

## Project # 2

### Analyzing Tweets/Blogs/Reviews

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

1. Set up an arrangement to collect/scrap data. For example, you could use #Lions to search tweets about Detroit Lions. For product reviews, you could use cites like: <https://www.bestbuy.com/site/reviews/samsung-27-8-cu-ft-4-door-french-door-refrigerator-with-food-showcase-fingerprint-resistant-black-stainless-steel/6323228?variant=A>
2. The collected tweets/reviews 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:
  - a. You may want to classify each tweet/blog post into favorable or unfavorable to the team and generate a chart showing daily/weekly score of favorable/unfavorable trend.
  - b. You may want to cluster daily tweets/blogs to see the emergence of different subtopics in the tweets.
  - c. You may want to build a classifier that predicts the number of stars a review will receive
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: November 30 and December 2 Presentations

December 6 Final Report

**The project will be done on an individual basis.**