Using IBM's Tonality Analysis of Language and Geolocated Tweets to Map Emotional Intensity.
Research Proposal

Isaac Callison (ic2d@mtmail.mtsu.edu)
Middle Tennessee State University

March 13, 2020

1   Introduction

There is a convergence of powerful technologies that allow for near- instantaneous notification of current events using available meta-data from social media platforms. The goal of this project is to gather and analyze tweets for emotive tonality, then display this on a Google "heatmap" of emotive intensity.

There are three separate technologies that will be interdependent in this project. First, using Twitter's developer API, geolocated tweets will be collected from a specific region and radius, then pre-processed to extract lattitude, longitude and text. Secondly, using IBM's Watson, the tweets will be assesed for emotive tonality. Finally a Google map will be dislayed with a heatmap layer graphing the intensity of these emotions.

2   Specific Aims

- Access the developer consoles and APIs of Twitter, IBM's Watson, and Google Maps to see if the three technology giants can mesh.

- Create an Jupyter Notebook that pulls in geolocated tweets that IBM's Watson can analyze for emotive tonality.

- Display a heatmap of emotion in a certain area based on intesity of the selecteed emotional state.

3   Background

As stated above, there will be several interdependent moving parts with this project. With regards to platform, a Jupyter Notebook will be used for this project to pull in modules, access API's, and make GET and POST requests. The datasets for the natural language processing will come from Twitter. The Twitter API allows for the triangulation of geolocated tweets[5]. Using developer authentication, and a GET request with certain location parameters, one can obtain a list of current tweets within a search radius in JSON format. A great deal of meta-data is returned in this format, but from these tweets one can glean a myriad of data, including the lattitude and longitude of the tweet.

The second part of this project is natural language processing with IBM's Watson using tonality analysis. IBM has a cloud computing program with various machine learning capabilities[3], one of which is tonality analysis. The natural language tonality processing that Watson offers can, among other things, extract emotion from a corpus. In this specific case, a variety of emotions, from sadness to excitedness, and the intensity of that specific emotion, can be derived at the sentence and document level.

A heatmap is a overlay feature offered by Google Maps. It can create a visualization to depict the intensity of data at range of geographical points. This is good when you have lot of data points of varying magnitude. When the Heatmap Layer is enabled, a colored overlay will appear on top of the map. By default, areas of higher intensity will be colored red, and areas of lower intensity will appear green[2].

4 Preliminary Results

A few steps have already been implemented to see this project to fruition. The three aformentioned services reguire developer accounts to be used. Twitter's API required an application and pre-approval. These credentials have been secured. Postman, a REST API testing software has been used to experiment with Twitter's API[4]. Further, a skeleton Jupyter Notebook has been created and uploaded to Github[1]. To date, this Notebook just pulls in and demonstrates the Google heatmap layer with some sample earthquake data and does some basic tonality analysis on a few sentences.

5 Work Plan

The Jupyter Notebook is the nexus of this project. Moving foward attempts will be made to integrate calls to the twitter API using the credentials referenced earlier. Passing in the lattitude and longitude of a specific region should yield tweets that can then be parsed. IBM provides a Watson Software Development Kit for integration into Python and Jupyter. The text from the obtained tweets will be passed into Watson for tonality processing. A Google Map will display a heatmap layer with the attendand emotion and intensity. For instance, in areas of a region where the tweets have low sadness, green shading will predominate.

6 Broader Impacts

We live in an age of ever expanding meta-data. As this increases, humanity will seek to harness this data through new technologies, for better or worse. In this case, Twitter data can potentially be used to inform and improve the lives of everyday citizens. If the above can be implemented, one could graph emotional intensity of differing regions based on twitter content.

References

[1] Google,. Heatmap API. `https://github.com/cthulhu1988/DangerFloof`, 2020. [Online; accessed 29-February-2020].

[2] Google,. Heatmap API. `https://github.com/cthulhu1988/DangerFloof`, 2020. [Online; accessed 29-February-2020].

[3] IBM Incorporated,. IBM's Watson: Cloud Computing. `https://cloud.ibm.com/`, 2020. [Online; accessed 29-February-2020].

[4] Postman Incorporated,. Postman API. `https://www.postman.com/`, 2020. [Online; accessed 29-February-2020].

[5] Twitter Incorporated,. Twitter Documentation: Geocode API. `https://developer.twitter.com/en/docs/geo/places-near-location/api-reference/get-geo-search`, 2020. [Online; accessed 29-February-2020].