

# LIGN 165 Final Project: Sarcasm Detection

Group Members: Victoria Tom, Kyuheon Lee, Nicolas Martin, Josh Duhay

# Question and Emphasis

- Research Question: How can we detect written forms of sarcasm?
- Focused on following previous research that answered this question using sentiment analysis

Sentiment analysis basics:

Positive phrase: I like, I love, I enjoy

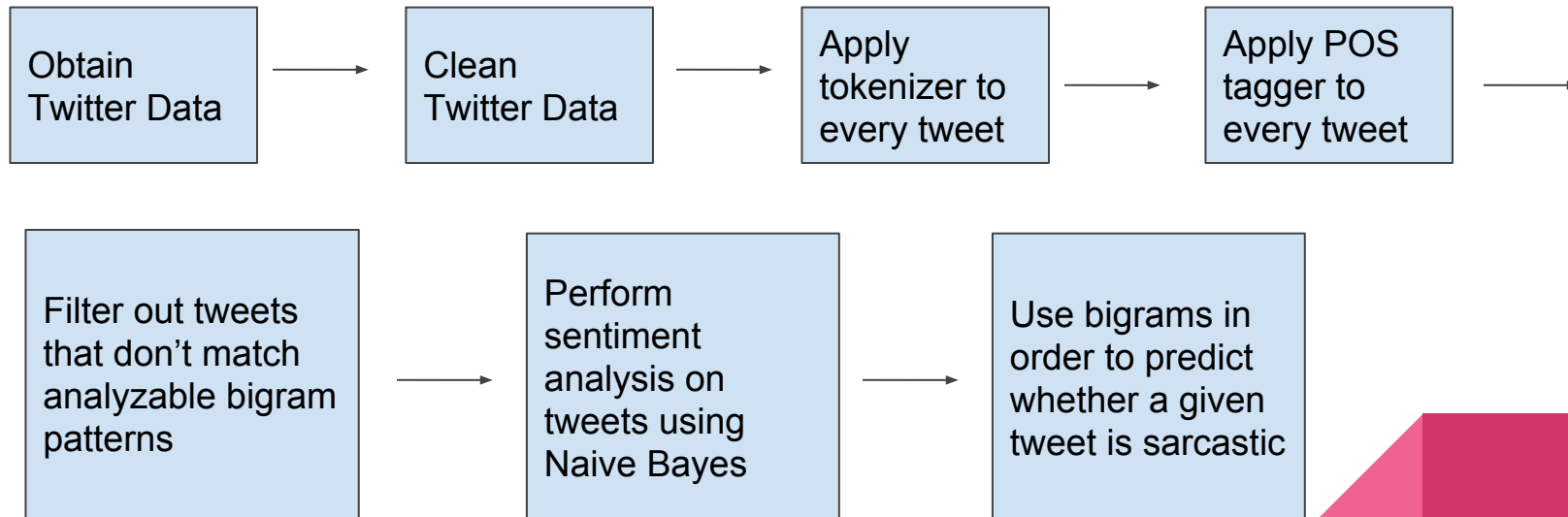
Negative phrases: I hate, I dislike, I despise

Sarcastic phrase: I love doing homework (positive/negative)

Non-sarcastic phrase: I hate doing homework  
(negative/negative)



# System Design and Components



# Examples

Sarcastic:

```
['yes', 'love', 'waking', 'and', 'her', 'little', 'brother', 'fighting', 'sarcasm']  
prediction : True  
sarcasm True
```

Non-sarcastic:

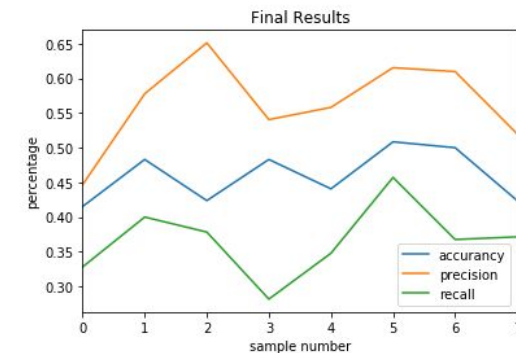
```
['bookmarked', 'and', 'will', 'check', 'out', 'after', 'get', 'some', 'sleep', 'thanks', 'sam']  
prediction : False  
sarcasm False
```



# Result/Performance

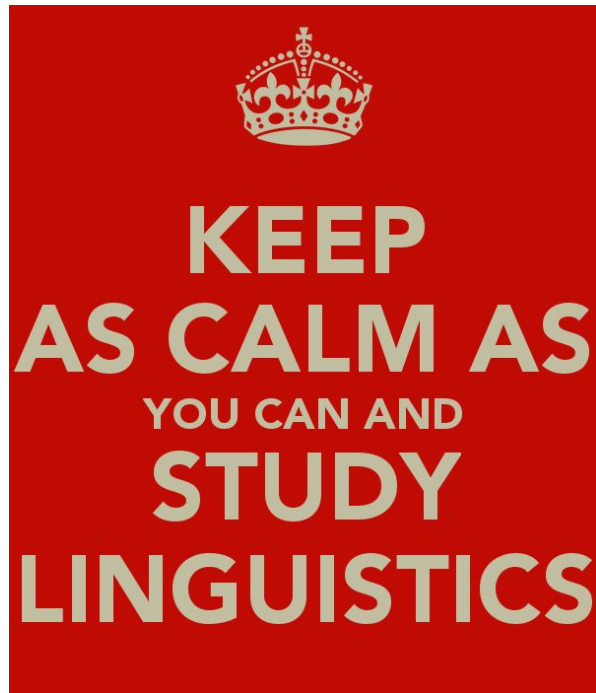
- We got better precision and recall values than the original study using similar methods
- Ours: recall = .4, precision = .63
- Theirs:

accuracy average : 0.459745762712  
precision average : 0.564946233435  
recall average : 0.366529761697



System	Recall	Precision	F score
<i>Positive and Negative Sentiment, Unordered</i>			
Liu05	.19	.37	.25
MPQA05	.27	.30	.29
AFINN11	.17	.30	.22

# Successes and Regrets



# List of External Resources

- nltk module : <http://www.nltk.org/>
  - POS tagger
  - Word tokenizer
- tweepy module: <http://www.tweepy.org/>
- University of Utah : <http://www.cs.utah.edu/~huangrh/official-sarcasm-cameraReady-v2.pdf>
  - Data sets (gold standard)
- <http://thinknook.com/twitter-sentiment-analysis-training-corpus-dataset-2012-09-22/>
- <http://www.laurentluce.com/posts/twitter-sentiment-analysis-using-python-and-nltk/>
- [Twitter Sentiment Analysis Dataset](#)

THANK YOU

