Sentiment Analysis in Twitter with Lightweight Discourse Analysis
- S.Mukherjee, P.Bhattacharyya

Rohit Kumar Jha

October 16, 2013
 Outline

 Introduction
  About
  Motivation
  Background
  Overview of Approach

 Details
  Related Works
  Algorithm

 Evaluation and Results
  Dataset 1
  Dataset 2
  Dataset 3
  Results

 Drawbacks

 References
Introduction
About

- A lightweight method for using discourse relations for polarity detection of tweets

- Incorporate discourse information in bag of words model to improve accuracy
Motivation

- Targeted towards web-based application that deal with noisy, unstructured text, like the tweets
- Can not afford to use heavy linguistic resources like parsing
- Bag of words model ignore discourse particles like but, since, although, etc
Background
Sentiment Analysis

- Also known as opinion mining, uses NLP, Machine Learning and Computational Linguistics to extract subjective information in source material.

- Polarity Detection involves determining of the sentiment/mood/opinion involved in the text is positive, negative or neutral.
ML Techniques

- Machine Learning techniques are often used for Sentiment Analysis

- Most commonly Naive Bayes classifier is used

- Give good results for structured text, containing little noise

- Require heavy processing and may not be useful when we need results on the fly
Bag of Words Model

- Use a word list where each word has been scored positivity/negativity or sentiment strength

- Supervised Learning Method

- Good result for short and unstructured texts

- Overall polarity determined by the aggregate of polarity of all the words in the text
Discourse Relation

- Use discourse relation like connectives and conditionals to incorporate discourse relation (Rule Based)

- Use probabilistic models for identifying elementary discourse units at clausal level and generating trees at the sentence level, using lexical and syntactic information from discourse-annotated corpus (Supervised)

- Use a discourse parser or a dependency parser to identify the scope of the discourse relations and the opinion frames
Overview of Approach

- Presence of linguistic constructs like connectives, modals, conditionals and negation can alter sentiment at the sentence level as well as the clausal or phrasal level

- "@user share 'em! i’m quite excited about Tintin, despite not really liking original comics. Probably because Joe Cornish had a hand in." The overall sentiment of this example is positive, although there is equal number of positive and negative words

- "Think i’ll stay with the whole ’sci-fi’ shit. but this time...a classic movie" Again the overall sentiment influenced due to the presence of "but"
Details
Related Works
Discourse Based Works

- *Maru (2000)* discussed probabilistic models for identifying elementary discourse units at clausal level and generating trees at the sentence level, using lexical and syntactic information from discourse-annotated corpus.

- *Somasundaran et al. (2009)* and *Asher et al. (2008)* discuss some discourse-based supervised and unsupervised approaches to opinion analysis.

- *Taboada et al. (2008)* leverage discourse to identify relevant sentences in the text for sentiment analysis, however focus to adjective alone in relevant portions of text.
Twitter Based Works

- *Alec et al. (2009)* describe a distant supervision-based approach for sentiment classification. They use hashtags in tweets to create training data and implement a multi-class classifier with topic-dependent clusters.

- *Barbosa et al. (2010)* propose an approach to sentiment analysis in Twitter using POS-tagged n-gram features and some Twitter specific features like hashtags.

- *Joshi et al. (2011)* propose a rule-based system, **C-Feel-It**, which classifies a tweet as positive or negative based on the opinion words present in it (Bag of Words Model).

- Most successful of all has been the bag of words model because of the noisy text.
Algorithm
Categorization of Discourse Relations
## Lightweight Twitter Sentiment Analysis

<table>
<thead>
<tr>
<th>Coherence Relations</th>
<th>Conjunctions</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Cause-effect</em></td>
<td>because; and so</td>
</tr>
<tr>
<td><em>Violated Expectations</em></td>
<td>although; but; while</td>
</tr>
<tr>
<td><em>Condition</em></td>
<td>if…(then); as long as; while</td>
</tr>
<tr>
<td><em>Similarity</em></td>
<td>and; (and) similarly</td>
</tr>
<tr>
<td><em>Contrast</em></td>
<td>by contrast; but</td>
</tr>
<tr>
<td><em>Temporal Sequence</em></td>
<td>(and) then; first, second, … before; after; while</td>
</tr>
<tr>
<td><em>Attribution</em></td>
<td>according to …; …said; claim that …; maintain</td>
</tr>
<tr>
<td></td>
<td>that …; stated</td>
</tr>
<tr>
<td><em>Example</em></td>
<td>for example; for instance</td>
</tr>
<tr>
<td><em>Elaboration</em></td>
<td>also; furthermore; in addition; note (furthermore)</td>
</tr>
<tr>
<td></td>
<td>that; (for, in, with) which; who; (for, in, on,</td>
</tr>
<tr>
<td></td>
<td>against, with) whom</td>
</tr>
<tr>
<td><em>Generalization</em></td>
<td>in general</td>
</tr>
</tbody>
</table>

**Table 1**: Contentful Conjunctions used to illustrate Coherence Relations (Wolf *et al.* 2005)
 Relevant Discourse Relations
1. Cause-effect: *(YES! I hope she goes with Chris)* so *(I can freak out like I did with Emmy Awards.)*

2. Violated Expectations: *(i'm quite excited about Tintin)*, **despite** *(not really liking original comics.)*

3. Condition: **If** *(MicroMax improved its battery life)*, *(it wud hv been a gr8 product).*

4. Similarity: *(I lyk Nokia)* and *(Samsung as well).*

5. Contrast: *(my daughter is off school very poorly)*, **but** *(brightened up when we saw you on gmtv today).*

6. Temporal Sequence: *(The film got boring)* **after a while.**

7. Attribution: *(Parliament is a sausage-machine: the world)* **according to** *(Kenneth Clarke).*

8. Example: *(Dhoni made so many mistakes...) for instance, *(he shud’ve let Ishant bowl wn he was peaking).*

9. Elaboration: **In addition** *(to the worthless direction)*, *(the story lacked depth too).*

10. Generalization: **In general,** *(movies made under the RGV banner)* *(are not worth a penny).*

**Table 2:** Examples of Discourse Coherent Relations
# Relation Classes

<table>
<thead>
<tr>
<th>Relations</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conj_Fol</td>
<td><em>but, however, nevertheless, otherwise, yet, still, nonetheless</em></td>
</tr>
<tr>
<td>Conj_Prev</td>
<td><em>till, until, despite, in spite, though, although</em></td>
</tr>
<tr>
<td>Conj_Infer</td>
<td><em>therefore, furthermore, consequently, thus, as a result, subsequently, eventually, hence</em></td>
</tr>
<tr>
<td>Conditionals</td>
<td><em>If</em></td>
</tr>
<tr>
<td>Strong_Mod</td>
<td><em>might, could, can, would, may</em></td>
</tr>
<tr>
<td>Weak_Mod</td>
<td><em>should, ought to, need not, shall, will, must</em></td>
</tr>
<tr>
<td>Neg</td>
<td><em>not, neither, never, no, nor</em></td>
</tr>
</tbody>
</table>

*Table 3: Discourse Relations and Semantic Operators Essential for Sentiment Analysis*
Sketch

- In Step 1, we mark all the conditionals and strong modals which are handled separately by the lexicon-based classifier and the supervised classifier.

- In Step 2 and Step 3, the weight of any word appearing before ConjPrev and after ConjFol or ConjInfer is incremented by 1.

- In Step 4, the polarity of all the words appearing within a window (NegWindow is taken as 5), from the occurrence of a negation operator and before the occurrence of a violating expectation conjunction, are reversed.

- Finally, we get the feature vector $w_{ij}$, $f_{ij}$, $flip_{ij}$ and $hyp_{ij}$ for all the words in the review.

- $f_{ij}$ is the weight of the word $w_{ij}$ in sentence $s_i$, initialized to 1, $flip_{ij}$ is a variable which indicates whether the polarity of $w_{ij}$ should be flipped or not, $hyp_{ij}$ is a variable which indicates the presence of a conditional or a strong modal in $s_i$. 
Feature Vector Classification

• This equation finds the weighted, signed polarity of a review

\[
\text{sign}\left(\sum_{i=1}^{m} \sum_{j=1}^{n_i} f_{ij} \times \text{flip}_{ij} \times p(w_{ij})\right)
\]

where \( p(w_{ij}) = \text{pol}(w_{ij}) \) if \( \text{hyp}_{ij} = 0 \)

\[
= \frac{\text{pol}(w_{ij})}{2} \quad \text{if} \quad \text{hyp}_{ij} = 1
\]

... Equation 1
Evaluation and Results
Dataset 1

- 8507 tweets collected based on a total of around 2000 different entities from 20 different domains

- Domains used for crawling data: Movie, Restaurant, Television, Politics, Sports, Education, Philosophy, Travel, Books Technology, Banking & Finance, Business, Music, Environment, Computers, Automobiles, Cosmetics brands, Amusement parks and Eatables and History

- Tweets processed before evaluation

- All the links in the tweets replaced by #link, all the user id’s in the tweets replaced by #user
Dataset 2

- Twitter API is used to collect another set of 15,214 tweets based on hashtags. Hashtags #positive, #joy, #excited, #happy etc. used to collect tweets bearing positive sentiment, hashtags like #negative, #sad, #depressed, #gloomy, #disappointed etc. used to collect negative tweets.

![Graph 1: Accuracy Comparison between C-Feel-It and Discourse System using Lexicon in Datasets 1 and 2](image)
Dataset 3

- Travel Review Dataset in Balamurali et al. (2011) - contains 595 polarity-tagged documents for each of the positive and negative classes

- Used to determine whether discourse-based approach performs well for structured text as well

<table>
<thead>
<tr>
<th>Sentiment Evaluation Criterion</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Bag-of-Words Model</td>
<td>69.62</td>
</tr>
<tr>
<td>Bag-of-Words Model + Discourse</td>
<td>71.78</td>
</tr>
</tbody>
</table>

*Table 5*: Accuracy Comparison between Bag-of-Words and Discourse System using Lexicon in *Dataset 3*
Results

- In Dataset 1, there is an accuracy improvement of around 4% over C-Feel-It for both 2-class and 3-class classification.

- Discourse system accuracy at 72.81% for 2-class classification is higher than that of the 3-class classification accuracy of 61.31%, showing 3-class classification of tweets is much more difficult than 2-class classification.

- In the Travel review dataset, lexicon-based classification yielded an accuracy improvement of 2% for the discourse model over simple bag-of-words model.
Drawbacks

- Lexicon-based classification suffers from the usage of a generic lexicon in the lexeme space, where it cannot distinguish between the various senses of a word.

- The lexicons do not have entries for the interjections like wow, duh etc. which are strong indicators of sentiment.

- The scope of the discourse marker has been heuristically taken till the sentence boundary or till the next discourse marker. Consider the sentence, "I wanted to follow my dreams and ambitions despite all the obstacles, but I did not succeed." Here want and ambition will get the polarity +2 each, as they appear before despite; obstacle will get a polarity -1 and not succeed will get a polarity -2. Thus the overall polarity is +1, whereas the overall sentiment should be negative.
References
Alec, G.; Lei, H.; and Richa, B. Twitter sentiment classification using distant supervision. Technical report, Standford University. 2009


Barbosa, L., and Feng, J. Robust sentiment detection on twitter from biased and noisy data. In 23rd International Conference on Computational Linguistics: Posters, 36–44. 2010

Bermingham, A., and Smeaton, A. Classifying sentiment in microblogs: is brevity an advantage ACM 1833–1836. 2010


Questions?