Feature Extraction and Analysis for Tweet Text - Application for Predicting Stock Trends

Student: Haikuan Liu
Supervisor: Dr. Timothy Graham
Co-supervisor: Prof. Hanna Suominen

COMP4560, The Australian National University
23 October 2017
Why experiments on Tweet Text?

- Unstructured
- More casual
- Popular
- Fast
- Challenging

= In some cases, more informative
A quantitative stock prediction system based on financial news. Robert P. Schumaker, Hsinchun Chen


Twitter for health – building a social media search engine to better understand and curate laypersons’ personal experiences. Hanna Suominen, Leif Hanlen and Cécile Paris
Approaches

Feature Extraction:
Grammatical Features: Dependency Parsing, Part of Speech, Named Entity, etc.
Other Features: Trends from market prices of a company.
  e.g. Up-down, Swing-Stable.

Analysis Algorithms:
Machine Learning e.g. SVM
Expected Outcomes and Extensions

Expected Outcomes:

Able to extract meaningful and precise information from Tweet text by grammatical features.

Could apply analysis algorithms to predict stock market trends.

Link stock market trends with grammatical features.

Extensions:

More than Tweet text
More than English
More than Stock Market