

# Research Questions

- Given set of online forum posts, how do we automate accurate classification into labels (helpful to forum moderators) defined by human experts?
- What subset of possible features and machine learning classifier construction achieves the highest accuracy?
- Can we mitigate the problem of having only a small subset of posts with a certain label in the training dataset?
- Are there discernible social patterns exhibited by cliques of forum users?

# Dataset Selection

- 2016-2017 CLPsych workshop organizers provided an official ~150,000 post dataset with ~1,200 labeled posts used for the task, but unbalanced distribution of post labels (50% green vs. 0.5% crisis)
- Other datasets used by teams for semi-supervised learning and bootstrapping include sub-Reddit posts on depression/anxiety/suicide and Twitter posts with (e.g.) #depression
- We will focus on supervised learning approach using ~1,200 posts from official dataset
- Concerns about American vs. Australian english and even ReachOut.com-specific language or norms, and general lack of success with previous semi-supervised learning work (and our own lack of expertise)...

# Approach

## Features:

- 1) Token unigrams and bigrams with context including previous posts in the thread
- 2) Forum metadata: author rank and board
- 3) Subjectivity/Imperative mood
- 4) Post and/or last sentence sentiment/emotion
- 5) Social link graph risk score

## Learning Algorithms:

- Multiclass SVM
- Staging Binary SVMs, use different sampling methods to produce balanced training set, especially for last stage where the Red and Crisis posts are classified.
- (Ensemble)

# Evaluation and Baseline

## Metrics:

- F1 score
- Recall
- Accuracy
- Precision

## Baseline:

- SVM with token unigrams (inspired by Kim et al. with official F1 score of 0.42 in 2016 CLPsych task)

# Plan and Timeline

6 Nov: Baseline feature vector + Weka results

13 Nov: Features 2-3 + Weka results

20 Nov: Features 4-5 + early custom classifier

27 Nov: Feature 6 + custom classifier

4 Dec: Refined feature vectors + custom classifier

11 Dec: Final write-up of results

\* Tim's and Julien's focus

