

Assignment 2

Deadline : Oct 25th 5 pm (1 hour before midterm)

Total Points : 100

Question

- Use the Twitter Api to fetch the Tweets for a topic. Create a script that asks for a search term and can perform 5 analysis on it.
- You can create 2 scripts. One to save the tweets and other for performing the analysis.

Useful Links

- <https://dev.twitter.com/rest/public/search> - **How to use Search API.**
- <https://dev.twitter.com/rest/reference/get/users/search> - **Understanding the JSON response from Twitter.**
- <https://docs.python.org/2/library/datetime.html> - **DateTime API.**
- https://www.tutorialspoint.com/python/time_strptime.htm
- https://www.tutorialspoint.com/python/time_strftime.htm
- <http://docs.python-requests.org/en/master/> - **Request API**
- <https://docs.python.org/3/library/argparse.html> - **argparse**

Script 1 - Search and Save :

- Use argparse to take the search term as argument.
- Create a function which checks if the folder is present. If not it creates the folder.
- Create a separate folder for each search term. Inside you can create sub-folders based on the date time of analysis or vice versa.
- Use the datetime module and request module to fetch the tweet and save them.
- Additionally you can create a separate function which takes a JSON as argument and filename. The function extracts useful data from the JSON files and saves the data to the filename that is provided.

Script 2 - Perform Analysis

- Use argparse and input() to ask user for input as to which analysis to perform.
- **Some of the analysis can be :**
- 1. **For a given search term (or across all search term) and a time window, what is the average number of friends a user has each day.**
Eg : date, search term, avg_friends.
Question : People who tweet regarding Trump have more or less users friends to Hillary. This can give idea about user reach.
- 2. Mapping the city to the state (you need an external dictionary containing the mapping) which states tweet more on the topic.
Eg : Date, search term, region/state, number of tweet
Question : Which states seem to tweet more regarding Hillary compared to Trump.
- 3. For each topic, show the top 10 retweeted tweets for every day during the time window.
Eg. date, search term, tweet text, # of times retweeted.
Question : For each day show what people are talking about the presidential debate.
- 4. For each topic, weighing each tweet by the follower_count * retweeted count, show top influential tweet for each state.