain actions by NFL
officials, like the spot on that last Titans' play today, make you think things

```

might not ex... <https://t.co/V42NPWT3UL>', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36b4'), 'text': 'RT @Michael_Jaayy: Tariq really his daddy's son #PowerBook2 #PowerGhost <https://t.co/AdhGdjBNup>', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36b9'), 'text': '@M_J_C13 I understand why he's calling it. But it's different when lock was practicing it last year as the starter... <https://t.co/QL5IEsPjXD>', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36bc'), 'text': 'If I'm too much, go find less', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36be'), 'text': 'RT @TFGmykL: Just realized my mutant was up for sale and someone bought it 2 days ago \n\nLooks like one of you guys is gonna win my 3rd muta...', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36bf'), 'text': '@bloodstoppin when catholics kneel and open their mouth so the body of a man can enter...', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36c2'), 'text': '@keeptalkinglol i guess i could do worse than Black Kimchi ', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36c6'), 'text': 'thanks bro. i don't feel bad enough about not making enough money at work \U0001f972', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ca'), 'text': '@guardian ...people hve died from the vaccines, blood clots, no matter how small the risk. People must be allowed t... <https://t.co/V6eoRu68dw>', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36cd'), 'text': 'Is she dead ? <https://t.co/10Fpt39zQQ>', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36d2'), 'text': '@CBSNews 7 Signs of weak immune system\n\nHow to know if you have weak immune system to avoid any complications \nhttps://t.co/DWo0VX30RU', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36d3'), 'text': 'RT @russiamey: do you ever just go:\n\n \n men ran haitani \n', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36de'), 'text': 'RT @bitchimcierra: ima be single forever niggas don't excite me nm ', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ea'), 'text': 'RT @vpungs: If I'm honest I don't think I've processed many emotions this year. Traumatic things have happened and I feel like there's been...', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36eb'), 'text': 'RT @alley_junk: @kamenism also people who cry abt China never really care abt the specific event beyond the sensationalized 2 week news cyc...', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ec'), 'text': 'RT @Teajahlashae: Somebody said "I can't wait to live, I just been surviving " and I felt that', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36ee'), 'text': 'bring the dead chinchilla chat member.', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36f3'), 'text': 'RT @avalanche_edin: Just bought 800 LPs and 4,500 CDs from an old customer. Everything from Low to techno with a massive helping of Scottis...', 'prediction': 0.0}

{'_id': ObjectId('61bfc4d9b8bc24cd819b36f4'), 'text': '@DRiveraTX @creatureNFT @dannycoleee Hahahah my apologies but I couldn't leave him there!!',

```
'prediction': 0.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36f8'), 'text': 'If you used my Social and
got cash or a car and have kids I don't feel sorry for you. Watch what
happens.', 'prediction': 0.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36fa'), 'text': '@sswidamh @erecalie r u
1000000% over them? or is there a small part of you that still would',
'prediction': 0.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36fc'), 'text': 'Aaron breaking the record
tonight', 'prediction': 0.0}
{'_id': ObjectId('61bfc4d9b8bc24cd819b36ff'), 'text': 'i'm sorry but sometimes i
hate my soft ass friends bc once we argue bxches jus irking ... then we got my
hard friends... https://t.co/7UcNQNUttw', 'prediction': 0.0}
```