Assignment 5 - aka "the project"

CS 173 – 2019 SP [XXX points]

Choose from any of the topic areas below. Work in a small team (up to 3 people), else minus 10%. Define a short project of your own design that demonstrates understanding of the topic. Write a minimum 2-page summary of the effort not exceeding the space provided below.

DO NOT REUSE A PROJECT YOU DID IN ANOTHER COURSE. A recycled project will be considered a violation of academic integrity in this case and may result in severe penalty, including failing the entire course (not just this assignment). Don't do it; do something new.

Rubrics - in a nutshell, show-off that you learned something cool that you are proud of
- well-sourced: spent time and energy finding something interesting
- well-scoped: a "goldilocks" sized project (not too big, not too small), estimate 10 hours
- well-resourced: identified a solid and adequately sized dataset (Kaggle, etc.)
- well-implemented: used appropriate packages, toolkits, or own source code
- well-summarized: a concise and clear description of the problem, data, and methods
- well-written: complete and coherent sentences, thorough yet concise description of project

Topics
- named-entity recognition
- sentiment analysis
- text classification
- word sense disambiguation
- query completion (auto-complete)
- hyponomy/hypernymy learning
- text generation
- other (clear with professor early in advance)

Methods or tools to consider: NLTK, Stanford CoreNLP, BRAT annotator, language models, vector-based methods, grammars and parsing, regular expressions, parts-of-speech tagging, conditional-random fields, distributional semantics, latent semantic analysis

Team members:

URL link to project source code & data:

Name: _____ StudentID: _____

Name: _____ StudentID: _____

Name: _____ StudentID: _____