Context-Aware Sentiment Detection

Albert Weichselbraun
February 10, 2011
Agenda

- Motivation
  - On the Need for Contextualization
  - Indicators for Missing Context

- Method
  - Context-Aware Sentiment Detection
  - Creation of Contextualized Sentiment Lexicons
  - Example
  - Cross-corpus Contextualized Sentiment Lexicons

- Evaluation

- Outlook and Conclusions
Motivation

- Pang et al. (2002): state of the art machine learning approaches do not unfold their full potential when applied to sentiment detection
- Lexicon-based approach
  - no labeled training corpus necessary
  - applicable across domains
  - throughput
### Positive
- The repair of my car was satisfying.
- This movie’s plot is unpredictable.
- The long peace brought wealth and safety to the people.

### Negative
- I had many complaints after my camera’s repair.
- The breaks of this car are unpredictable.
- This peace is a lie.
Frequency Diagram

Idealized Neutral Term

And

Motivation
Frequency Diagram

**Idealized Ambiguous Term**

**Accident**
Frequency Diagram

- Idealized Smooth Ambiguous Term
- Expensive

Motivation
Use a contextualized sentiment lexicon
- Based on ordinary sentiment lexicons
- Contains stable sentiment terms and ambiguous terms
- Uses context terms for disambiguation

Derived from online reviews (Amazon, TripAdvisor)
Refined Sentiment Detection

\[ s_{total} = \sum_{t_i \in \text{doc}} n(t_i)[s(t_i) + s'(t_i | c)] \]

with

\[ n(t_i) = \begin{cases} 
-1.0 & \text{if } t_i \text{ has been negated} \\
+1.0 & \text{otherwise.} 
\end{cases} \]

Context Detection

\[ c = \{ c_1, \ldots c_n \} \]

\[ p(C_+ | c) = \frac{p(C_+) \cdot \prod_{i=1}^{n} p(c_i | C_+)}{\prod_{i=1}^{n} p(c_i)} \]
Method - Contextualized Sentiment Lexicon

Corpus Creation

Sentiment Lexicon

Ambiguous Term Detection (s')

Training Corpus

Ambiguous Terms

Context Term Detection (c)

Naive Bayes

Contextualized Sentiment Lexicon

Sentiment Detection

Text Document

Sentiment Detection

Sentiment Value
Method - Contextualized Sentiment Lexicon

- Identify ambiguous terms \( (s') \)
  → frequency diagrams

\[
\begin{align*}
\sigma_i & \geq \nu \quad (3) \\
\mu_i + \sigma_i & \geq \omega \quad (4a) \\
\mu_i - \sigma_i & \leq -\omega \quad (4b)
\end{align*}
\]

- Learn context terms \( (c) \) for disambiguation
  → conditional probabilities

- Recalculate the sentiment value of the contextualized sentiment terms
The service staff was *friendly*. They accomplished the *repair* of my car’s motor very *quickly*. After driving it for another three months I can say that the motor is as *reliable* as it was before.

<table>
<thead>
<tr>
<th>positive context terms</th>
<th>negative context terms</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>reliable</em></td>
<td>slowly</td>
</tr>
<tr>
<td>long-lasting</td>
<td>re-do</td>
</tr>
<tr>
<td>affordable</td>
<td>unreliable</td>
</tr>
<tr>
<td>pick-up-service</td>
<td>waiting</td>
</tr>
<tr>
<td>replacement-car</td>
<td>expensive</td>
</tr>
<tr>
<td>cooperative</td>
<td>cheater</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
The service staff was *friendly*. They accomplished the *repair* of my car’s motor very *quickly*. After driving it for another three months I can say that the motor is as *reliable* as it was before.

| Context Term (c_i) | P(C_+|c_i) |
|-------------------|-----------|
| reliable          | 0.80      |
| friendly          | 0.70      |
| quickly           | 0.65      |

⇒ repair is used in a positive context → positive sentiment
## Method - Real World Examples

<table>
<thead>
<tr>
<th>Ambiguous Term</th>
<th>$SV_{\text{orig}}$</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>busy</td>
<td>1</td>
<td>The hotel is located on a <strong>busy</strong> road.</td>
</tr>
<tr>
<td>complaint</td>
<td>-1</td>
<td>My <strong>only complaint</strong> would be the service.</td>
</tr>
<tr>
<td>cool</td>
<td>-1</td>
<td>Our room felt like a <strong>really cool</strong> European apartment with a rooftop terrace.</td>
</tr>
<tr>
<td>expensive</td>
<td>-1</td>
<td>The room was one of the more <strong>expensive</strong> hotels in Vienna but still <strong>excellent</strong>.</td>
</tr>
<tr>
<td>quality</td>
<td>1</td>
<td><strong>Poor quality</strong> copies with one edge always dark.</td>
</tr>
<tr>
<td>better</td>
<td>1</td>
<td>Let’s <strong>hope</strong> they work <strong>better</strong>.</td>
</tr>
<tr>
<td>cost</td>
<td>-1</td>
<td>Toner <strong>cost</strong> is way <strong>behind</strong> competitors.</td>
</tr>
</tbody>
</table>
Cross-corpus Contextualized Sentiment Lexicons

**Helpful**

- $t/C_+$

**Harmful**

- $t/C'_-$

---

**Relative Frequency**

- $C_+$

- $C'_-$

- $C_-$

- $C'_+$

**Sentiment Value**

- $-1.0$

- $-0.5$

- $0.0$

- $0.5$

- $1.0$

---

**Method**

16 / 28
Cross-corpus Contextualized Sentiment Lexicons

Three-step process

- Determine the helpfulness of all context terms
- Discard harmful context terms
- Merge remaining context terms into a large contextualized lexicon
Step 1 - Determine the helpfulness of context terms
Cross-corpus Contextualized Sentiment Lexicons

Step 2 - Discard harmful context terms

Diagram:
- Harmful
- Helpful
- Neutral
Cross-corpus Contextualized Sentiment Lexicons

Step 3 - Merging
Evaluation - Approach

Evaluations

1. Comparison to a baseline
   – Do we outperform a lexicon-based approach which does not consider context?

2. Intra-domain sentiment detection
   – Does the removal of unstable sentiment terms has a positive effect?

3. Cross-domain sentiment detection
   – Determine the cross-domain performance of a generic contextualized sentiment lexicon.

4. Comparison to a machine learning approach
   – Intra-domain and cross-domain performance.
Evaluation - Setting

- Method: 10-fold cross validation
- Test corpora:
  - Equal number of positive and negative reviews.
  - Amazon: 2,500 reviews
  - TripAdvisor: 1,800 reviews
### Evaluation - Context Aware Sentiment Detection

#### Corpus: Amazon

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th></th>
<th>Context Aware</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{R}$</td>
<td>$\bar{P}$</td>
<td>$\bar{F}_1$</td>
<td>$\bar{R}$</td>
</tr>
<tr>
<td>Pos</td>
<td>0.80</td>
<td>0.64</td>
<td>0.71</td>
<td>0.75</td>
</tr>
<tr>
<td>Neg</td>
<td>0.53</td>
<td>0.74</td>
<td>0.62</td>
<td>0.71</td>
</tr>
</tbody>
</table>

#### Corpus: TripAdvisor

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th></th>
<th>Context Aware</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{R}$</td>
<td>$\bar{P}$</td>
<td>$\bar{F}_1$</td>
<td>$\bar{R}$</td>
</tr>
<tr>
<td>Pos</td>
<td>0.96</td>
<td>0.60</td>
<td>0.74</td>
<td>0.97</td>
</tr>
<tr>
<td>Neg</td>
<td>0.34</td>
<td>0.90</td>
<td>0.49</td>
<td>0.46</td>
</tr>
</tbody>
</table>
### Evaluation - Intra-Domain Performance

#### Test corpus: Amazon

<table>
<thead>
<tr>
<th></th>
<th>Domain-specific (Amazon)</th>
<th>Generic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{R}$</td>
<td>$\bar{P}$</td>
<td>$\bar{F}_1$</td>
</tr>
<tr>
<td><strong>Pos</strong></td>
<td>0.75</td>
<td>0.75</td>
<td>0.74</td>
</tr>
<tr>
<td><strong>Neg</strong></td>
<td>0.71</td>
<td>0.79</td>
<td>0.73</td>
</tr>
</tbody>
</table>

#### Test corpus: TripAdvisor

<table>
<thead>
<tr>
<th></th>
<th>Domain-specific (TripAdvisor)</th>
<th>Generic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{R}$</td>
<td>$\bar{P}$</td>
<td>$\bar{F}_1$</td>
</tr>
<tr>
<td><strong>Pos</strong></td>
<td>0.97</td>
<td>0.66</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Neg</strong></td>
<td>0.46</td>
<td>0.95</td>
<td>0.61</td>
</tr>
</tbody>
</table>
### Evaluation - Cross Domain Performance

#### Test corpus: Amazon

<table>
<thead>
<tr>
<th></th>
<th>Domain-specific (TripAdvisor)</th>
<th>Generic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R$</td>
<td>$P$</td>
</tr>
<tr>
<td>Pos</td>
<td>0.76</td>
<td>0.67</td>
</tr>
<tr>
<td>Neg</td>
<td>0.58</td>
<td>0.73</td>
</tr>
</tbody>
</table>

#### Test corpus: TripAdvisor

<table>
<thead>
<tr>
<th></th>
<th>Domain-specific (Amazon)</th>
<th>Generic</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R$</td>
<td>$P$</td>
<td>$F_1$</td>
</tr>
<tr>
<td>Pos</td>
<td>0.84</td>
<td>0.69</td>
<td>0.75</td>
</tr>
<tr>
<td>Neg</td>
<td>0.58</td>
<td>0.8</td>
<td>0.66</td>
</tr>
</tbody>
</table>
### Evaluation - Naïve Bayes (NLTK)

<table>
<thead>
<tr>
<th>Training</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>TripAdvisor</td>
<td>Test</td>
</tr>
<tr>
<td>⊕</td>
<td>TripAdvisor</td>
</tr>
<tr>
<td></td>
<td>⊕ 87</td>
</tr>
<tr>
<td></td>
<td>⊕ 89</td>
</tr>
<tr>
<td></td>
<td>⊕ 75</td>
</tr>
<tr>
<td></td>
<td>⊕ 60</td>
</tr>
<tr>
<td>Amazon</td>
<td></td>
</tr>
</tbody>
</table>
Summary

- Lexicon-based approaches:
  - Simple, no labelled data required
  - Applicable across domains
  - High throughput
  - Can serve as a baseline

- Machine Learning approaches:
  - Powerful, but *domain-specific*
  - Require labelled training data

→ The introduced approach combines these advantages (cross-domain, high throughput, high performance)
Outlook and Conclusions

- Considering context in sentiment detection
- Creation cross-domain contextualized sentiment lexicons
- Outperforms generic approaches
- Future work:
  - Different context scopes (paragraph, documents, text windows)
  - Consider other machine learning approaches