

Team Members: Karinna Loo, Allison Schwartz, Kate Miller

Project Category: Implementation, Empirical Analysis

Idea: Ivy League Tweet Analyzer. User will be able to see tweets with coordinates from around ivy league universities, grouped by campus. Tweets from outside a tight radius will be discarded. From the most popular nouns, verbs, and adjectives from a particular school, a conglomerate “most popular sentence” will be generated and displayed to the user according to a template selected by the user (subject/object/verb, subject/verb/adverb, etc). The most popular hashtags, most common emoticons, and other fields as needed, from each campus will also be displayed. We will use a pre-built java-based natural language processor, TweetNLP (<http://www.ark.cs.cmu.edu/TweetNLP/>), to identify the parts of speech that appear in the tweets in order to reconstruct them.

Describe which topics will be used in your project and how: Document Search. The twitter API will be used to produce a json or csv containing the tweets, which will then be parsed to produce the analysis described above. Social Networks for the interface with the twitter API.

Short work breakdown:

1. Create user interface (initially command line, graphical if time permits) - all
2. Set up coordinates of universities within program and select an appropriate radius - Kate
3. Read in tweets from twitter API and store in document form - Kate
4. Select on stored tweets and allocate to specific universities - Kate
5. Process university tweets for most popular parts of speech, most popular hashtag, most popular emoticons - Karinna & Allison
6. Generate most popular sentence per user input - Karinna & Allison
7. Display all results - all