ASPECT-LEVEL SENTIMENT ANALYSIS: Implicit Features Detection in Consumers Reviews

Kim Schouten
(schouten@ese.eur.nl)

Flavius Frasincar
(frasincar@ese.eur.nl)

Franciska de Jong
(f.m.g.dejong@eshcc.eur.nl)
Introduction

• People like to voice their opinion
• This is especially true on the Web
• We specifically focus on user-generated reviews
  • (e.g., Amazon, Yelp, etc.)
• These data can be harnessed for business purposes
  • For consumers, it is an important source of information when looking to purchase something
  • For producers, it is a valuable source of consumer feedback
    • Usually honest
    • Free!
Introduction

• The aim of Aspect-Level Sentiment Analysis is to find a quadruple

\[(s, g, h, t)\]

• \(s\) = sentiment score
• \(g\) = the target on which the sentiment is expressed
• \(h\) = the holder: the one expressing the sentiment
• \(t\) = the time when the sentiment was expressed
Introduction

• Traditionally, sentiment analysis is performed at document or sentence level
• Assumption that only one topic is discussed there
• Why not look for the actual topics being discussed and attach sentiment scores to those?
• This is called aspect-based sentiment analysis
Introduction

• Usually one document or sentence describes one entity
• However, multiple facets or aspects are described for that entity within a document, often with conflicting sentiment scores:
  • “The pizza was perfect, but the waiters were rude”
• Aspects can be fine-grained (pro: very detailed)
  • “pizza” and “waiters”
• Or coarse-grained (pro: easier to compare across reviews)
  • “food” and “service”
• Aspect-based sentiment analysis has two main tasks: finding aspects, and finding their sentiment scores
Introduction

• Both fine-grained and coarse-grained aspects can be implicit

• Since coarse-grained aspects are more general and abstract, these tend to be implied more often than fine-grained aspects

• Coarse-grained aspects are also referred to as aspect categories
ASPECT-LEVEL SENTIMENT ANALYSIS:
Coarse-grained Aspect Detection in Consumers Reviews

Kim Schouten
(schouten@ese.eur.nl)

Flavius Frasincar
(frasincar@ese.eur.nl)

Franciska de Jong
(f.m.g.dejong@eshcc.eur.nl)
Coarse-grained Aspect Detection Method

• Our intuition is that certain words are expected to co-occur a lot with certain coarse aspects (e.g., ‘delicious’ -> food, ‘expensive’ -> price)

• Hence, if we encounter those words, we can predict that these related coarse aspects are present in that sentence

• By looking at the co-occurrence of each word in a sentence with each of the possible coarse aspects, we can compute a score for each aspect
Our agreed favorite is the orrechiete with sausage and chicken (usually the waiters are kind enough to split the dish in half so you get to sample both meats).
Training Phase – Creating $C$

Annotated sentences $\rightarrow$ Co-occurrence matrix

$$freq(a, w) / freq(w)$$

$$= C$$
Training Phase – Pruning \( C \)

**Co-occurrence matrix**

Using simple linear search, two threshold variables (\( \text{min}_\text{cooc} \) and \( \text{min}_\text{freq} \)) are optimized

\[
\frac{\text{freq}(a, w)}{\text{freq}(w)} > \text{min}_\text{cooc} \\
\text{freq}(w) > \text{min}_\text{freq}
\]

Only entries that satisfy both constraints are retained in \( C \)
Processing unseen data

Simply put: $A = SC$

Unseen sentence  Co-occurrence matrix

$T$

<table>
<thead>
<tr>
<th>Words (lemmas)</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent in sentence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present in sentence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Aspects

<table>
<thead>
<tr>
<th>Words (lemmas)</th>
<th>...</th>
<th>...</th>
<th>...</th>
<th>...</th>
<th>...</th>
<th>...</th>
<th>...</th>
<th>...</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent in sentence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present in sentence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\approx$

Aspect scores

Any aspect that has a score is assigned to the sentence
Results

• Two data sets with restaurant reviews
  • First one has ~3800 sentences with 4 different aspect categories
  • Second one has ~1300 sentences with 13 different aspect categories

• One data set with laptop reviews
  • (~1700 sentences with 82 different aspect categories)

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>$F_1$</th>
<th>SemEval best $F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restaurants 1</td>
<td>66.7%</td>
<td>73.9%</td>
<td>70.1%</td>
<td>85.3%</td>
</tr>
<tr>
<td>Restaurants 2</td>
<td>64.7%</td>
<td>62.2%</td>
<td>63.4%</td>
<td>61.9%</td>
</tr>
<tr>
<td>Laptops</td>
<td>40.1%</td>
<td>42.3%</td>
<td>41.4%</td>
<td>49.6%</td>
</tr>
</tbody>
</table>

• Very good performance for a relatively simple method that uses no advanced machine learning techniques
**Alternative SVM**

- Trained a basic bag-of-words model SVM classifier
  - One (binary) SVM model for each aspect
  - Per sentence determine whether that coarse aspect is present or not
- Results on the same data:

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>$F_1$</th>
<th>Co-occurrence $F_1$</th>
<th>SemEval best $F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restaurants 1</td>
<td>81.3%</td>
<td>75.2%</td>
<td>78.1%</td>
<td>70.1%</td>
<td>85.3%</td>
</tr>
<tr>
<td>Restaurants 2</td>
<td>77.5%</td>
<td>57.7%</td>
<td>66.1%</td>
<td>63.4%</td>
<td>61.9%</td>
</tr>
<tr>
<td>Laptops</td>
<td>60.9%</td>
<td>34.7%</td>
<td>44.2%</td>
<td>41.4%</td>
<td>49.6%</td>
</tr>
</tbody>
</table>

- Very good performance for a relatively simple method that uses no advanced machine learning techniques
On-going and future work

• Use of ontologies or other knowledge bases to make use of domain knowledge
  • Move towards a more concept-driven, or semantics-driven form of aspect-level sentiment analysis

• Use of more advanced machine learning techniques
  • Latent Dirichlet Allocation might be used to find coarse aspect
  • Recurrent Neural Networks (with LSTM)
QUESTIONS?