Assignment 5:

Due November 18

1. 18K comments about the 2016 president election debate were collected from Youtube:

https://www.dropbox.com/s/qnbw2ulzr85w4wg/The%20First%20Presidential%20Debate.zip?dl=0

(1) (10pt) write a program to analyze sentiment in the text comments in the "textDisplay" fields. Some lexicon resources can be found at:

http://nlp.cs.rpi.edu/course/fall16/sentimentdictionary.zip

And generate sentiment (positive/negative/neutral) distribution for each president candidate.

If you want to try supervised models, you could consider using the training data from 2012 president election:

http://nlp.cs.rpi.edu/course/fall16/sentiment_train.xls

(2) (5pt extra) use structured information such as user IDs (the same user tends to have consistent sentiment toward the same candidate) and likes/dislikes.

(3) (5pt extra) detect specific issues and aspects (e.g., foreign policy, economy) and associate them with sentiment analysis results.