Predicting the Semantic Orientation of Adjective

Vasileios Hatzivassiloglou and Kathleen R. McKeown
Presented By Yash Satsangi
Aim

• To validate that conjunction put constraints on conjoined adjectives and this information can be used to detect their semantic orientation

• Based on above information cluster adjectives into two groups representing adjectives with positive and negative orientation.
Constraint On Conjoined Adjectives

• Validate constraints from conjunction on positive/negative semantic orientation of adjectives
• Honest ‘and’ peaceful – same orientation
• Talented ‘but’ Irresponsible – opposite orientation
• Thus conjunction affect semantic orientation
• Synonyms may have same semantic orientation
• Antonyms may have opposite semantic orientation (hot and cold).
Approach

• Extract conjunction from corpus with their morphological relation
• A log-linear regression model to predict orientation of two different adjectives
• A clustering algorithm separates the adjectives into two subset of same or opposite orientation.
Data

• 21 million word 1987 Wall Street Journal Corpus annotated with part-of-speech tags
• Remove adjectives occurring less than 20 times and those which had no orientation.
• Manually assign orientation to each adjective based on use of adjective
• Multiple validation of labeled adjectives was done.
• Final Set – 1336 adjective – 657 positive and 679 negative – with 96.97% inter-reviewer agreement.
Validating the Hypothesis

• Run parser on 21 million words dataset to get 15,048 conjunction tokens involving 9,296 pairs of distinct adjective pairs.

• Each conjunction was classified into:
  1.) conjunction used ; 2.) type of modification ;
  3.) modified noun

• Count percentage of conjunction in each category with adjectives of same or different orientation
## Validating Hypothesis

<table>
<thead>
<tr>
<th>Conjunction category</th>
<th>Conjunction types analyzed</th>
<th>% same-orientation (types)</th>
<th>% same-orientation (tokens)</th>
<th>P-Value (for types)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All conjunctions</td>
<td>2,748</td>
<td>77.84%</td>
<td>72.39%</td>
<td>&lt; 1 ∙ 10⁻¹⁶</td>
</tr>
<tr>
<td>All <em>and</em> conjunctions</td>
<td>2,294</td>
<td>81.73%</td>
<td>78.07%</td>
<td>&lt; 1 ∙ 10⁻¹⁶</td>
</tr>
<tr>
<td>All <em>or</em> conjunctions</td>
<td>305</td>
<td>77.05%</td>
<td>60.97%</td>
<td>&lt; 1 ∙ 10⁻¹⁶</td>
</tr>
<tr>
<td>All <em>but</em> conjunctions</td>
<td>214</td>
<td>30.84%</td>
<td>25.94%</td>
<td>2.09 ∙ 10⁻⁸</td>
</tr>
<tr>
<td>All attributive <em>and</em> conjunctions</td>
<td>1,077</td>
<td>80.04%</td>
<td>76.82%</td>
<td>&lt; 1 ∙ 10⁻¹⁶</td>
</tr>
<tr>
<td>All predicative <em>and</em> conjunctions</td>
<td>860</td>
<td>84.77%</td>
<td>84.54%</td>
<td>&lt; 1 ∙ 10⁻¹⁶</td>
</tr>
<tr>
<td>All appositive <em>and</em> conjunctions</td>
<td>30</td>
<td>70.00%</td>
<td>63.64%</td>
<td>0.04277</td>
</tr>
</tbody>
</table>
Validating Hypothesis

• For almost all the cases p-values are low. Hence the statistics are significant.
• There are very small differences in behavior of conjunctions
• ‘and’ usually joins adjectives of same orientation
• ‘but’ is opposite and joins adjectives of different orientation
Baseline Method to Predict Link

• Simple baseline method – to call each link as same orientation will give 77.84% accuracy
• Adjective con-joined by ‘but’ are mostly of opposite orientation
• Morphological relationship (e.g. : adequate-inadequate) contains information as well
Better Idea – Use regression model

• Train a log Linear Regression Model
  \[ \eta = w^T x \]

• \( x \) is the observed count of adjective pair in various conjunction category.

• To avoid over fitting they used subsets of data.

• Process of iterative stepwise refinement leads to building up of final model
Result of Prediction

<table>
<thead>
<tr>
<th>Prediction method</th>
<th>Morphology used?</th>
<th>Accuracy on reported same-orientation links</th>
<th>Accuracy on reported different-orientation links</th>
<th>Overall accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always predict same orientation</td>
<td>No</td>
<td>77.84%</td>
<td>—</td>
<td>77.84%</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>78.18%</td>
<td>97.06%</td>
<td>78.86%</td>
</tr>
<tr>
<td>But rule</td>
<td>No</td>
<td>81.81%</td>
<td>69.16%</td>
<td>80.82%</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>82.20%</td>
<td>78.16%</td>
<td>81.75%</td>
</tr>
<tr>
<td>Log-linear model</td>
<td>No</td>
<td>81.53%</td>
<td>73.70%</td>
<td>80.97%</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>82.00%</td>
<td>82.44%</td>
<td>82.05%</td>
</tr>
</tbody>
</table>

- Log Linear Regression models performs slightly better than baseline
- Mainly used to group adjectives into same group
Grouping Adjectives into same pack

• Log Linear model generates a dissimilarity score between two adjective between 0 and 1
• Same and different adjectives thus form a graph
• Iterative Optimization procedure is used to partition graph into clusters.
• Minimize :
  \[ \Phi(P) = \sum_{i=1}^{2} \left( \frac{1}{|C_i|} \sum_{x,y \in C_i} d(x, y) \right) \]
• Hierarchical Clustering
Labeling Clusters

• Same authors in ‘95 showed that a semantically unmarked member of gradable adjectives is the most frequent.
• Now semantic markedness exhibit a strong correlation with orientation
• Unmarked member always have positive orientation
• So group with higher average frequency contains positive terms.
Evaluating Clustering of Adjectives

• Separate the Adjective set A into training and testing groups by selecting a parameter named $\alpha$.

• $\alpha$ is the parameter which decides the number of link of each adjective in the selected training and test set.

• Higher $\alpha$ creates subset of A such that more adjectives are connected to each other.
Clustering Results

| $\alpha$ | Number of adjectives in test set ($|A_\alpha|$) | Number of links in test set ($|L_\alpha|$) | Average number of links for each adjective | Accuracy | Ratio of average group frequencies |
|---|---|---|---|---|---|
| 2 | 730 | 2,568 | 7.04 | 78.08% | 1.8699 |
| 3 | 516 | 2,159 | 8.37 | 82.56% | 1.9235 |
| 4 | 369 | 1,742 | 9.44 | 87.26% | 1.3486 |
| 5 | 236 | 1,238 | 10.49 | 92.37% | 1.4040 |

- Highest accuracy obtained when highest number of links were present.
- Every time - ratio of group frequency correctly identified the positive subgroup
### Classification Example

<table>
<thead>
<tr>
<th>Classified as positive:</th>
<th>Classified as negative:</th>
</tr>
</thead>
<tbody>
<tr>
<td>bold decisive disturbing generous good</td>
<td>ambiguous cautious cynical evasive</td>
</tr>
<tr>
<td>honest important large mature patient</td>
<td>harmful hypocritical inefficient insecure</td>
</tr>
<tr>
<td>peaceful positive proud sound</td>
<td>irrational irresponsible minor outspoken</td>
</tr>
<tr>
<td>stimulating straightforward strange</td>
<td>pleasant reckless risky selfish tedious</td>
</tr>
<tr>
<td>talented vigorous witty</td>
<td>unsupported vulnerable wasteful</td>
</tr>
</tbody>
</table>
Performance

• To measure performance of algorithm a series of simulation experiments were run.
• Parameter P measures how well each link is predicted independently – Precision
• Parameter k – number of distinct adjective each adjectives appears in conjunction with.
• Generate Random Graph between nodes such that each node participated in k links and P% of all nodes connected same orientation and classify them
Results

(a) $P = 0.75$

(b) $P = 0.8$

(c) $P = 0.85$

(d) $P = 0.9$
Conclusion

• A good ‘and’ comprehensive method for classification of semantic orientation of adjectives.
• Can be used to find antonyms without accessing any semantic information
• Can be extended to nouns and verbs.
Thank You!