

**North South University**

**Department of Electrical & Computer Engineering**

**Assignment-1 ( Resubmission )**

|  |  |
| --- | --- |
| **Course Code:** | CSE 440.1 |
| **Course Name:** | Artificial Intelligence |

|  |  |
| --- | --- |
| **Prepared By:** | Ridowan Ahmed |
| **Email:** | [navid.rashik@gmail.com](mailto:navid.rashik@gmail.com) |
| **Facebook:** | <https://www.facebook.com/navid.rashik> |

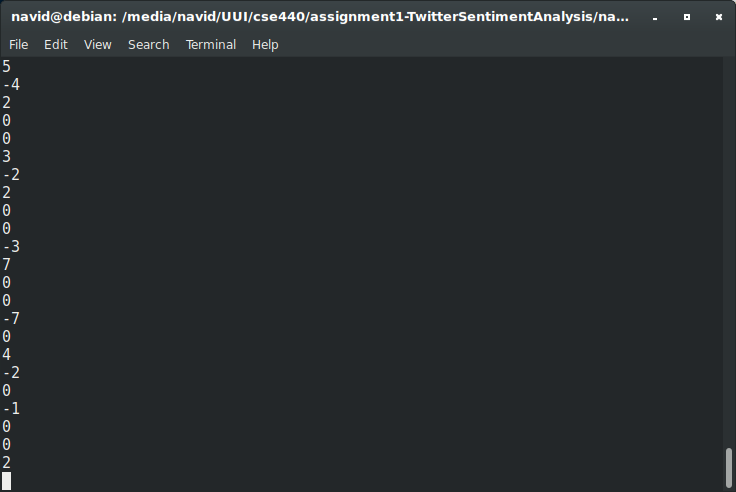
|  |  |
| --- | --- |
| **Student ID:** | **Student Name:** |
| **1410951042** | Navid Anjum Chowdhury |

**tweetstream.py**

In this file we need to fill up api key, api secret, token key, token secret from our twitter developer account.

**tweet\_sentiment.py:**

The result of running this file is :



We can see that after running this file we can find the sentiment of each tweet from our downloaded data. Some important methods are described below:

**def isLang(tweet\_json, lang='en'):**

In this function we check the json data structure if language is en or not

**def genSentDict(sent\_file):**

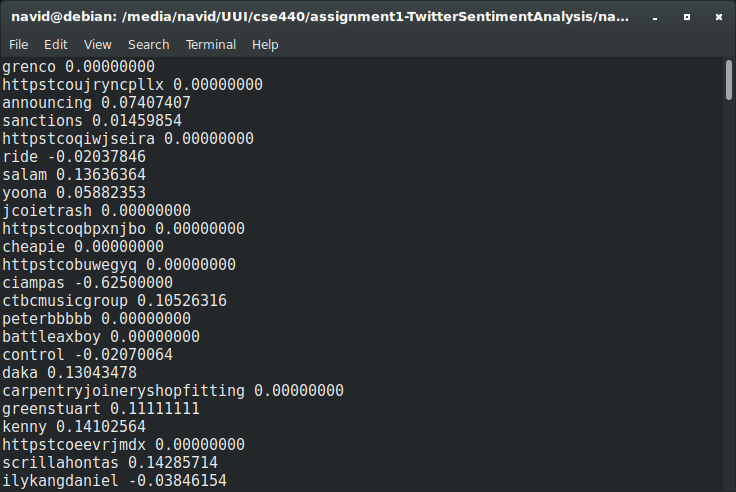
Used tweet\_json['text'].encode('utf-8') to convert the tweet into string format

**def getENTweet(tweet\_json):**

* Used stringSentence.lower() to make font lower.
* Then used tweet\_terms\_unsplited.split() to convert a list from string.
* def getSentScoreOfTweet(tweet\_json, sentDict):
* Here re.sub('[^A-Za-z]+', '', term) remove anything but a-z and A-Z from each term of each tweet in json file.
* Then calculated sentiment score and keep in dictionary.

**term\_sentiment.py**

The result of running this file is :



In this program we assume the words sentiment based on the sentiment of Afin111 file. Many words are found in tweets which are not in our sentiment file so this will analyze the those word based on given words and put a sentiment value on those word. In above result we can see that some of those word have sentiment values based on the analysis.

Some important methods are described below:

**def refinePredSentDict(newSentDict):**

In this funciton I used interpred imported from scipy.interpolate which was for mapping a range to another range.

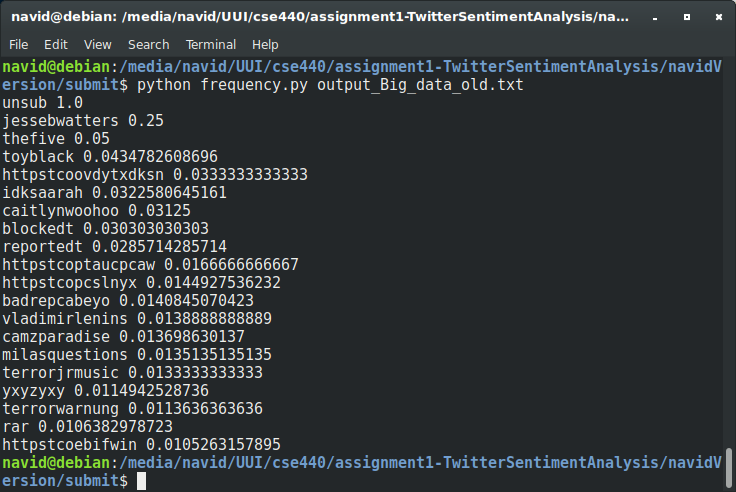
By doing score = 1.0 \* mapping(score) I have made an array(something value) to something value

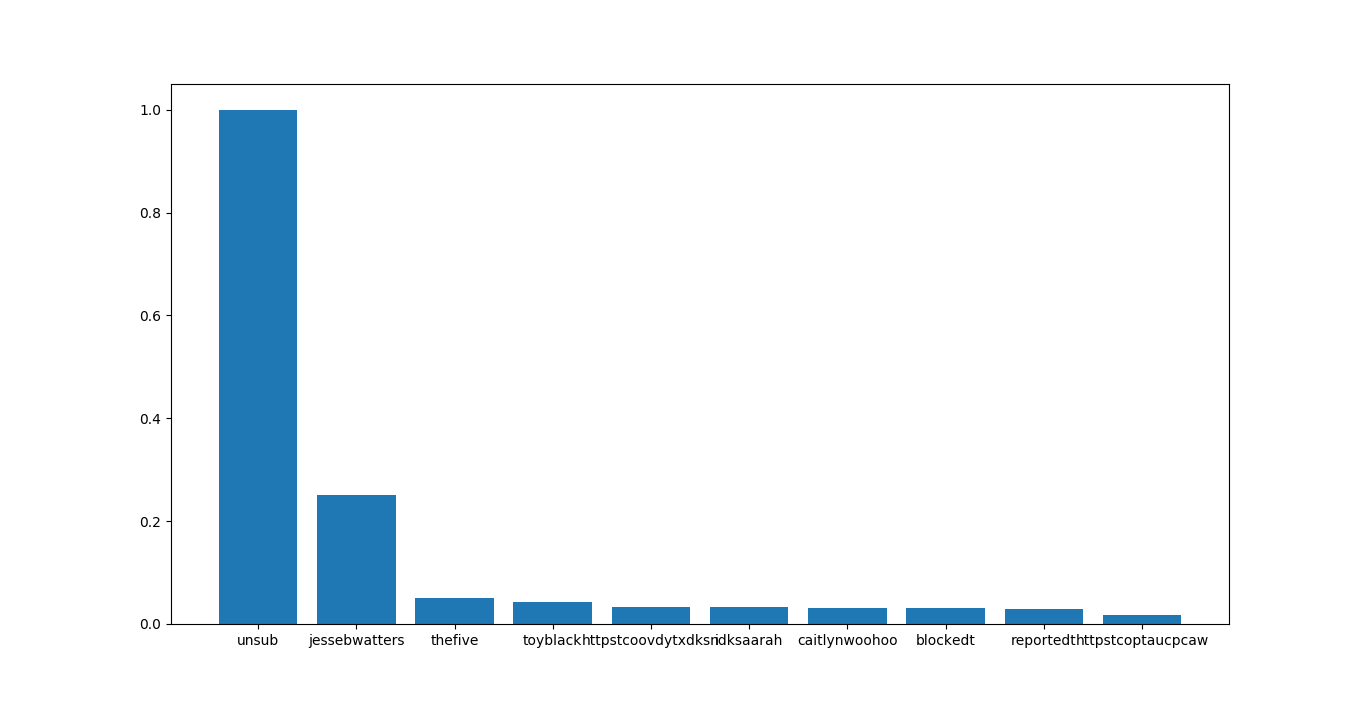
**def genPredSentDict(score, numTerms, uTerms, predSentDict):**

Added sentiment of the tweet where the unknown term is present and then also added the number of terms present in that tweet and this happens when unknown term appear in different tweet. Also update the dictionary for peritcular unknown turm.

**frequency.py**

The result of running this file is :



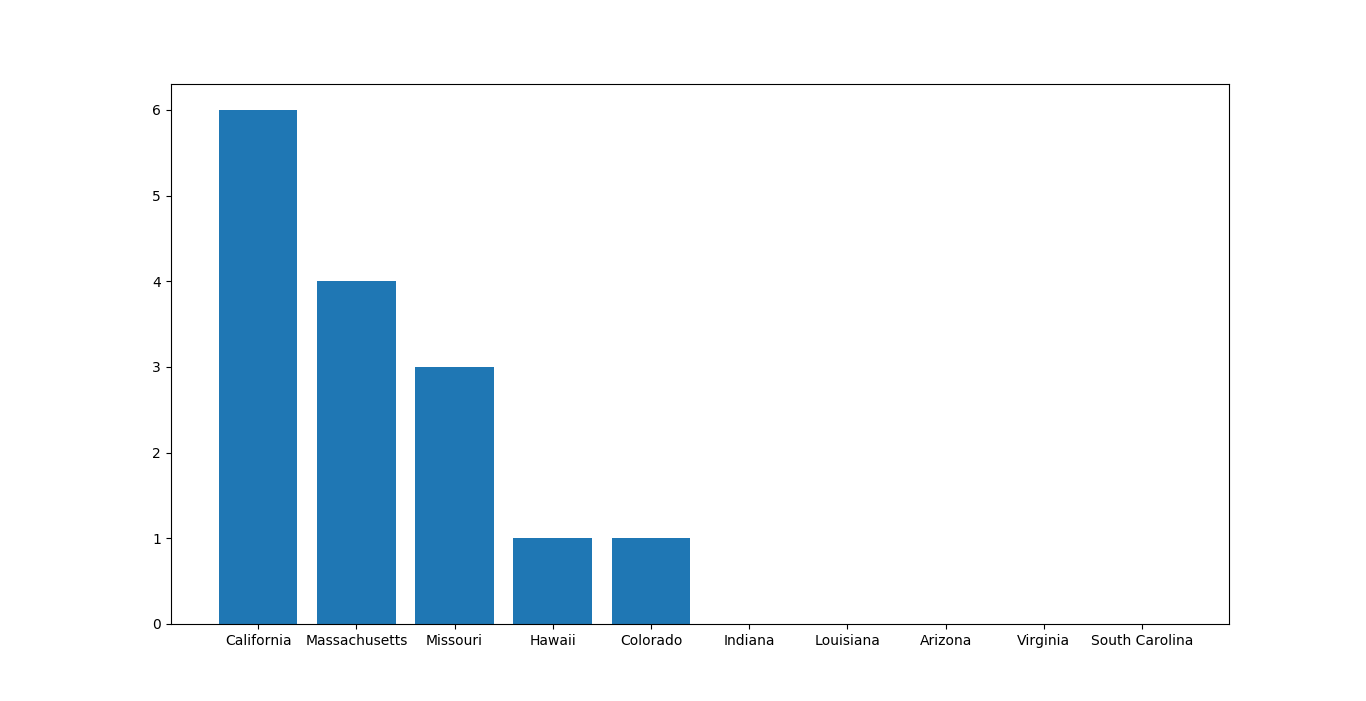
In the file It will return top 10 used word In the collected tweet and this graph represent the same thing in the console.

**def frequency(tweets\_file):**

Calculate how many time a word used in whole jeson file tweet and devide thet with all word used in tweet text. This gives a ratio.

**happiest\_state.py:**

The result of running this file is :

This program find the positive tweet from united states and make top 10 list of happiest state. This graph shows the states name on x axis and y axis represents the positive tweet from these states.

**def isCountry(tweet\_json, country = 'United States'):**

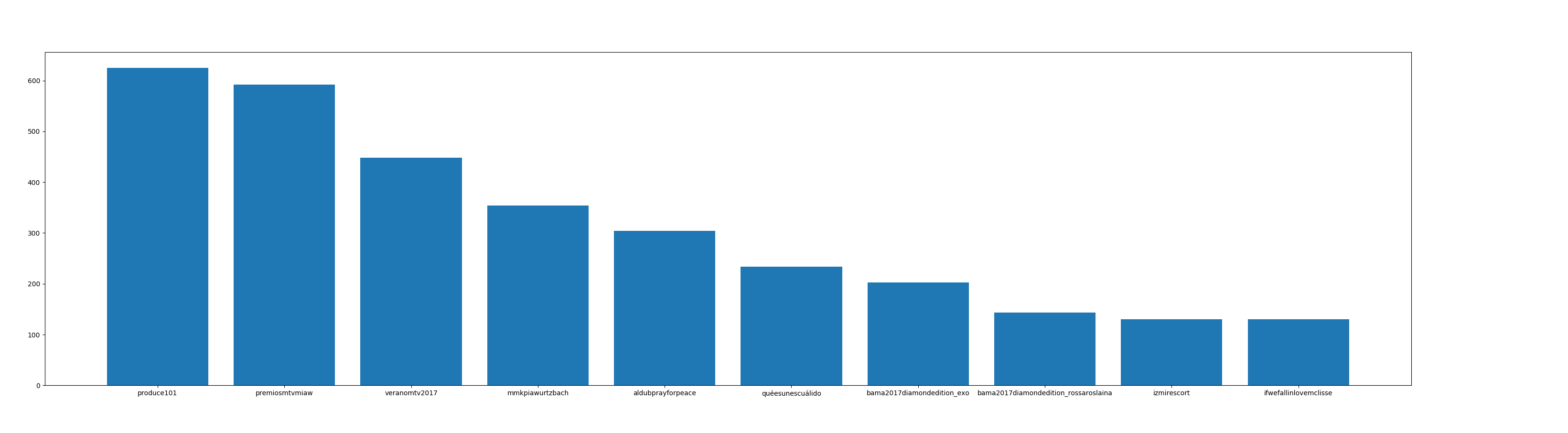
tweet\_json['place']['country'] returns full name of the country.

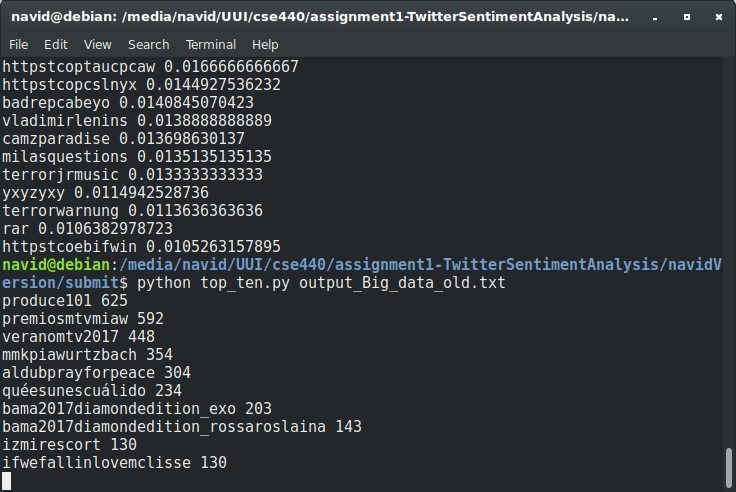
**def isStateInUSA(tweet\_json, stateList):**

In the parameter statelist represent a dictionary. This method return if the state is in usa or not.

**Top\_ten.py:**

The result of running this file is :





In the result we can see that the most used 10 tags and how many times they were found in the json file . This is the purpose of this program.

**def gettopTenHashTags(tweets\_file, n = 10):**

((tweet\_json['entities']['hashtags'])[0])['text'].lower() at first this one extract a list which contains a dictionary

then by using [0] I am selecting first item of the list which is the dictionary now after having the dictionary I have used the dictionary key name text to extract its value to hashtag