

Twitter API – Tweet Extraction, Storage and Sentiment Analysis

Goal:

To fetch tweets from twitter, using Twitter API. Store them in MongoDB. Get familiar with StreamListener calls and NoSQL database operations. Use these tweets for sentiment analysis or try building an application.

Roadmap:

03/29 Implemented the python script to fetch the tweets. Still working on optimizing the code. Uploaded the sample json file with tweets using the keyword. ['emoji'] Looking forward to start with the back-end part soon.

04/17 In my next update, I started working with the MongoDB database, which is a document based NoSQL database. Spent some time in installing the software and setting up the mongo server. Got familiarized with the commands and finally integrated my code with the database. Now all the tweets that I was fetching, I'm also storing them in the MongoDB collection. One advantage that I've felt by using MongoDB over any other relational database is that it is a difficult task to store these different tweets in form of a tuple or csv in a relational database, since most of the tweets have different structure. MongoDB eases this functionality by storing the entire tweet as a document in preferably a json format. Moreover, it'll be easier to fetch tweets back from storage and analyze. One final part of the project remains - which is the analysis of these tweets. I hope to continue and complete the remaining part soon.

05/06 Final Update - In this update and in the final part of my project, I worked on writing a Python script to fetch tweets from Twitter using Python's Tweepy Module and perform sentiment analysis on the data. NLTK suite was used to carry forward the actual analysis. Used TextBlob library to process the textual data. We can test the code by entering a keyword of our choice and then categorizing the corresponding tweets as Positive, Negative or Neutral.