# Course Projects

Instructor: Alan Ritter

# Final Project

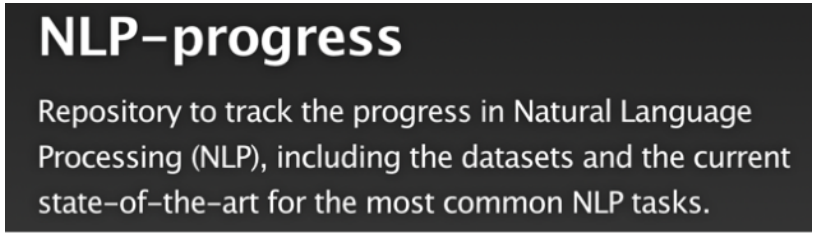
- Groups of 3-4
- Sign up here:
  - https://goo.gl/forms/FK6Yv84ac2fhFyH42
- **Scope:** on the order of one of the programming assignments
  - But, need to define the problem you are addressing, propose a solution, design evaluation and write up results in a formal report.

# Selecting a Topic

- Find a problem/application you are interested in where you think ML can help.
- Experiment with one of the algorithms we discussed in class.
- **First question:** what is the dataset?

# Datasets

- Sentiment Analysis (and other NLP) Datasets:
  - http://nlpprogress.com/english/sentiment_analysis.html
- Semeval Tasks:
  - http://alt.qcri.org/semeval2018/index.php?id=tasks
- Kaggle Competitions
  - https://www.kaggle.com/competitions




# Requirements

- 4 Page Report
  - Due December 10
  - **Late reports will not be accepted**
- Briefly describe the problem you are solving and your solution.
- Include empirical analysis of your approach
  - Report performance on dev / test set
  - Compare against some reasonable baseline method.

# Advice

- **First question:** is the data available?
- Try to get a simple baseline working as early as possible to determine whether your project idea is feasible.
- Start with a manageable-sized dataset
  - Then scale up…

# Grading Rubric

For the reasonably well-prepared reader, is it clear what was done and why? Is the report well-written and well-structured?
Clarity (1-5):

How original is the approach? Does this project break new ground in topic, methodology, or content? How exciting and innovative is the work that it describes?
Originality / Innovativeness (1-5):

First, is the technical approach sound and well-chosen? Second, can one trust the claims of the report -- are they supported by proper experiments, proofs, or other argumentation?
Soundness / Correctness (1-5):

Does the author make clear where the problems and methods sit with respect to existing literature? Are any experimental results meaningfully compared with the best prior approaches?
Meaningful Comparison (1-5):

Overall (1-5):

# Group Formation Time

- **(5 minutes)**
- https://gather.town/invite?token=M4bC1odjRE6p3CdUD5gso8tw0vAJs2ZC