Project # 2

Analyzing Tweets/Blogs

This project is designed to introduce you to do text mining of tweets or blogs. You can decide to work with tweets or blogs; it is your choice. The major steps of the project are as follows:

1. Set up an arrangement to aggregate blogs or tweets on a daily basis on a topic of current interest. For example, you could use #Lions to search tweets about Detroit Lions. You will collect tweets/blogs for at least one week.
2. The collected tweets/blog posts will constitute your data for mining.
3. The next step is to decide on your actual mining task. You are free to choose your task and accordingly you will figure out the needed steps. Examples of some tasks are:
   1. You may want to classify each tweet/blog post into favorable or unfavorable to the team and generate a chart showing daily score of favorable/unfavorable trend.
   2. You may want to cluster daily tweets/blogs to see the emergence of different subtopics in the tweets.
   3. You may want to track the spread of tweets.
4. Prepare a report clearly outlining each step along with details of every tool used and the data collected, and a summary of your results.

Due Date: December 11

The project will be done on an individual basis. One presentation on project will be required. The presentations will take place on 12/3 and 12/5. The presentation order will be reverse alphabetical of last names. Final report with results is due on December 11.