Problem  We are interested in jointly predicting sarcasm, disagreement and different aspects of argumentation in online debate forum. Intuitively, we think sarcasm and disagreement are correlated: people may post sarcastic comments when they disagree with some previous posters. Current work on automatic sarcasm detection applies supervised learning using Twitter data with hashtags such as # sarcasm [1]. On the other hand, a lot of work is done on disagreement detection utilizing online debate data [2]. Since we think these sarcasm and disagreement are related, and no work has been done to jointly learn both, we hope to improve the performance of both tasks by considering them together.

Method  Currently, we plan to try two approaches. Firstly, we intend to try multi-label SVM with lexical, syntactical and sentiment features (and potentially more if we see fit). We are also interested in applying multi-task learning where we treat detection of sarcasm, disagreement and other argument labels as different tasks.

Dataset  We intend to use the Internet Argument Corpus [3], which has around 11,216 discussions and 109,553 posts taken from online debate forums. The corpus has been annotated by Mechanical Turk with various labels for analysis. For our project, we are particularly interested in the sarcasm, agree-disagree, fact-feeling and negotiate-attack labels. The category with the least amount of labeled posts is sarcasm. Since around 8% of the posts contain sarcasm, we think we have a reasonable amount of data for each type of labels.

Evaluation  As mentioned before, there are existing work for sarcasm and disagreement detection separately. We can evaluate our methods by comparing the performance of our model against the state-of-the-art approaches in those tasks separately in terms of accuracy, precision and recall.

References
