Mini Project: Final Presentation

Karan Goel
Jyotesh Singh

Project Objective

- *To identify sentiments expressed and opinion holders in topically relevant tweets*
- Task Breakdown:
  - Topical Relevance of Tweets
  - Crowdsourcing/Tagging of Test Sets
  - Sentiment Analysis Preliminary (Baseline)
  - Entity Discovery
  - Triplet Generation
  - Sentiment Analysis Final

Topical relevance of Tweets

- Used a dictionary of topics on tweet corpus
- Top 20 (most popular) topics
- 4500 topically relevant tweets
- Utilized exact match occurrence of topic in tweet

Crowdsourcing/tagging of sets

- Extracted well-distributed randomized subsets of previously determined topical tweets (distribution among topics)
- Was to be manually annotation by students in Prof. Mausam's NLP course – but no end result
- Interim methodology:
  - Manually tagged 200 tweets out of these for sentiment
  - Manually tagged 600 tweets for entities

Sentiment Analysis (Preliminary)

- Sentiment analysis using SentiWordNet list & emoticon regex
- Identied +ve {P} & -ve {N} sentiment words in SentiWordNet (strong minimum threshold) - helps eliminate some context sensitivity
- Split tweets into {Sentimental, Non-Sentimental} where Sentimental if sentiment word (from {P,N}) found or emoticon found
- **Results:** Extremely poor discernment of sentiments due to lack of context, and uncertain thresholds.

Stanford Sentiment Analysis (Preliminary)

- Using Stanford CoreNLP, carried out sentiment detection for a small set of tweets

  ["Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank" - Socher, et al]

- **Conclusion:**

  Qualitative analysis showed poor results in both cases due to clear misattributions in a large number of tweets.

Entity Discovery (for Opinion Holder)

- Two methodologies:
  - Maximal String Matching of YAGO entities (scraped via PostgreSQL)
  - AIDA: Max-Planck institute's environmentally aware engine which uses Stanford named entity recognizer in conjunction with YAGO ontology
  - Tested on 600 manually marked tweet set

Method	Precision	Recall
AIDA	66.02%	52.61%
MSM	63.26%	32.3%

  - AIDA also run on 10 articles to get 75.7% Precision & 67% Recall

Triplet generation (baseline)

- Existing code optimized for articles
- Utilizes NLP concepts based on Stanford CoreNLP, negation handling and subject-object resolution
- Computational approach
- Original code: uses opinion holding verbs (“oppose”, “supports” etc.)
- Replaced by: Sentimental words from SentiWordNet (nouns, verbs, adjectives)

Triplet generation (baseline) (results)

- 22.46% Precision, 48.44% Recall of Sentiments
- 20% Precision in identification of opinion holder
- Results given correctly identifies sentiment:
  - 74.19% Precision in opinion holder
  - 70.97% Precision in object

- **Conclusion:**
  - Poor performance due to non-verb sentiment words
  - If sentimental verb, relative ease of splitting into subject-object
    - Implication: Simplistic sentimental sentences in twitter?

Classifying Sentiment using ML

- Literature
  - Kouloumpis, et al (”Twitter sentiment analysis: The good the bad and the omg!”)
    - Suggests POS features reduce performance
  - Socher et al. (“Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank”)
  - NCSU (Healey & Ramaswamy, ”Visualizing Twitter Sentiments”)
    - Computational approach to sentiment classification

Hybrid Approach

- Based on the literature, a mixed approach with features
  - Top 5 Sentimental Word Scores (from SentiWordNet) in the Tweet
  - Presence of
    - Hashtags
    - Positive Emoticons
    - Negative Emoticons
    - Links
    - Negation
  - Count of Sentimental
    - Adjectives
    - Nouns
    - Verbs

Generation of Features

- Parse tweet for binary features
- Using Stanford NLP toolkit along with SentiWordNet to extract verb, adjective and noun counts
- Also find top k sentimental words

Results

- Sanders Tweet Corpus (~3000 tweets)
  - SVM based vote classifier (70.20% precision)

	True pos	True neg	True neu	class precision
pred. pos	67	17	28	59.82%
pred. neg	9	69	33	62.16%
pred. neu	374	400	1892	70.97%
class recall	14.89%	14.20%	96.88%	

Results

- Koulampis et al. Twitter Corpus (500 tweets)
  - SVM classifier (72.00% precision)
  - Macro-F measure of 73

	True pos	True neu	True neg	class precision
pred. pos	11	1	1	84.62%
pred. neu	7	16	4	59.26%
pred. neg	0	1	9	90.00%
class recall	61.11%	88.89%	64.29%	

Conclusions

- Results are promising, but need a large, reliable training set