Information Extraction from Microtext

Kimberly Paterson
Advisor: Dr. Alex Dekhtyar

Twitter, Spoons, & Outages

- Spoons
  - Telling netflix what Twitter users say about services
- Concern
  - No "why" associated with outage
- Goal to define "details" of outage
  - Location (geographic & platform)
  - Description of "cause"
  - Threat levels

Tools & Techniques

- Information Extraction
  - Extract relevant information from tweets
- Machine Learning
  - Labelling "Outage" tweets
  - Generating rules
- Latent Dirichlet Allocation (LDA)
  - Topic modeling
- "Bag-of-words" approach

Novelness

- Most semantic information extracted from Twitter is sentiment or opinion mining
  - I would call that "superficial"
- Event Detection
  - natural disasters
  - political coos
  - movie releases