

# Final Project

For this final project, the author will perform sentiment analysis (lexicon and machine learning) of publicly available tweets to compare the public perception of Donald Trump before the 2016 and 2020 US Presidential Elections.

Name: Lee Waters

ID: dld011

## Define LexiconClassifier() class

```
In [1]: # DO NOT MODIFY THE CODE IN THIS CELL
class LexiconClassifier():
    def __init__(self):
        self.positive_words = set()
        with open('positive-words.txt', encoding = 'utf-8') as iFile:
            for row in iFile:
                self.positive_words.add(row.strip())

        self.negative_words = set()
        with open('negative-words.txt', encoding='iso-8859-1') as iFile:
            for row in iFile:
                self.negative_words.add(row.strip())

    def predict(self, sentence):
        num_pos_words = 0
        num_neg_words = 0
        for word in sentence.lower().split():
            if word in self.positive_words:
                num_pos_words += 1
            elif word in self.negative_words:
                num_neg_words += 1

        pred = 'neutral'
        if num_pos_words > num_neg_words:
            pred = 'positive'
        elif num_pos_words < num_neg_words:
            pred = 'negative'

        return pred

    def count_pos_words(self, sentence):
        num_pos_words = 0
        for word in sentence.lower().split():
            if word in self.positive_words:
                num_pos_words += 1
        return num_pos_words

    def count_neg_words(self, sentence):
        num_neg_words = 0
        for word in sentence.lower().split():
            if word in self.negative_words:
                num_neg_words += 1
        return num_neg_words
```

## Sentiment analysis

For this task, the author uses `LexiconClassifier()` to classify all of the tweets in a real Twitter dataset. The code does the following:

1. Keeps track of the number of tweets
2. Keeps track of the number of positive and negative tweets
3. Keeps track of the user that tweets the most
4. Keeps track of the total number of unique users
5. Keeps track of the average number of tweets per user (how many tweets does each user make, on average)
6. Keeps track of the most positive and negative tweets.

Note to self:

- Do **not** store all of the tweets in a list. This will use too much memory because of the size of the dataset. It is okay to store all of the user's screen names.
- The dataset is big! Work on a subset of the dataset to make sure code works, i.e., "break" after the first 100 lines.

**2016**

```

In [2]: %%time

import csv

# define variables

lc = LexiconClassifier() # initiate class

total_number_of_tweets_2016 = 0 #1
total_number_of_positive_tweets_2016 = 0 #2
total_number_of_negative_tweets_2016 = 0 #2

user_tweets_2016 = {} # dict of screen_name: # of tweets
user_with_most_tweets_2016 = '' #3
tweets_by_biggest_user_2016 = 0 #3
total_number_of_users_2016 = 0 #4
average_number_tweets_per_user_2016 = 0 #5

most_positive_tweet_2016 = '' #6
most_negative_tweet_2016 = '' #6
most_positive_tweet_score_2016 = 0 #6
most_negative_tweet_score_2016 = 0 #6

# run code

twitter_dataset = open('trump-2016.csv', 'r', encoding='iso-8859-1')
tweet_string = csv.reader(twitter_dataset, delimiter=',')

count = 0 # skip first 100 tweets because machine learning will train/ test on this

for row in tweet_string:

    count += 1

    if count > 100:
        # extract tweet and screen_name, update dict, increment

        tweet_text = row[3]
        screen_name = row[1]

        user_tweets_2016[screen_name] = user_tweets_2016.get(screen_name, 0) + 1

        total_number_of_tweets_2016 += 1

        # count pos and neg words, calculate sentiment score and prediction

        num_pos_words = lc.count_pos_words(tweet_text)
        num_neg_words = lc.count_neg_words(tweet_text)

        sentiment_score = num_pos_words - num_neg_words
        sentiment_prediction = lc.predict(tweet_text)

        # record most positive/negative tweets/users

        if sentiment_score > most_positive_tweet_score_2016:
            most_positive_tweet_2016 = tweet_text
            most_positive_tweet_score_2016 = sentiment_score

        elif sentiment_score < most_negative_tweet_score_2016:
            most_negative_tweet_2016 = tweet_text
            most_negative_tweet_score_2016 = sentiment_score

        # increment positive/ negative tweets

        if sentiment_prediction == 'positive':
            total_number_of_positive_tweets_2016 += 1

        elif sentiment_prediction == 'negative':
            total_number_of_negative_tweets_2016 += 1

#debug
# if total_number_of_tweets_2016 > 10:
#     break

```

```
# total users, avg. tweets per user, user with most tweets quantity
total_number_of_users_2016 = len(user_tweets_2016)
average_number_tweets_per_user_2016 = sum(user_tweets_2016.values()) / total_number_of_users_2016

for key, value in user_tweets_2016.items():
    if value > tweets_by_biggest_user_2016:
        user_with_most_tweets_2016 = key
        tweets_by_biggest_user_2016 = value

twitter_dataset.close()
```

Wall time: 16.5 s

```
In [3]: print("Total Number of Tweets: {}".format(total_number_of_tweets_2016))
print("Total Number of Positive Tweets: {}".format(total_number_of_positive_tweets_2016))
print("Total Number of Negative Tweets: {}".format(total_number_of_negative_tweets_2016))
print("Total Number of Neutral Tweets: {}\\n".format(total_number_of_tweets_2016-total_number_of_positive_tweets_2016-total_number_of_negative_tweets_2016))

print("Most Positive Tweet")
print(most_positive_tweet_2016)
print(most_positive_tweet_score_2016)
print()

print("Most Negative Tweet")
print(most_negative_tweet_2016)
print(most_negative_tweet_score_2016)
print()

print("Total Number of Users: {}".format(total_number_of_users_2016))
print("Average Number of Tweets per User: {}".format(average_number_tweets_per_user_2016))
print("User with the most tweets: {} ({}).format(user_with_most_tweets_2016, tweets_by_biggest_user_2016))
```

Total Number of Tweets: 647548  
 Total Number of Positive Tweets: 271978  
 Total Number of Negative Tweets: 109701  
 Total Number of Neutral Tweets: 265869

Most Positive Tweet

@Ted\_Strickland TRUMP NOW Trump Forever Trump Now Trump Forever Trump Now Trump Forever Trump Now Trump Forever  
 8

Most Negative Tweet

@cnni @cnnbrk @CNN LIARS LIARS LIARS LIARS LIARS LIARS LIARS LIARS LIARS LIARS LIARS LIARS LIARS LIARS LIARS LIAR  
 S!! <https://t.co/jrnPHFET6T>  
 -15

Total Number of Users: 257204  
 Average Number of Tweets per User: 2.517643582525933  
 User with the most tweets: PolitickDick (1001)

```
In [4]: assert(isinstance(total_number_of_tweets_2016, int) or isinstance(total_number_of_tweets_2016, float))
assert(isinstance(total_number_of_positive_tweets_2016, int) or isinstance(total_number_of_positive_tweets_2016, float))
assert(isinstance(total_number_of_negative_tweets_2016, int) or isinstance(total_number_of_negative_tweets_2016, float))
assert(isinstance(most_positive_tweet_2016, str))
assert(isinstance(most_negative_tweet_2016, str))
assert(isinstance(user_with_most_tweets_2016, str))
assert(total_number_of_tweets_2016 == 647548)
print("Assert finished successfully!")
```

Assert finished successfully!

## Manual Analysis

Perform manual analysis of the predictions. Modify the code to load the tweet text.

```
In [5]: twitter_dataset = open('trump-2016.csv', 'r', encoding='iso-8859-1')
tweet_string = csv.reader(twitter_dataset, delimiter=',')

num_tweets_to_print = 100
num_tweets = 0
accuracy_count = 0

for row in tweet_string:

    num_tweets += 1
    tweet_text = row[3]
    manual_annotation = row[4]

    num_pos_words = lc.count_pos_words(tweet_text)
    num_neg_words = lc.count_neg_words(tweet_text)

    sentiment_score = num_pos_words - num_neg_words
    sentiment_prediction = lc.predict(tweet_text)

    if manual_annotation == sentiment_prediction:
        accuracy_count += 1

    print("Tweet {}: {}".format(num_tweets, tweet_text))
    print("Tweet {}: Manual Annotation: {}".format(num_tweets, manual_annotation))
    print("Tweet {}: Lexicon Prediction: {}".format(num_tweets, sentiment_prediction))
    print()

    if num_tweets == num_tweets_to_print:
        break

twitter_dataset.close()
```

Tweet 1: RT @HuffPostBiz: Trump's new economic plan is short on specifics, other than tax help for wealthy <https://t.co/vGefdXaiDP> <https://t.co/Obh9â> |  
 Tweet 1: Manual Annotation: negative  
 Tweet 1: Lexicon Prediction: positive

Tweet 2: RT @jasonfekete: Great read here: In 2007, Trump was forced to own up to his falsehoods. And he did. Sort of. <https://t.co/yM6JGjrWbD>  
 Tweet 2: Manual Annotation: neutral  
 Tweet 2: Lexicon Prediction: positive

Tweet 3: RT @nytopinion: Forget politics, Trump is a disgusting human being. His children should be ashamed of him. <https://t.co/EErtmVZr4R> <https://â> |  
 Tweet 3: Manual Annotation: negative  
 Tweet 3: Lexicon Prediction: negative

Tweet 4: RT @newtgingrich: Hillary Clinton snubs police union <https://t.co/5b2xCcWg9l>  
 Tweet 4: Manual Annotation: positive  
 Tweet 4: Lexicon Prediction: neutral

Tweet 5: RT @randyprine: Every GOP leader needs to repudiate Trump. <https://t.co/03x0gfClDd>  
 Tweet 5: Manual Annotation: negative  
 Tweet 5: Lexicon Prediction: negative

Tweet 6: RT @hughhewitt: Should b lead story of the AM: <https://t.co/Q2zWG3y1KD>. It isn't b/c of @realDonaldTrump gaffe. Crime/corruption overshadowâ |  
 Tweet 6: Manual Annotation: neutral  
 Tweet 6: Lexicon Prediction: positive

Tweet 7: RT @dredeyedick: If Trump was "joking", why say "& that will be a horrible day."?This was no joke. It was fomenting Insurrection. <https://tâ> |  
 Tweet 7: Manual Annotation: negative  
 Tweet 7: Lexicon Prediction: neutral

Tweet 8: RT @JenniferMayer: Hillary Clintonâ s Embrace of Kissinger Is Inexcusable: "Kissinger is a unique monster" <https://t.co/X3IDNoDISP> @GregGranâ |  
 Tweet 8: Manual Annotation: positive  
 Tweet 8: Lexicon Prediction: neutral

Tweet 9: RT @sanepolitico: Denis O'Brien donates to the Clinton foundation.  
 Tweet 9: Manual Annotation: neutral  
 Tweet 9: Lexicon Prediction: neutral

Tweet 10: RT @OnPointRadio: â Trump's said about 150 things that would disqualify anybody else from running for President.â @GovGaryJohnson <https://t.â> |  
 Tweet 10: Manual Annotation: negative  
 Tweet 10: Lexicon Prediction: neutral

Tweet 11: RT @Dleather25\_: Donald Trump is not what the media portrays him to be, he is an American Patriot with a good heart #MAGA ð ºð , <https://t.co/Pâ> |  
 Tweet 11: Manual Annotation: positive  
 Tweet 11: Lexicon Prediction: positive

Tweet 12: Is NBC News Owned by Hillary Clinton? <https://t.co/1zNlCDiTzx>  
 Tweet 12: Manual Annotation: positive  
 Tweet 12: Lexicon Prediction: neutral

Tweet 13: All By Herself: Kelly Ayotte Feels Heat Back Home As Only New England Senator Supporting Trump: With Sen. Col... <https://t.co/9HXT5pNXdP>  
 Tweet 13: Manual Annotation: neutral  
 Tweet 13: Lexicon Prediction: positive

Tweet 14: RT @ZDFheute: AuÃ enminister Steinmeier wird beim Gedanken an einen US-PrÃsidenten Trump "echt bange". Er sei hier "nicht neutral", sagte seâ |  
 Tweet 14: Manual Annotation: neutral  
 Tweet 14: Lexicon Prediction: positive

Tweet 15: The Clinton Foundation Exposed | Charles Ortel and Stefan Molyneux <https://t.co/tbeEPKPrmf>  
 Tweet 15: Manual Annotation: positive  
 Tweet 15: Lexicon Prediction: neutral

Tweet 16: RT @aipsmith: Newt Gingrich has said that Trump will be "at least as reliable" as Andrew Jackson as Pres. A historian responds: <https://t.câ> |  
 Tweet 16: Manual Annotation: negative  
 Tweet 16: Lexicon Prediction: positive

Tweet 17: TWITTERTORTIAL: GOP women abandoning Trump. After his disgusting non-joke about Hillary/2nd Amendment, why hasn't everyone done the same?  
 Tweet 17: Manual Annotation: negative  
 Tweet 17: Lexicon Prediction: negative

Tweet 18: RT @realDonaldTrump: My thoughts and prayers are with the two police officers shot in Sebastian County, Arkansas. #LESM  
 Tweet 18: Manual Annotation: positive  
 Tweet 18: Lexicon Prediction: neutral

Tweet 19: RT @DavidAFrench: Why would a political party sell its soul to a man who believes only in himself? <https://t.co/FEuureW8wy>  
 Tweet 19: Manual Annotation: negative  
 Tweet 19: Lexicon Prediction: neutral

Tweet 20: But how can you blame someone for doing something "absolutely incredible" @realDonaldTrump? <https://t.co/xnM4ilbRox>  
 Tweet 20: Manual Annotation: positive  
 Tweet 20: Lexicon Prediction: negative

Tweet 21: @RealKidPoker @snagglechud Is trump killing people?  
 Tweet 21: Manual Annotation: negative  
 Tweet 21: Lexicon Prediction: neutral

Tweet 22: @Sandman\_HQ @morgan\_murphy <https://t.co/yxc9BXierN>  
 Tweet 22: Manual Annotation: neutral  
 Tweet 22: Lexicon Prediction: neutral

Tweet 23: RT @Slate: Trump's "Second Amendment people" joke was targeting gun owners, too: <https://t.co/C3PgZ0oD9B>  
 Tweet 23: Manual Annotation: positive  
 Tweet 23: Lexicon Prediction: negative

Tweet 24: RT @sherrilee7: Christians in name only. <https://t.co/tBekEa7G1F>  
 Tweet 24: Manual Annotation: negative  
 Tweet 24: Lexicon Prediction: neutral

Tweet 25: RT @Slate: Donald Trump's child care proposal would be totally useless to families that need it: <https://t.co/nfeZ0fuqDP> <https://t.co/fIKnEa7G1F>  
 Tweet 25: Manual Annotation: negative  
 Tweet 25: Lexicon Prediction: negative

Tweet 26: RT @sanepolitico: The Clinton foundation shares executives etc with Teneo Consulting.  
 Tweet 26: Manual Annotation: neutral  
 Tweet 26: Lexicon Prediction: neutral

Tweet 27: My language already got tightened as to Trump but he's getting worse and worse. #Hillary 2nd amendment subtle hint - play of words hey ?  
 Tweet 27: Manual Annotation: negative  
 Tweet 27: Lexicon Prediction: neutral

Tweet 28: RT @realDonaldTrump: Thank you Wilmington, North Carolina! #MakeAmericaGreatAgain <https://t.co/ZnnaSPF5or>  
 Tweet 28: Manual Annotation: positive  
 Tweet 28: Lexicon Prediction: positive

Tweet 29: RT @fmei7777: Bolsonaro vai odiar a cerimônia. Trump também. Pelo menos nisso acertamos.  
 Tweet 29: Manual Annotation: neutral  
 Tweet 29: Lexicon Prediction: positive

Tweet 30: President Donald J. Trump's first executive order. Welcome to the United States of Ostentacia. <https://t.co/QvD0pEy1rX>  
 Tweet 30: Manual Annotation: negative  
 Tweet 30: Lexicon Prediction: positive

Tweet 31: RT @realDonaldTrump: My thoughts and prayers are with the two police officers shot in Sebastian County, Arkansas. #LESM  
 Tweet 31: Manual Annotation: positive  
 Tweet 31: Lexicon Prediction: neutral

Tweet 32: @Richard\_A\_King @mitchellvii @pnehlen Trump would put a stake through it's heart. Hillary and Ryan can't wait to wake it back up.  
 Tweet 32: Manual Annotation: neutral  
 Tweet 32: Lexicon Prediction: positive

Tweet 33: RT @NYMag: New emails revive concerns about cronyism in Clinton's State Department: <https://t.co/owSNem7IKZ> <https://t.co/viWgkKd0XW>  
 Tweet 33: Manual Annotation: positive  
 Tweet 33: Lexicon Prediction: negative

Tweet 34: RT @NTKRNow: Emails Renew Questions About Clinton Foundation and State Dept. Overlap <https://t.co/jANRAXKgEx>  
 Tweet 34: Manual Annotation: positive  
 Tweet 34: Lexicon Prediction: neutral

Tweet 35: RT @realDonaldTrump: Thank you Wilmington, North Carolina! #MakeAmericaGreatAgain <https://t.co/ZnnaSPF5or>  
 Tweet 35: Manual Annotation: positive  
 Tweet 35: Lexicon Prediction: positive

Tweet 36: RT @ZeddRebel: Trump voter: "I like Trump because he says exactly what he means." Trump: 'Somebody shoot my opponent' Trump voter: 'He didnâ |  
 Tweet 36: Manual Annotation: positive  
 Tweet 36: Lexicon Prediction: positive

Tweet 37: RT @toptc: What happens when you wear a pro-Trump T-shirt in this neighborhood? You get beaten with a cro wbar @Followt96844528 <https://t.co/câ |>  
 Tweet 37: Manual Annotation: neutral  
 Tweet 37: Lexicon Prediction: neutral

Tweet 38: RT @Lrihendry: TRENDING! Eminem Switches to Trump, urges his fans to follow! @realDonaldTrump #TrumpPence 2016 <https://t.co/HK9VgQCYae>  
 Tweet 38: Manual Annotation: positive  
 Tweet 38: Lexicon Prediction: positive

Tweet 39: RT @WayneRoot: This is why I believe Google, Apple & high tech firms are committing fraud to hurt Tru mp and elect Hillary. Biased <https://â |>  
 Tweet 39: Manual Annotation: positive  
 Tweet 39: Lexicon Prediction: negative

Tweet 40: RT @realDonaldTrump: My thoughts and prayers are with the two police officers shot in Sebastian County, A rkansas. #LESM  
 Tweet 40: Manual Annotation: positive  
 Tweet 40: Lexicon Prediction: neutral

Tweet 41: @realDonaldTrump ð ¤ ð ¤ <https://t.co/Erct1jz3t3>  
 Tweet 41: Manual Annotation: neutral  
 Tweet 41: Lexicon Prediction: neutral

Tweet 42: @HillaryClinton what is your economic plan? Our economy is not doing so well.  
 Tweet 42: Manual Annotation: positive  
 Tweet 42: Lexicon Prediction: neutral

Tweet 43: RT @kylegriffin1: Former McCain adviser: Trump is 'much more reckless than Sarah Palin â and that's qui te a standard'. <https://t.co/EIwcFKcâ |>  
 Tweet 43: Manual Annotation: negative  
 Tweet 43: Lexicon Prediction: neutral

Tweet 44: Similar to our NH poll: Trump down by 15, Ayotte down by 10. <https://t.co/S17HvITwpT> <https://t.co/M4hMjH6ls0>  
 Tweet 44: Manual Annotation: neutral  
 Tweet 44: Lexicon Prediction: positive

Tweet 45: â His candidacy is now toxic to democracyâ : â Morning Joeâ panel denounces Trump for Clinton threat <https://t.co/Kx5RE7foRk>  
 Tweet 45: Manual Annotation: negative  
 Tweet 45: Lexicon Prediction: negative

Tweet 46: ð § <https://t.co/YRahRtdpSP>  
 Tweet 46: Manual Annotation: neutral  
 Tweet 46: Lexicon Prediction: neutral

Tweet 47: RT @Phil\_Lewis\_: Harambe didn't die for you to slander his name like this. <https://t.co/0ybcRcVB74>  
 Tweet 47: Manual Annotation: neutral  
 Tweet 47: Lexicon Prediction: negative

Tweet 48: MSM infers even Don will vote "for the novelty girl", and then none of us care who takes it from there if she drops. <https://t.co/cgUE9jdwu0>  
 Tweet 48: Manual Annotation: neutral  
 Tweet 48: Lexicon Prediction: positive

Tweet 49: I wholeheartedly agree with assessment. <https://t.co/T2DppeJbBK>  
 Tweet 49: Manual Annotation: neutral  
 Tweet 49: Lexicon Prediction: positive

Tweet 50: RT @timkaine: David's letter to his daughter about why he can't support Donald Trump really hits home. ht tps://t.co/uyDaDdcpab  
 Tweet 50: Manual Annotation: negative  
 Tweet 50: Lexicon Prediction: positive

Tweet 51: #Trump not being arrested for suggesting #Clinton should be assassinated is just another way rich white g uys get preferential treatment.  
 Tweet 51: Manual Annotation: negative  
 Tweet 51: Lexicon Prediction: positive

Tweet 52: This is amazing! <https://t.co/hHawCi94v>



Tweet 52: Manual Annotation: neutral  
 Tweet 52: Lexicon Prediction: neutral

Tweet 53: RT @hopefulseekr: Emails raise new questions on ties between Clinton Foundation, State Department via the @FoxNews App <https://t.co/5IoCtCqâ> |  
 Tweet 53: Manual Annotation: positive  
 Tweet 53: Lexicon Prediction: neutral

Tweet 54: @seanhannity. Sean you and greta and bill hemmer and oreilly are the only ones on fox that at least cut t rump some slack. The rest are bad.  
 Tweet 54: Manual Annotation: positive  
 Tweet 54: Lexicon Prediction: positive

Tweet 55: RT @politicalgates: Donald Trump knew exactly what he was saying... <https://t.co/cRiVgl2P0Z> <https://t.co/hZ2WcWJpnZ>  
 Tweet 55: Manual Annotation: neutral  
 Tweet 55: Lexicon Prediction: positive

Tweet 56: RT @LindaSuhler: Rudy Giuliani Calls Clinton Corruption "Illegal" #CrookedHillary â i. #TrumpPence16 #MAGA <https://t.co/uV8I4844lU> <https://tâ> |  
 Tweet 56: Manual Annotation: positive  
 Tweet 56: Lexicon Prediction: negative

Tweet 57: RT @Ecnebs: @andersonDrLJA Tired of this stuff coming out & talked about but NOTHING done about it. Clinton & crew slip thru EVERYTHING itsâ |  
 Tweet 57: Manual Annotation: positive  
 Tweet 57: Lexicon Prediction: negative

Tweet 58: Maybe @Wikileaks will help! --& Obama DOJ Refused FBI Request For Full Investigation Into Clinton Foundation <https://t.co/X6mYvXCPFh> #MAGA  
 Tweet 58: Manual Annotation: positive  
 Tweet 58: Lexicon Prediction: negative

Tweet 59: @realDonaldTrump @dbongino @CNN U should be admitted pal. U are one sick puppy.I pray nothing happens to Hillary.#guttersnipe #thug  
 Tweet 59: Manual Annotation: negative  
 Tweet 59: Lexicon Prediction: negative

Tweet 60: RT @pourmecoffee: It's a total mystery how a conspiracy-mongering demagogue like Trump could have taken over the GOP. <https://t.co/FNinZD5uâ> |  
 Tweet 60: Manual Annotation: negative  
 Tweet 60: Lexicon Prediction: positive

Tweet 61: RT @ZeddRebel: Trump voter: "I like Trump because he says exactly what he means." Trump: 'Somebody shoot my opponent' Trump voter: 'He didnâ |  
 Tweet 61: Manual Annotation: negative  
 Tweet 61: Lexicon Prediction: positive

Tweet 62: RT @shaman\_don: 10/10 I approve of this title <https://t.co/2dCqZLur6H>  
 Tweet 62: Manual Annotation: neutral  
 Tweet 62: Lexicon Prediction: positive

Tweet 63: Trump est instoppable malgrÃ toutes les gaffes qu'il peut faire. C'est le bed of nails dÃcrit par John Oliver. /1  
 Tweet 63: Manual Annotation: neutral  
 Tweet 63: Lexicon Prediction: positive

Tweet 64: RT @wikileaks: Clinton campiagn has developed a new PR strategy. All future corruption revelations about her are the work of Russia <https://â> |  
 Tweet 64: Manual Annotation: positive  
 Tweet 64: Lexicon Prediction: neutral

Tweet 65: RT @10thAmendment: #CLINTON LACKEYS and LIB REPORTERS Threatened to KILL JULIAN ASSANGE â But That Was OK? <https://t.co/lyPLxZ8Mvo> #SethRicâ |  
 Tweet 65: Manual Annotation: positive  
 Tweet 65: Lexicon Prediction: negative

Tweet 66: wow really makes you think <https://t.co/xKAbxbpoNb>  
 Tweet 66: Manual Annotation: neutral  
 Tweet 66: Lexicon Prediction: positive

Tweet 67: RT @chuckwoolery: Dan Rather: Media Must Publicly Shame Donald Trump Supporters - Breitbart <https://t.co/JHehtjNs3R> via @BreitbartNews  
 Tweet 67: Manual Annotation: negative  
 Tweet 67: Lexicon Prediction: neutral

Tweet 68: RT @axiome8: Hillary Clinton : Comportement Bizarre et doutes importants sur sa santÃ <https://t.co/JeM3zaFc7o> <https://t.co/16trGuDy5Y>  
 Tweet 68: Manual Annotation: neutral  
 Tweet 68: Lexicon Prediction: negative

Tweet 69: Clinton up 13 in WI among LVs in the 4-way (like it will be in Nov). Among all registered only up 9. <https://t.co/sbubHh6sHQ>  
 Tweet 69: Manual Annotation: neutral  
 Tweet 69: Lexicon Prediction: neutral

Tweet 70: 1 2 3 4 We Declare A \*Trump War <https://t.co/WtHPeSCKt8>  
 Tweet 70: Manual Annotation: positive  
 Tweet 70: Lexicon Prediction: neutral

Tweet 71: RT @brianoflondon: Now this is how you make a proper death threat on live TV. Donâ t mess about with ambiguous â jokesâ . <https://t.co/oWJT5rNâ> |  
 Tweet 71: Manual Annotation: neutral  
 Tweet 71: Lexicon Prediction: negative

Tweet 72: RT @OccuWorld: Trumpâ s vomit is media gold <https://t.co/1zbZPNBVZe>  
 Tweet 72: Manual Annotation: neutral  
 Tweet 72: Lexicon Prediction: neutral

Tweet 73: RT @MeannMike: If there's still a line of human decency that exists in American politics, he crossed it. GOP, what say you? <https://t.co/Sâ> |  
 Tweet 73: Manual Annotation: negative  
 Tweet 73: Lexicon Prediction: positive

Tweet 74: RT @democracynow: â Donald Trump has done business all his life w/ mobsters & criminals because it â s a way to make moneyâ - @DavidCayJ <https://t.co/1zbZPNBVZe> |  
 Tweet 74: Manual Annotation: negative  
 Tweet 74: Lexicon Prediction: positive

Tweet 75: RT @MadanPayal1: @rahulgandhi4in @PMOIndia @narendramodi @mlkhattar @ZeeNews @htTweets One of Election Manifesto should be Home for all @Suâ |  
 Tweet 75: Manual Annotation: neutral  
 Tweet 75: Lexicon Prediction: neutral

Tweet 76: RT @SilerRabbit: #WednesdayWisdom New emails from a FOIA request prove Clinton Foundation #PayToPlay. Just like the #DNCLeaks do. <https://t.co/1zbZPNBVZe> |  
 Tweet 76: Manual Annotation: positive  
 Tweet 76: Lexicon Prediction: positive

Tweet 77: RT @JillyWisdom: Ivanka Trump says she supports working moms. The women who design her fashion... <https://t.co/CqtSnAJRM2> by @ApothecaryJerâ |  
 Tweet 77: Manual Annotation: positive  
 Tweet 77: Lexicon Prediction: positive

Tweet 78: RT @RoyalWJohnson: What Clinton and Trump would do for parents: Ask most working parents what makes family life hard, and you ge... <https://t.co/1zbZPNBVZe> |  
 Tweet 78: Manual Annotation: neutral  
 Tweet 78: Lexicon Prediction: positive

Tweet 79: RT @CounterJihadUS: Did Clinton Emails Get Iranian Scientist Killed? <https://t.co/00gvxNxHLM> <https://t.co/06eMgvZLlh>  
 Tweet 79: Manual Annotation: positive  
 Tweet 79: Lexicon Prediction: neutral

Tweet 80: @PollsAndVotes says there are signs the Bernie Sanders voters are coming back to Dems and Clinton. #mulawpoll  
 Tweet 80: Manual Annotation: neutral  
 Tweet 80: Lexicon Prediction: neutral

Tweet 81: @realDonaldTrump Hindus like you in India & World because U hate Muslims Terror who are creating problems for the world  
 Tweet 81: Manual Annotation: positive  
 Tweet 81: Lexicon Prediction: negative

Tweet 82: RT @davidschneider: Poor Donald Trump. He's almost at the end of his list of things to say and do to get out of becoming President. <https://t.co/1zbZPNBVZe> |  
 Tweet 82: Manual Annotation: negative  
 Tweet 82: Lexicon Prediction: negative

Tweet 83: Trump's Son Said Women "Invite" Harassment and Women Are Tearing Him To Shreds (Tweets) <https://t.co/9ZHYcoJUom>  
 Tweet 83: Manual Annotation: negative  
 Tweet 83: Lexicon Prediction: negative

Tweet 84: RT @Reuters: One in five Republicans want Donald Trump to drop out: Reuters/Ipsos poll. <https://t.co/YD5m0fNzNl> <https://t.co/5hEADYksi>  
 Tweet 84: Manual Annotation: negative  
 Tweet 84: Lexicon Prediction: positive

Tweet 85: @alaskarebel And yet he doesn't call for Trump's med records to see if he has dementia. Yeah, that makes

sense.

Tweet 85: Manual Annotation: negative  
Tweet 85: Lexicon Prediction: neutral

Tweet 86: @tripgabriel @elizabethforma @realDonaldTrump Taunting Trump NOT EVEN CLOSE to his putting target on HRC, or any of the other cr\*p he spews  
Tweet 86: Manual Annotation: negative  
Tweet 86: Lexicon Prediction: negative

Tweet 87: RT @BruceBartlett: Donald Trump's long, long record of lying about almost everything. Documented in testimony under oath. <https://t.co/fZwTâ> |  
Tweet 87: Manual Annotation: negative  
Tweet 87: Lexicon Prediction: negative

Tweet 88: @realDonaldTrump another 2A fan of yours?  
Tweet 88: Manual Annotation: neutral  
Tweet 88: Lexicon Prediction: neutral

Tweet 89: RT @tinaissa: It is absolutely horrifying what Trump said about Clinton Maybe there needed to be more outrage when GOP said this about @POâ |  
Tweet 89: Manual Annotation: negative  
Tweet 89: Lexicon Prediction: negative

Tweet 90: @LaughtTheLamp That observation also matches with my understanding of how and when libertarians backed Trump over some other option  
Tweet 90: Manual Annotation: neutral  
Tweet 90: Lexicon Prediction: positive

Tweet 91: @HillaryClinton Stay strong, stay focused! My admiration for your stamina is boundless!  
Tweet 91: Manual Annotation: negative  
Tweet 91: Lexicon Prediction: positive

Tweet 92: RT @realDonaldTrump: My thoughts and prayers are with the two police officers shot in Sebastian County, Arkansas. #LESM  
Tweet 92: Manual Annotation: positive  
Tweet 92: Lexicon Prediction: neutral

Tweet 93: RT @BigStick2013: How Donald Trump will make It to the White House <https://t.co/XSkWwAus4p>  
Tweet 93: Manual Annotation: neutral  
Tweet 93: Lexicon Prediction: positive

Tweet 94: @realDonaldTrump lmao, I love it when you and your team do damage control, literally every day, stick to a teleprompter ð #hilary4president  
Tweet 94: Manual Annotation: negative  
Tweet 94: Lexicon Prediction: neutral

Tweet 95: @realDonaldTrump Is this something Presidential material should be concerned with?  
Tweet 95: Manual Annotation: negative  
Tweet 95: Lexicon Prediction: negative

Tweet 96: RT @ZeddRebel: Trump voter: "I like Trump because he says exactly what he means." Trump: 'Somebody shoot my opponent' Trump voter: 'He didnâ |  
Tweet 96: Manual Annotation: negative  
Tweet 96: Lexicon Prediction: positive

Tweet 97: New! Daily News Says Trump Should End His Campaign <https://t.co/aQ7bV8cNAX>  
Tweet 97: Manual Annotation: negative  
Tweet 97: Lexicon Prediction: positive

Tweet 98: @CAIRNational LEAVE TRUMP ALONE! <https://t.co/vI59fJXN8b>  
Tweet 98: Manual Annotation: positive  
Tweet 98: Lexicon Prediction: positive

Tweet 99: RT @jacqui4peace: @LeChatNoire4 @Figment\_Imagine You are exactly right LeChatNoir, 3rd party loons ONLY go after her, NEVER after TRUMP, itâ |  
Tweet 99: Manual Annotation: negative  
Tweet 99: Lexicon Prediction: positive

Tweet 100: RT @foxandfriends: .@TomFitton: Newly uncovered emails show the Clinton Foundation working with the State Department to get favors done httpâ |  
Tweet 100: Manual Annotation: positive  
Tweet 100: Lexicon Prediction: neutral

```
In [6]: print("Accuracy: {}".format(accuracy_count/num_tweets))
```

Accuracy: 0.33

## Machine Learning

For this step, load the training and test sentiment datasets. The data is loaded into 4 lists of strings: X\_txt\_train, X\_txt\_test, y\_test, y\_train.

```
In [7]: X_txt_train = []
        y_train = []

        X_txt_test = []
        y_test = []

        X_txt_new = []

        with open('trump-2016-train.csv', encoding='iso-8859-1') as file_handle:
            my_csv = csv.reader(file_handle, delimiter = ',')
            for row in my_csv:
                X_txt_train.append(row[3])
                y_train.append(row[4])

        with open('trump-2016-test.csv', encoding='iso-8859-1') as file_handle:
            my_csv = csv.reader(file_handle, delimiter = ',')
            for row in my_csv:
                X_txt_test.append(row[3])
                y_test.append(row[4])

        with open('trump-2016-new.csv', encoding='iso-8859-1') as file_handle:
            my_csv = csv.reader(file_handle, delimiter = ',')
            for row in my_csv:
                X_txt_new.append(row[3])
```

```
In [8]: assert(type(X_txt_train) == type(list()))
        assert(type(X_txt_train[0]) == type(str()))
        assert(type(X_txt_test) == type(list()))
        assert(type(X_txt_test[0]) == type(str()))
        assert(type(y_test) == type(list()))
        assert(type(y_train) == type(list()))
        assert(len(X_txt_test) == 20)
        assert(len(y_test) == 20)
        assert(len(X_txt_train) == 80)
        assert(len(y_train) == 80)
        assert(len(X_txt_new) == 647548)
        print("Asserts Completed Successfully!")
```

Asserts Completed Successfully!

Note to self:

- For this task, create a feature matrix using CountVectorizer and train a LinearSVC model from scikit-learn
- On the train split, use GridSearchCV to find the best LinearSVC C values (0.0001, 0.001, 0.001, 0.01, 0.1, 1, 10, or 100) based on the micro f1 scoring metric (hint: "micro" average) and set the cv parameter to 5
- Also, with the CountVectorizer, only use unigrams (i.e., set ngram\_range = (1,1))
- Note that GridSearchCV will retrain the final classifier using the best parameters, so you don't need to do it manually.

```
In [9]: from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.svm import LinearSVC
from sklearn.metrics import accuracy_score
from sklearn.model_selection import GridSearchCV
from sklearn.metrics import precision_score, recall_score, f1_score

import numpy as np
np.random.seed(42)
import random
random.seed(42)

# WRITE CODE HERE

vec = CountVectorizer(ngram_range=(1,1))

X_train = vec.fit_transform(X_txt_train)
X_test = vec.transform(X_txt_test)

params = {'C':[0.0001, 0.001, 0.01, 0.1, 1, 10, 100]}

svc = LinearSVC()

clf = GridSearchCV(svc, params, scoring='f1_micro', cv=5)

clf.fit(X_train, y_train)

validation_score = clf.best_score_
print("Validation F1: {:.4f}".format(validation_score))

svm_test_predictions = clf.predict(X_test)

accuracy = accuracy_score(y_test, svm_test_predictions)

print("Accuracy: {:.4f}".format(accuracy))

Validation F1: 0.5250
Accuracy: 0.6500

C:\Users\lee\anaconda3\lib\site-packages\sklearn\svm\_base.py:976: ConvergenceWarning: Liblinear failed to converge, increase the number of iterations.
  warnings.warn("Liblinear failed to converge, increase "
```

```
In [10]: from scipy.sparse import csr_matrix
assert(type(X_train) == type(csr_matrix(0)) or type(X_train) == type(np.array(0)))
assert(type(X_test) == type(csr_matrix(0)) or type(X_test) == type(np.array(0)))
assert(X_train.shape[0] == len(X_txt_train))
assert(X_test.shape[0] == len(X_txt_test))
assert(X_train.shape[1] == X_test.shape[1])
assert(type(accuracy) == type(float()) or type(accuracy) == type(np.float64()))
print("Asserts Completed Successfully!")

Asserts Completed Successfully!
```

```
In [11]: X_new = vec.transform(X_txt_new)
svm_new_predictions = clf.predict(X_new)
```

```
In [12]: assert(X_new.shape[0] == len(X_txt_new))
print("Asserts Completed Successfully!")

Asserts Completed Successfully!
```

```
In [13]: y_predict = svm_new_predictions.tolist()

total_number_of_positive_tweets_2016_m1 = 0
total_number_of_negative_tweets_2016_m1 = 0
total_number_of_neutral_tweets_2016_m1 = 0

for item in y_predict:
    if item == "positive":
        total_number_of_positive_tweets_2016_m1 += 1
    elif item == "negative":
        total_number_of_negative_tweets_2016_m1 += 1
    elif item == "neutral":
        total_number_of_neutral_tweets_2016_m1 += 1
```

```
In [14]: assert(len(y_predict)==total_number_of_positive_tweets_2016_ml+total_number_of_negative_tweets_2016_ml+total_number_of_neutral_tweets_2016_ml)
print("Asserts Completed Successfully!")
```

Asserts Completed Successfully!

```
In [15]: print("Total Number of Tweets: {}".format(len(y_predict)))
print("Total Number of Positive Tweets: {}".format(total_number_of_positive_tweets_2016_ml))
print("Total Number of Negative Tweets: {}".format(total_number_of_negative_tweets_2016_ml))
print("Total Number of Neutral Tweets: {}".format(total_number_of_neutral_tweets_2016_ml))
```

Total Number of Tweets: 647548  
Total Number of Positive Tweets: 198030  
Total Number of Negative Tweets: 231600  
Total Number of Neutral Tweets: 217918

**2020**

```

In [16]: %%time

import csv

# define variables

lc = LexiconClassifier() # initiate class

total_number_of_tweets_2020 = 0 #1
total_number_of_positive_tweets_2020 = 0 #2
total_number_of_negative_tweets_2020 = 0 #2

user_tweets_2020 = {} # dict of screen_name: # of tweets
user_with_most_tweets_2020 = '' #3
tweets_by_biggest_user_2020 = 0 #3
total_number_of_users_2020 = 0 #4
average_number_tweets_per_user_2020 = 0 #5

most_positive_tweet_2020 = '' #6
most_negative_tweet_2020 = '' #6
most_positive_tweet_score_2020 = 0 #6
most_negative_tweet_score_2020 = 0 #6

count = 0 # skip first 100 tweets because machine learning will train/ test on this

# run code

twitter_dataset = open('trump-2020.csv', 'r', encoding='utf8')
tweet_string = csv.reader(twitter_dataset, delimiter=',')

for row in tweet_string:

    count += 1

    if count > 100:

        # extract tweet and screen_name, update dict, increment

        tweet_text = row[3]
        screen_name = row[1]

        user_tweets_2020[screen_name] = user_tweets_2020.get(screen_name, 0) + 1

        total_number_of_tweets_2020 += 1

        # count pos and neg words, calculate sentiment score and prediction

        num_pos_words = lc.count_pos_words(tweet_text)
        num_neg_words = lc.count_neg_words(tweet_text)

        sentiment_score = num_pos_words - num_neg_words
        sentiment_prediction = lc.predict(tweet_text)

        # record most positive/negative tweets/users

        if sentiment_score > most_positive_tweet_score_2020:
            most_positive_tweet_2020 = tweet_text
            most_positive_tweet_score_2020 = sentiment_score

        elif sentiment_score < most_negative_tweet_score_2020:
            most_negative_tweet_2020 = tweet_text
            most_negative_tweet_score_2020 = sentiment_score

        # increment positive/ negative tweets

        if sentiment_prediction == 'positive':
            total_number_of_positive_tweets_2020 += 1

        elif sentiment_prediction == 'negative':
            total_number_of_negative_tweets_2020 += 1

#debug
# print(screen_name, tweet_text)
# if total_number_of_tweets_2020 > 10:

```

Wall time: 1min 26s

```
Total Number of Tweets: 3053797
Total Number of Positive Tweets: 1176515
Total Number of Negative Tweets: 623914
Total Number of Neutral Tweets: 1253368

Most Positive Tweet
@GOP GOOD GREAT RIGHT DECISION GOOD GREAT RIGHT DECISION GOOD GREAT RIGHT DECISION GOOD GREAT RIGHT DECISION GOOD G
REAT RIGHT DECISION GOOD GREAT RIGHT DECISION GOOD GREAT RIGHT DECISION GOOD GREAT RIGHT DECISION GOOD GREAT RIGHT
DECISION GOOD GREAT RIGHT DECISION
30

Most Negative Tweet
@realDonaldTrump @OANN Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie
Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie
Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie Lie
-70

Total Number of Users: 1005918
Average Number of Tweets per User: 3.0358309524235576
User with the most tweets: tsartbot (444)
```

Assert finished successfully!

Perform manual analysis of the predictions. Modify the code to load the tweet text.



```
In [19]: twitter_dataset = open('trump-2020-rate.csv', 'r', encoding='iso-8859-1')
tweet_string = csv.reader(twitter_dataset, delimiter=',')

num_tweets_to_print = 100
num_tweets = 0
accuracy_count = 0

for row in tweet_string:

    num_tweets += 1
    tweet_text = row[3]
    manual_annotation = row[4]

    num_pos_words = lc.count_pos_words(tweet_text)
    num_neg_words = lc.count_neg_words(tweet_text)

    sentiment_score = num_pos_words - num_neg_words
    sentiment_prediction = lc.predict(tweet_text)

    if manual_annotation == sentiment_prediction:
        accuracy_count += 1

    print("Tweet {}: {}".format(num_tweets, tweet_text))
    print("Tweet {}: Manual Annotation: {}".format(num_tweets, manual_annotation))
    print("Tweet {}: Lexicon Prediction: {}".format(num_tweets, sentiment_prediction))
    print()

    if num_tweets == num_tweets_to_print:
        break

twitter_dataset.close()
```

Tweet 1: RT @GOPChairwoman: The contrast between @realDonaldTrump's #AmericaFirst trade policies and Biden's 40 years of job-killing policies couldâ |  
 Tweet 1: Manual Annotation: positive  
 Tweet 1: Lexicon Prediction: negative

Tweet 2: RT @MattGertz: Here's the difference between Fox's news and an opinion side. The opinion side speculates Biden has dementia; what makesâ |  
 Tweet 2: Manual Annotation: neutral  
 Tweet 2: Lexicon Prediction: neutral

Tweet 3: "Arizona reports record spike in new coronavirus cases & deaths ahead of Pence's visit" His megachurch stunt took place 5 days ago in Phoenix. <https://t.co/Z3YmMzqore>  
 Tweet 3: Manual Annotation: negative  
 Tweet 3: Lexicon Prediction: negative

Tweet 4: RT @LetItRain233: Joe Biden help a press conference. Did media ask him about his involvement in framing General Flynn? #fakenews  
 Tweet 4: Manual Annotation: positive  
 Tweet 4: Lexicon Prediction: neutral

Tweet 5: RT @Mike\_Pence: Today marks a HISTORIC day for our Nation! On January 29, 2020 President @realDonaldTrump signed the #USMCA into law and toâ |  
 Tweet 5: Manual Annotation: positive  
 Tweet 5: Lexicon Prediction: neutral

Tweet 6: RT @AndrewBatesNC: "Who's the absent candidate now?... 'Sleepy Joe' is signaling that he's very much awake -- and dialed into a moment whereâ |  
 Tweet 6: Manual Annotation: negative  
 Tweet 6: Lexicon Prediction: neutral

Tweet 7: @LangeTurid @USMSS99 @rightNtruthMat @realDonaldTrump Debt to the world bank? No... But speaking of debt, Trump blew up the deficit in a "great" economy  
 Tweet 7: Manual Annotation: negative  
 Tweet 7: Lexicon Prediction: neutral

Tweet 8: @tekstone @voltage\_333 @HowieHawkins @SPofUSA Do you not understand how massive of an opportunity 2020 for a third party such as @GreenPartyUS to attract literally millions of voters? The competition will never be this easy again.  
 Tweet 8: Manual Annotation: negative  
 Tweet 8: Lexicon Prediction: positive

Tweet 9: @Jorgensen4POTUS @RealSpikeCohen Just found out about you and so far I love your policies and what you have to say. Ofc I still need to read more about your stances and plans but Iâm very impressed with what Iâm hearing (especially your social security)  
 Tweet 9: Manual Annotation: neutral  
 Tweet 9: Lexicon Prediction: positive

Tweet 10: RT @HKrassenstein: BREAKING: The Taliban has just admitted to accepting money from Russia to kill U.S. Troops. This is not a "HOAX." Vlaâ |  
 Tweet 10: Manual Annotation: neutral  
 Tweet 10: Lexicon Prediction: negative

Tweet 11: RT @IngrahamAngle: Biden will move quickly to reverse this if elected. His donors will resume helping China build their military machine aâ |  
 Tweet 11: Manual Annotation: negative  
 Tweet 11: Lexicon Prediction: positive

Tweet 12: @95YSIUxcGfGGXnj @VijayNeupane1 @AngelaBelcamino @realDonaldTrump Dude, get a grip  
 Tweet 12: Manual Annotation: negative  
 Tweet 12: Lexicon Prediction: neutral

Tweet 13: RT @ILACCommissioner: @ShawnFinchum @jeremymbarr Hannity asked What's your plans for a second term? Trump said : "Djwo sjeb is sue us euwâ |  
 Tweet 13: Manual Annotation: neutral  
 Tweet 13: Lexicon Prediction: neutral

Tweet 14: RT @DavidCornDC: So....it wasn't a "hoax"? If Trump is calling a threat to US forces a "hoax," then O'Brien and others ought to resign. Serâ |  
 Tweet 14: Manual Annotation: neutral  
 Tweet 14: Lexicon Prediction: neutral

Tweet 15: RT @brett\_mcgurk: That we knew Putin may have a policy to target our troops while Trump is regularly on the phone and inviting Putin back tâ |  
 Tweet 15: Manual Annotation: negative  
 Tweet 15: Lexicon Prediction: positive

Tweet 16: @ExcaliberXX @HormanTodd @realDonaldTrump After murders of young people?  
 Tweet 16: Manual Annotation: negative  
 Tweet 16: Lexicon Prediction: neutral

Tweet 17: RT @tedlieu: Dear @realDonaldTrump: Why would your Administration brief NATO that Russia was targeting coalition & US troops in Afghanistanâ |  
 Tweet 17: Manual Annotation: negative  
 Tweet 17: Lexicon Prediction: neutral

Tweet 18: RT @BrianKarem: 128K deaths later . . . these words will haunt @realDonaldTrump forever <https://t.co/VGqEiBArB1>  
 Tweet 18: Manual Annotation: negative  
 Tweet 18: Lexicon Prediction: negative

Tweet 19: RT @3sidedstory: The Great Awakening is happening. #MAGA2020  
 Tweet 19: Manual Annotation: positive  
 Tweet 19: Lexicon Prediction: positive

Tweet 20: RT @Kris\_Sacrebleu: Meet the Supporters Trump Has Lost - The New York Times â I will never vote for another Republican in my life because oâ |  
 Tweet 20: Manual Annotation: negative  
 Tweet 20: Lexicon Prediction: neutral

Tweet 21: RT @ThePlumLineGS: Fox News is really all in on the Biden "cognitive decline" smear. @MattGertz has a good deconstruction of the scam: <httpâ> |  
 Tweet 21: Manual Annotation: negative  
 Tweet 21: Lexicon Prediction: positive

Tweet 22: RT @CaslerNoel: Since they now claim you read Donnie, check this out please. @realDonaldTrump <https://t.co/yOCqsVsdu2>  
 Tweet 22: Manual Annotation: neutral  
 Tweet 22: Lexicon Prediction: neutral

Tweet 23: RT @cuteskrys: parabÃns vc passou para o nÃvel 7, veja tudo que jÃ aconteceu -quase terceira guerra mundial -covid -bolsonaro e trump -ovniâ |  
 Tweet 23: Manual Annotation: neutral  
 Tweet 23: Lexicon Prediction: positive

Tweet 24: RT @evilmidget1: @realDonaldTrump Aren't you getting dizzy from all the "spinning" you're doing? Face it. You're a traitor to our countryâ |  
 Tweet 24: Manual Annotation: negative  
 Tweet 24: Lexicon Prediction: negative

Tweet 25: RT @JohnJHarwood: US Rep and Iraq War vet Seth Moulton on Trump/Russian bounties: â if this is not treasonous behavior, I donâ t know what iâ |  
 Tweet 25: Manual Annotation: negative  
 Tweet 25: Lexicon Prediction: negative

Tweet 26: RT @AMoneyResists: @realDonaldTrump Jr Someone sounds triggered even more than usual today. And yes, @JoeBiden knows how to tweet, he would jusâ |  
 Tweet 26: Manual Annotation: negative  
 Tweet 26: Lexicon Prediction: neutral

Tweet 27: RT @Independent: Indigenous tribes condemn Trumpâ s border wall destruction as president protects Confederate monuments <https://t.co/Wor3v3Aâ> |  
 Tweet 27: Manual Annotation: negative  
 Tweet 27: Lexicon Prediction: negative

Tweet 28: RT @amvetsupport: That moron trump vows to veto the Defense Bill if it includes renaming bases. So once again, military salaries and defensâ |  
 Tweet 28: Manual Annotation: negative  
 Tweet 28: Lexicon Prediction: neutral

Tweet 29: RT @tedlieu: Dear @realDonaldTrump: Why would your Administration brief NATO that Russia was targeting coalition & US troops in Afghanistanâ |  
 Tweet 29: Manual Annotation: negative  
 Tweet 29: Lexicon Prediction: neutral

Tweet 30: @HILITINGHOCKEY @mini\_bubbly @Trump\_Detester @RobinADupuis @RandomOntarion @NorthBelle4 @Scribulatora @Fryman401 @WaytowichNeil @Learning2GetBy @CoastalCndn @CanadensisMax @kaythird @earthriset NB kept twisting my posts, so I told her if she didn't stop  
 Tweet 30: Manual Annotation: neutral  
 Tweet 30: Lexicon Prediction: neutral

Tweet 31: RT @NikkiMcR: Twitch wants to "make sure that Trump is the same as SimpLord69" -@Slasher in an instantly iconic moment on Fox Business <httpâ> |  
 Tweet 31: Manual Annotation: negative  
 Tweet 31: Lexicon Prediction: positive

Tweet 32: RT @NY\_runaway: The Vietnam War was a war for Americaâ s economy at the cost of many thousands of American lives. Today we knowingly allowâ |  
 Tweet 32: Manual Annotation: negative  
 Tweet 32: Lexicon Prediction: neutral

Tweet 33: RT @mmpadellan: In addition to painting "Black Lives Matter" on the street in front of Trump Tower, they should also change the name of 5thâ |  
 Tweet 33: Manual Annotation: negative  
 Tweet 33: Lexicon Prediction: positive

Tweet 34: @michaelawat @NYCMayor @realDonaldTrump You support BLM burning down our black community in St. Paul? Thousands of blacks lost their jobs! You are supporting BLM a terrorist organization! Look what they did to us! <http://t.co/m9kZhliKB1>  
 Tweet 34: Manual Annotation: negative  
 Tweet 34: Lexicon Prediction: neutral

Tweet 35: RT @thebradfordfile: I donâ t feel sorry for Joe Biden. His family is 100% responsible for making him a national punchline.  
 Tweet 35: Manual Annotation: positive  
 Tweet 35: Lexicon Prediction: negative

Tweet 36: RT @catturd2: @HillaryClinton @JoeBiden He canâ t tweet at all.  
 Tweet 36: Manual Annotation: positive  
 Tweet 36: Lexicon Prediction: neutral

Tweet 37: RT @libshipwreck: Never Trump Republicans may make some barbed quips and their organizations might make some cutting adsâ but they are not yâ |  
 Tweet 37: Manual Annotation: neutral  
 Tweet 37: Lexicon Prediction: positive

Tweet 38: RT @tedlieu: Dear @realDonaldTrump: Why would your Administration brief NATO that Russia was targeting coalition & US troops in Afghanistanâ |  
 Tweet 38: Manual Annotation: negative  
 Tweet 38: Lexicon Prediction: neutral

Tweet 39: @JoeBiden So you're talking OBAMACARE FOR ALL MEDICAIRE  
 Tweet 39: Manual Annotation: positive  
 Tweet 39: Lexicon Prediction: neutral

Tweet 40: RT @SawyerHackett: Letâ s assume Trump didnâ t read the intelligence reports.... That means he learned about it on Friday, and still playedâ |  
 Tweet 40: Manual Annotation: negative  
 Tweet 40: Lexicon Prediction: positive

Tweet 41: RT @realDonaldTrump: â No corroborating evidence to back reports.â Department of Defense. Do people still not understand that this is all aâ |  
 Tweet 41: Manual Annotation: positive  
 Tweet 41: Lexicon Prediction: neutral

Tweet 42: @realDonaldTrump <https://t.co/ahsLpboGHj>  
 Tweet 42: Manual Annotation: neutral  
 Tweet 42: Lexicon Prediction: neutral

Tweet 43: RT @washingtonpost: Opinion: Trumpâ s ugly new spin on Russia bounties actually incriminates him <http://t.co/IbEieC3GJn>  
 Tweet 43: Manual Annotation: negative  
 Tweet 43: Lexicon Prediction: negative

Tweet 44: RT @sara\_hi\_c: Donald trump says if we stop testing we wonâ t have cases ... so if I stop looking at my bills I donâ t have them anymore ? ð  
 Tweet 44: Manual Annotation: negative  
 Tweet 44: Lexicon Prediction: positive

Tweet 45: RT @W7VOA: According to @NYGovCuomo, @POTUS should 'come clean with the American people' about the #coronavirus threat, and "admit you wereâ |  
 Tweet 45: Manual Annotation: negative  
 Tweet 45: Lexicon Prediction: positive

Tweet 46: RT @washingtonpost: Opinion: Trumpâ s ugly new spin on Russia bounties actually incriminates him <http://t.co/IbEieC3GJn>  
 Tweet 46: Manual Annotation: negative  
 Tweet 46: Lexicon Prediction: negative

Tweet 47: @ClareTyne @mesainy @HKrassenstein @realDonaldTrump @NYCMayor @JoeBiden Yep. Torturing that guy because the Dems donâ t care.  
 Tweet 47: Manual Annotation: positive  
 Tweet 47: Lexicon Prediction: negative

Tweet 48: RT @GOPChairwoman: Joe Biden has fully embraced the radical leftâ s agenda: He supports Green New Deal policies that would destroy millionsâ |  
 Tweet 48: Manual Annotation: positive  
 Tweet 48: Lexicon Prediction: negative

Tweet 49: @JoeBiden Debate President Trump. PROVE you donâ t have dementia. #DementiaJoeCantDebate #JoeBidenScared

ToDebate

Tweet 49: Manual Annotation: positive

Tweet 49: Lexicon Prediction: neutral

Tweet 50: RT @SenDuckworth: Trump's message to our troops: Sure I may be selling you out to Putin, but I'm also threatening to veto your pay raise juâ |

Tweet 50: Manual Annotation: negative

Tweet 50: Lexicon Prediction: negative

Tweet 51: @jonathanvswan @jaketapper The Impeached @realDonaldTrump will always blame others.

Tweet 51: Manual Annotation: negative

Tweet 51: Lexicon Prediction: negative

Tweet 52: @LouisHorvath7 @realDonaldTrump @NYCMayor Not renting to people of color was the trump business policy. They thought theyâ d get away with it but the Feds fined them.

Tweet 52: Manual Annotation: negative

Tweet 52: Lexicon Prediction: positive

Tweet 53: RT @itsJeffTiedrich: @realDonaldTrump @NYCMayor why should a Black Lives Matter sign antagonize police? the only people antagonized by Blacâ |

Tweet 53: Manual Annotation: negative

Tweet 53: Lexicon Prediction: negative

Tweet 54: RT @realDonaldTrump: â No corroborating evidence to back reports.â Department of Defense. Do people still not understand that this is all aâ |

Tweet 54: Manual Annotation: positive

Tweet 54: Lexicon Prediction: neutral

Tweet 55: RT @GeorgeTakei: Trump needs to stop protecting dead Confederate soldiers and start protecting living American ones.

Tweet 55: Manual Annotation: negative

Tweet 55: Lexicon Prediction: neutral

Tweet 56: RT @kylegriffin1: As the pandemic cripples parts of America, Politico reports that four senior Trump officials â Jared Kushner, Mark Meadowsâ |

Tweet 56: Manual Annotation: negative

Tweet 56: Lexicon Prediction: neutral

Tweet 57: RT @ProjectLincoln: When will President Trump condemn Putin? Why won't President Trump simply say he will defend our troops from Russia aâ |

Tweet 57: Manual Annotation: negative

Tweet 57: Lexicon Prediction: positive

Tweet 58: RT @HKrassenstein: BREAKING: The Taliban has just admitted to accepting money from Russia to kill U.S. Troops. This is not a "HOAX." Vlaâ |

Tweet 58: Manual Annotation: negative

Tweet 58: Lexicon Prediction: negative

Tweet 59: @debbymcmanus @mmpadellan @HillaryClinton @JoeBiden Their best not being the gang!!!

Tweet 59: Manual Annotation: positive

Tweet 59: Lexicon Prediction: positive

Tweet 60: @realDonaldTrump You disrespected our soldiers in Afghanistan so your not and better.

Tweet 60: Manual Annotation: negative

Tweet 60: Lexicon Prediction: neutral

Tweet 61: RT @robreiner: Time for Donald Trump to exit the White House and enter Prison.

Tweet 61: Manual Annotation: negative

Tweet 61: Lexicon Prediction: positive

Tweet 62: @SykesCharlie Thereâ s no way trump is doing better in Wisconsin than Michigan. Doesnâ t make any sense. Might not be the best polling

Tweet 62: Manual Annotation: negative

Tweet 62: Lexicon Prediction: positive

Tweet 63: RT @tedlieu: Dear @realDonaldTrump: Why would your Administration brief NATO that Russia was targeting coalition & US troops in Afghanistanâ |

Tweet 63: Manual Annotation: negative

Tweet 63: Lexicon Prediction: neutral

Tweet 64: Newt Gingrich: Left-wing media using polls to hurt Trump and Republicans <https://t.co/ryUDoSvYcp> #FoxNews

Tweet 64: Manual Annotation: positive

Tweet 64: Lexicon Prediction: neutral

Tweet 65: RT @realDonaldTrump: Canâ t believe how badly @CNN has done in the newly released TV ratings. They are so far below @FoxNews (thank you Presâ |

Tweet 65: Manual Annotation: positive

Tweet 65: Lexicon Prediction: negative

Tweet 66: @realDonaldTrump #FakeNews #FakePresident

Tweet 66: Manual Annotation: negative  
Tweet 66: Lexicon Prediction: neutral

Tweet 67: @jonathanvswan This is like the 4th or 5th time this story has come out since Trump took office.  
Tweet 67: Manual Annotation: neutral  
Tweet 67: Lexicon Prediction: positive

Tweet 68: RT @nolean123: @mmpadellan So every piece of mail Trump receives at Trump Tower will have Obama's name on it. Perfect  
Tweet 68: Manual Annotation: negative  
Tweet 68: Lexicon Prediction: positive

Tweet 69: RT @martin\_kj: There is no confirmation that Russia has stopped paying a bounty for dead American troops and Trump still has not condemned it  
Tweet 69: Manual Annotation: negative  
Tweet 69: Lexicon Prediction: neutral

Tweet 70: @realDonaldTrump Can't wait until Putin makes "Trump" your maiden name.  
Tweet 70: Manual Annotation: negative  
Tweet 70: Lexicon Prediction: neutral

Tweet 71: @PamelaStovall6 @ChuckGrassley @realDonaldTrump Democrats are always saving @GOP dinosaurs like grassley after they fvck up  
Tweet 71: Manual Annotation: positive  
Tweet 71: Lexicon Prediction: positive

Tweet 72: @DebraMessing @NYGovCuomo <https://t.co/A5Dv1CV0pw> Conservative media has falsely spread misinformation. They have smeared the name of the free press and they're complicit to treason with Donald Trump.  
Tweet 72: Manual Annotation: negative  
Tweet 72: Lexicon Prediction: negative

Tweet 73: RT @marklevinshow: 1. Funny how Biden barely criticizes Putin, and on the China virus doesn't criticize Xi. Instead, he smears Trump. That's  
Tweet 73: Manual Annotation: positive  
Tweet 73: Lexicon Prediction: negative

Tweet 74: @maier\_amsden @HeerJeet That's totally fair, but at the very least, Trump is surrounded by people telling him exactly how to solve this and countries that have successfully done it. Whereas everyone in the 2000's seemed to struggle with how to exit the war  
Tweet 74: Manual Annotation: negative  
Tweet 74: Lexicon Prediction: positive

Tweet 75: RT @ajplus: 400+ companies will boycott Facebook ads starting today over its failure to tackle hate speech and making exceptions for President Trump  
Tweet 75: Manual Annotation: negative  
Tweet 75: Lexicon Prediction: negative

Tweet 76: RT @sara\_hi\_c: Donald trump says if we stop testing we won't have cases ... so if I stop looking at my bills I don't have them anymore ?  
Tweet 76: Manual Annotation: negative  
Tweet 76: Lexicon Prediction: positive

Tweet 77: @FlyThaiMMA @0rf @skenigsberg @JoeBiden @BernieSanders welcome to politics in the tulip age of your crumbling empire  
Tweet 77: Manual Annotation: positive  
Tweet 77: Lexicon Prediction: neutral

Tweet 78: RT @TeaPainUSA: The real headline is Trump is attackin' his Intelligence Agencies while defendin' Putin. Again.  
Tweet 78: Manual Annotation: negative  
Tweet 78: Lexicon Prediction: positive

Tweet 79: RT @brhodes: Trump will dishonestly throw anyone under the bus to protect himself, even our troops and the intelligence community that helped him  
Tweet 79: Manual Annotation: negative  
Tweet 79: Lexicon Prediction: positive

Tweet 80: RT @CCM1956: Nancy Pelosi says we cannot accept a second term for President Trump. It's ITS NOT FOR YOU TO DECIDE. Luckily, for #WeThePeople  
Tweet 80: Manual Annotation: positive  
Tweet 80: Lexicon Prediction: neutral

Tweet 81: RT @JohnFugelsang: (1). Here's a thread for anyone who's been conned into believing Donald Trump cares about the US military or our troops  
Tweet 81: Manual Annotation: negative  
Tweet 81: Lexicon Prediction: positive

Tweet 82: RT @starsandstripes: .@realDonaldTrump criticized @NYCMayor Bill de Blasio's plan to paint a Black Lives Matter mural on the street outside  
Tweet 82: Manual Annotation: negative

Tweet 82: Lexicon Prediction: negative

Tweet 83: Trump said "experience" was his edge over Joe Biden. A new poll shows voters disagree by a remarkable margin. <https://t.co/Zug0sTGUUr>

Tweet 83: Manual Annotation: negative

Tweet 83: Lexicon Prediction: positive

Tweet 84: @realDonaldTrump #BLM and @Billde Blasio share a hatred of Jews and love for Marxist theology. They are perfect together.

Tweet 84: Manual Annotation: negative

Tweet 84: Lexicon Prediction: positive

Tweet 85: RT @HillaryClinton: If elected president, you can bet @joebiden will never tweet about his TV ratings.

Tweet 85: Manual Annotation: negative

Tweet 85: Lexicon Prediction: neutral

Tweet 86: RT @readylets go27: Business Group Complains Trump H-1B Reform Boosting U.S. Graduates. Big tech is whining that they will have to hire Americans!

Tweet 86: Manual Annotation: positive

Tweet 86: Lexicon Prediction: neutral

Tweet 87: RT @HillaryClinton: If elected president, you can bet @joebiden will never tweet about his TV ratings.

Tweet 87: Manual Annotation: negative

Tweet 87: Lexicon Prediction: neutral

Tweet 88: RT @brhodes: Every political appointee who works for Trump and remains silent is endorsing and enabling the notion that Russian bounties on Americans!

Tweet 88: Manual Annotation: negative

Tweet 88: Lexicon Prediction: positive

Tweet 89: RT @evanmcMurry: ".@NYGovCuomo to Pres. Trump: "Come clean with the American people...At least have the courage to admit what everybody else!"

Tweet 89: Manual Annotation: negative

Tweet 89: Lexicon Prediction: positive

Tweet 90: RT @NotDevinsMom: Someone tell Trump he won the Nobel Peace Prize and he has to go to Iran to pick it up.

Tweet 90: Manual Annotation: negative

Tweet 90: Lexicon Prediction: positive

Tweet 91: RT @no\_silenced: How many Blankets can AOC buy for the poor migrant Children at the Border with that Black Lives Matters Graffiti De Blasio!

Tweet 91: Manual Annotation: positive

Tweet 91: Lexicon Prediction: negative

Tweet 92: RT @politicususa: Hundreds of former officials who served under George W. Bush are about to endorse Joe Biden for president. <https://t.co/â>

Tweet 92: Manual Annotation: negative

Tweet 92: Lexicon Prediction: positive

Tweet 93: RT @realDonaldTrump: "...horrible BLM chant, 'Pigs In A Blanket, Fry 'Em Like Bacon'. Maybe our GREAT Police, who have been neutralized and!"

Tweet 93: Manual Annotation: neutral

Tweet 93: Lexicon Prediction: positive

Tweet 94: RT @EricLevitz: On the one hand, it is possible that the polls will narrow between now and November. On the other hand, Trump's takeaway fails!

Tweet 94: Manual Annotation: neutral

Tweet 94: Lexicon Prediction: neutral

Tweet 95: RT @TheRickWilson: After all the ways they've debased themselves, what's a little treason thrown into the mix? Senate Republicans defend that!

Tweet 95: Manual Annotation: negative

Tweet 95: Lexicon Prediction: negative

Tweet 96: RT @danpfeiffer: Right now, Biden is demonstrating the fatal flaw in the Trump strategy to paint Biden as senile. Trump has not been this childish!

Tweet 96: Manual Annotation: negative

Tweet 96: Lexicon Prediction: positive

Tweet 97: RT @realDonaldTrump: My Executive Order to protect Monuments, Statues etc., IS IN FULL FORCE AND EFFECT. In excess of a 10 year prison term!

Tweet 97: Manual Annotation: neutral

Tweet 97: Lexicon Prediction: neutral

Tweet 98: RT @realDonaldTrump: My Executive Order to protect Monuments, Statues etc., IS IN FULL FORCE AND EFFECT. In excess of a 10 year prison term!

Tweet 98: Manual Annotation: neutral

Tweet 98: Lexicon Prediction: neutral

Tweet 99: RT @RedTracoon: The Trump administration just lent a troubled trucking company \$700 million. The company

```
was worth only $70 million httpsâ |
Tweet 99: Manual Annotation: negative
Tweet 99: Lexicon Prediction: positive
```

```
Tweet 100: RT @realDonaldTrump: ....horrible BLM chant, â Pigs In A Blanket, Fry â Em Like Baconâ . Maybe our GR
EAT Police, who have been neutralized andâ |
Tweet 100: Manual Annotation: neutral
Tweet 100: Lexicon Prediction: positive
```

```
In [20]: print("Accuracy: {}".format(accuracy_count/num_tweets))
```

```
Accuracy: 0.27
```

## Machine Learning

For this step, load the training and test sentiment datasets. The data is loaded into 4 lists of strings: X\_txt\_train, X\_txt\_test, y\_test, y\_train.

```
In [21]: X_txt_train = []
y_train = []

X_txt_test = []
y_test = []

X_txt_new = []

with open('trump-2020-train.csv', encoding='iso-8859-1') as file_handle:
    my_csv = csv.reader(file_handle, delimiter = ',')
    for row in my_csv:
        X_txt_train.append(row[3])
        y_train.append(row[4])

with open('trump-2020-test.csv', encoding='iso-8859-1') as file_handle:
    my_csv = csv.reader(file_handle, delimiter = ',')
    for row in my_csv:
        X_txt_test.append(row[3])
        y_test.append(row[4])

with open('trump-2020.csv', encoding='iso-8859-1') as file_handle:
    my_csv = csv.reader(file_handle, delimiter = ',')
    count = 0
    for row in my_csv:
        count += 1
        if count > 100:
            X_txt_new.append(row[3])
```

```
In [22]: assert(type(X_txt_train) == type(list()))
assert(type(X_txt_train[0]) == type(str()))
assert(type(X_txt_test) == type(list()))
assert(type(X_txt_test[0]) == type(str()))
assert(type(y_test) == type(list()))
assert(type(y_train) == type(list()))
assert(len(X_txt_test) == 20)
assert(len(y_test) == 20)
assert(len(X_txt_train) == 80)
assert(len(y_train) == 80)
assert(len(X_txt_new) == 3053797)
print("Asserts Completed Successfully!")
```

```
Asserts Completed Successfully!
```

Note to self:

- For this task, create a feature matrix using CountVectorizer and train a LinearSVC model from scikit-learn
- On the train split, use GridSearchCV to find the best LinearSVC C values (0.0001, 0.001, 0.001, 0.01, 0.1, 1, 10, or 100) based on the micro f1 scoring metric (hint: "micro" average) and set the cv parameter to 5
- Also, with the CountVectorizer, only use unigrams (i.e., set ngram\_range = (1,1))
- Note that GridSearchCV will retrain the final classifier using the best parameters, so you don't need to do it manually.



```
In [23]: from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.svm import LinearSVC
from sklearn.metrics import accuracy_score
from sklearn.model_selection import GridSearchCV
from sklearn.metrics import precision_score, recall_score, f1_score

import numpy as np
np.random.seed(42)
import random
random.seed(42)

# WRITE CODE HERE

vec = CountVectorizer(ngram_range=(1,1))

X_train = vec.fit_transform(X_txt_train)
X_test = vec.transform(X_txt_test)

params = {'C':[0.0001, 0.001, 0.01, 0.1, 1, 10, 100]}

svc = LinearSVC()

clf = GridSearchCV(svc, params, scoring='f1_micro', cv=5)

clf.fit(X_train, y_train)

validation_score = clf.best_score_
print("Validation F1: {:.4f}".format(validation_score))

svm_test_predictions = clf.predict(X_test)

accuracy = accuracy_score(y_test, svm_test_predictions)

print("Accuracy: {:.4f}".format(accuracy))

Validation F1: 0.7000
Accuracy: 0.5000

C:\Users\lee\anaconda3\lib\site-packages\sklearn\svm\_base.py:976: ConvergenceWarning: Liblinear failed to converge, increase the number of iterations.
  warnings.warn("Liblinear failed to converge, increase "
```

```
In [24]: from scipy.sparse import csr_matrix
assert(type(X_train) == type(csr_matrix(0)) or type(X_train) == type(np.array(0)))
assert(type(X_test) == type(csr_matrix(0)) or type(X_test) == type(np.array(0)))
assert(X_train.shape[0] == len(X_txt_train))
assert(X_test.shape[0] == len(X_txt_test))
assert(X_train.shape[1] == X_test.shape[1])
assert(type(accuracy) == type(float()) or type(accuracy) == type(np.float64()))
print("Asserts Completed Successfully!")
```

Asserts Completed Successfully!

```
In [25]: X_new = vec.transform(X_txt_new)
svm_new_predictions = clf.predict(X_new)
```

```
In [26]: assert(X_new.shape[0] == len(X_txt_new))
print("Asserts Completed Successfully!")
```

Asserts Completed Successfully!

```
In [27]: y_predict = svm_new_predictions.tolist()

total_number_of_positive_tweets_2020_ml = 0
total_number_of_negative_tweets_2020_ml = 0
total_number_of_neutral_tweets_2020_ml = 0

for item in y_predict:
    if item == "positive":
        total_number_of_positive_tweets_2020_ml += 1
    elif item == "negative":
        total_number_of_negative_tweets_2020_ml += 1
    elif item == "neutral":
        total_number_of_neutral_tweets_2020_ml += 1
```

```
In [28]: assert(len(y_predict)==total_number_of_positive_tweets_2020_ml+total_number_of_negative_tweets_2020_ml+total_number_of_neutral_tweets_2020_ml)
print("Asserts Completed Successfully!")
```

Asserts Completed Successfully!

```
In [29]: print("Total Number of Tweets: {}".format(len(y_predict)))
print("Total Number of Positive Tweets: {}".format(total_number_of_positive_tweets_2020_ml))
print("Total Number of Negative Tweets: {}".format(total_number_of_negative_tweets_2020_ml))
print("Total Number of Neutral Tweets: {}\n".format(total_number_of_neutral_tweets_2020_ml))
```

Total Number of Tweets: 3053797

Total Number of Positive Tweets: 282857

Total Number of Negative Tweets: 2729178

Total Number of Neutral Tweets: 41762