Mini Project 3 Reflection

Alix McCabe

February 2015

1 Overview

I used Facebook as my data source, pulling in my newsfeed to gain data from status updates. I utilized the pattern.en sentiment analyzer to gain insights about the subjectivity and positivity of my facebook friends' statuses. From this, I created a graph of all of my friends' statuses based on their sentiment.

2 Implementation

In order to have access to my facebook data, I decided to pickle it after importing it. This enabled me to have access to the data more quickly than if I were to import it from scratch every time I ran the code. It was also much easier to manipulate the text file than the raw facebook output. Once this was done, I evaluated the results line by line to determine their sentiment. I was able to determine which of the results from the import were statuses and which were not by a simple boolean. As long as the sentiment isn't zero, the entry is added to the list of values to be plotted. I put this if statement in because of all of the additional information the facebook newsfeed request generates.

I tried a multitude of other ways to strip out just the statuses and found this to be the most effective. One of the other ways I tried was to create a dictionary of "buzzwords" such as "life event","became friends", and "shared with" that indicate a randomly generated facebook post. Of course, I didn't want to evaluate these posts since they weren't statuses, so I had a function check whether the buzz word was in the entry I was evaluating and delete the entry if it was. Ultimately, I decided not to go with this method of because it was much more efficient to merely check that the sentiment was not equal to zero, indicating that the entry was a random list of numbers, a random name, or an autonomously generated post.

3 Results

I analyzed how objective vs. how subjective my facebook friends' statuses are. My results were very strongly weighted, but stand up to qualitative reasoning. As we can see, the cluster of statuses is in the top right quadrant, which indicates the highest levels of subjectivity and positivity.

There are a few highly subjective negative statuses and a few objective positive statuses, which also stands to reason. However, there are no objective negative statuses, which stand to reason since when people post on facebook they rarely post non-subjective negative statuses (at least in my experience, negative status are generally heavily charged with personal sentiment).

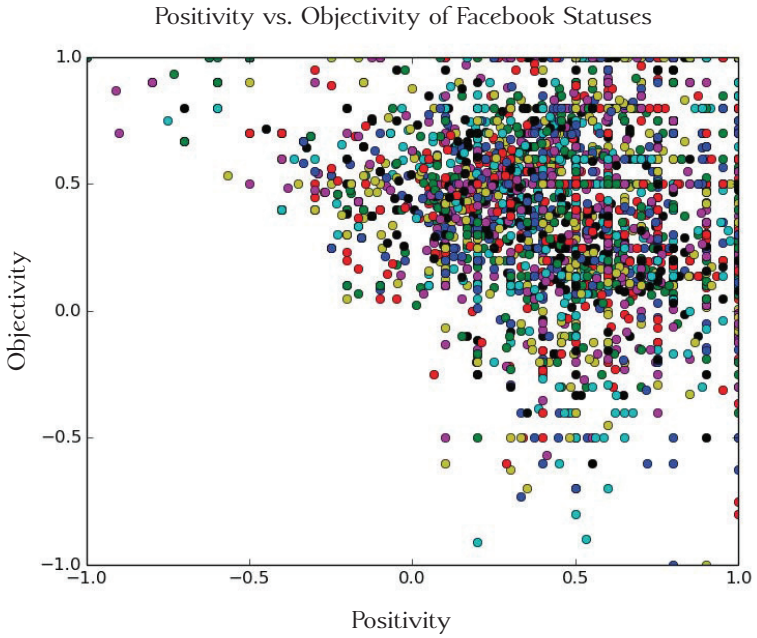

Figure 1: This is the plot of our statuses with on an x-y plane with objectivity on the y axis and subjectivity on the x-axis.

4 Reflection

I thoroughly enjoyed exploring the various APIs associated with this project, and wish that I had had time to do something with each of them. It was definitely a challenge to decide which to utlize and what to do with the data gathered. It was also a challenge for me to structure the code completely independently, since this is my first experience doing so. I felt that I tried many, many startegies that ended up working only to end with extremely simple code. I'm proud of my final code and how efficient it is, though I acknowledge that I spent a huge amount of time with trial and error trying to find strategies that would yield my desired results.