Contextual Phrase-Level Polarity Analysis

Apoorv Agarwal
Fadi Biadsy
Kathleen McKeown
Sentiment Analysis

Subjectivity Analysis

Polarity Analysis

Prior Polarity Lexicon

Incorporate Context

Document Sentence Phrase

Positive Negative Neutral
Our Contributions

- Use Dictionary of Affect in Language (DAL) to suggest a scoring scheme to enable automatic scoring of majority content words.

- Propose a feature that is a combination of the 3 scores given to words in DAL that differentiates between high and low subjective words.

- Suggest new contextual features based on N-gram of polar constituents of subjective phrases.
A sentence may have positive, negative and neutral opinions
Challenges

Greece has great food but I find the strike to be annoying

- A sentence may have positive, negative and neutral opinions
- It is difficult to accurately mark subjective phrase boundaries
Challenges

- A sentence may have positive, negative and neutral opinions
- It is difficult to accurately mark subjective phrase boundaries
- Negations and connectives change prior polarity
Dictionary of Affect in Language (DAL)

- 8742 English word dictionary to measure emotional meaning of texts
- Assigns 3 scores to each word on a scale of 1(low) - 3(high)
  - Pleasantness (ee)
  - Activeness (aa)
  - Imagery (ii)

<table>
<thead>
<tr>
<th>Word</th>
<th>ee</th>
<th>aa</th>
<th>ii</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affect</td>
<td>1.75</td>
<td>1.85</td>
<td>1.60</td>
</tr>
<tr>
<td>Affection</td>
<td>2.77</td>
<td>2.25</td>
<td>2.00</td>
</tr>
<tr>
<td>Slug</td>
<td>1.00</td>
<td>1.18</td>
<td>2.40</td>
</tr>
<tr>
<td>Energetic</td>
<td>2.25</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Flower</td>
<td>2.75</td>
<td>1.07</td>
<td>3.00</td>
</tr>
</tbody>
</table>

3 scores are uncorrelated (Cowie et. al., 2001)
Contains different scores for inflectional forms
8742 English word dictionary to measure emotional meaning of texts
Assigns 3 scores to each word on a scale of 1(low) - 3(high)
- Pleasantness ($ee$)
- Activeness ($aa$)
- Imagery ($ii$)

<table>
<thead>
<tr>
<th>Word</th>
<th>$ee$</th>
<th>$aa$</th>
<th>$ii$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affect</td>
<td>1.75</td>
<td>1.85</td>
<td>1.60</td>
</tr>
<tr>
<td>Affection</td>
<td>2.77</td>
<td>2.25</td>
<td>2.00</td>
</tr>
<tr>
<td>Slug</td>
<td>1.00</td>
<td>1.18</td>
<td>2.40</td>
</tr>
<tr>
<td>Energetic</td>
<td>2.25</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Flower</td>
<td>2.75</td>
<td>1.07</td>
<td>3.00</td>
</tr>
</tbody>
</table>

3 scores are uncorrelated (Cowie et. al., 2001)
**Dictionary of Affect in Language (DAL)**

- 8742 English word dictionary to measure emotional meaning of texts
- Assigns 3 scores to each word on a scale of 1(low) - 3(high)
  - Pleasantness (*ee*)
  - Activeness (*aa*)
  - Imagery (*ii*)

<table>
<thead>
<tr>
<th>Word</th>
<th><em>ee</em></th>
<th><em>aa</em></th>
<th><em>ii</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Affect</td>
<td>1.75</td>
<td>1.85</td>
<td>1.60</td>
</tr>
<tr>
<td>Affection</td>
<td>2.77</td>
<td>2.25</td>
<td>2.00</td>
</tr>
<tr>
<td>Slug</td>
<td>1.00</td>
<td>1.18</td>
<td>2.40</td>
</tr>
<tr>
<td>Energetic</td>
<td>2.25</td>
<td>3.00</td>
<td>3.00</td>
</tr>
<tr>
<td>Flower</td>
<td>2.75</td>
<td>1.07</td>
<td>3.00</td>
</tr>
</tbody>
</table>

- 3 scores are uncorrelated (Cowie et. al., 2001)
- Contains different scores for inflectional forms
The Taj has great food but I ....
... but I found their service to be lacking

CLASSIFIER

Positive
Negative
Neutral
Multi-Perspective Question Answering (MPQA) corpus

Gold Standard: Manual annotation tag (positive, negative, neutral) given to subjective phrases in the corpus
**Basic Scoring Scheme**

- **Phrase**: $W_1, W_2, \ldots, W_n$
- **DAL**: 74%
- **WordNet**
  - Found
  - Not Found
- **Negation Machine**
  - 97% words scored
- **DAL**: 23%
- **Produces a list of**
  - Antonyms
  - Synonyms
- **3% words not scored**
  - Eg: ulterior

- **Searched in**
  - 97% words scored
  - Negation Machine
  - DAL

---

Contextual Phrase-Level Polarity Analysis
A. Agarwal, F. Biadsy, K. McKeown
Activation - Evaluation (AE) space score (Cowie et. al. 2001)

\[ AE = \sqrt{ee^2 + aa^2} \]
Norm

**Activation - Evaluation (AE) space score (Cowie et. al. 2001)**

\[ AE = \sqrt{(ee^2 + aa^2)} \]

**Subjectivity \( \propto \) \( \frac{1}{Imagery} \)**

Eg: *goodies vs good*
**Norm**

Activation - Evaluation (AE) space score (Cowie et. al. 2001)

\[ AE = \sqrt{(ee^2 + aa^2)} \]

Subjectivity \( \propto \frac{1}{\text{Imagery}} \)

Eg: goodies vs good

\[ \text{norm} = \frac{\sqrt{ee^2 + aa^2}}{ii} \]
Contextual Features

This announcement was met with *unanimous* condemnation

\[\text{[NP]}_{\text{NEU}} \quad \text{[VP]}_{\text{NEU}} \quad \text{[PP]}_{\text{TARGET}}^{\text{NEG}}\]

Subjective phrase as marked in the corpus

Expanded to a chunk (our TARGET phrase)

Lexical Features

\ldots \ldots \text{ as in previous work (Wilson et. al., 2005)}

Syntactic Features

- N-grams over polar chunks, e.g. bigram: \([VP]_{\text{NEU}}[PP]_{\text{TARGET}}^{\text{NEG}}\)
- Minimum and maximum ee scores of chunks in the target phrase
- Count of syntactic categories of chunks associated with their prior polarity to the left and right of target phrase and in the target phrase
Experimental Set-up

- **MPQA corpus**
  - # of positive phrases: 2779
  - # of negative phrases: 6471
  - # of neutral phrases: 7993

- Random down sampling to get a balanced data-set

- Logistic classifier, 10-fold cross validation

- Baseline: Word N-gram
3-way Classifier

Positive vs. Negative vs. Neutral

% Accuracy

<table>
<thead>
<tr>
<th></th>
<th>DAL scores</th>
<th>+ POS</th>
<th>+ Chunks</th>
<th>+ N-gram (All)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chunks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-gram</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend

- Word
- N-gram baseline

Chunking performs very well!
2-way Classifier

Positive vs. Negative

DAL scores perform very well!
Conclusion

- Introduce completely automated system for scoring subjective phrases using DAL and WordNet

- Introduce new contextual features based on N-grams of constituents

- Don’t need accurate phrase boundary

- Limitation: do not handle polysemy
Future Work

- Study if there’s a correlation between subjectivity and polarity

- Use same framework for subjectivity and intensity analysis by tagging chunks with the imagery and activeness score respectively