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

Conjunction category	Conjunction types analyzed	% same-orientation (types)	% same-orientation (tokens)	P-Value (for types)
All conjunctions	2,748	77.84%	72.39%	$< 1 \cdot 10^{-16}$
All *and* conjunctions	2,294	81.73%	78.07%	$< 1 \cdot 10^{-16}$
All *or* conjunctions	305	77.05%	60.97%	$< 1 \cdot 10^{-16}$
All *but* conjunctions	214	30.84%	25.94%	$2.09 \cdot 10^{-8}$
All attributive *and* conjunctions	1,077	80.04%	76.82%	$< 1 \cdot 10^{-16}$
All predicative *and* conjunctions	860	84.77%	84.54%	$< 1 \cdot 10^{-16}$
All appositive *and* conjunctions	30	70.00%	63.64%	0.04277

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

$$\eta = \mathbf{w}^T \mathbf{x}$$

- **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

Prediction method	Morphology used?	Accuracy on reported same-orientation links	Accuracy on reported different-orientation links	Overall accuracy
Always predict same orientation	No	77.84%	–	77.84%
	Yes	78.18%	97.06%	78.86%
But rule	No	81.81%	69.16%	80.82%
	Yes	82.20%	78.16%	81.75%
Log-linear model	No	81.53%	73.70%	80.97%
	Yes	82.00%	82.44%	82.05%

- 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 :

$$\Phi(\mathcal{P}) = \sum_{i=1}^{2} \left(\frac{1}{|C_i|} \sum_{\substack{x,y \in C_i \\ x \neq y}} d(x,y) \right)$$

- 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 α.
- α is the parameter which decides the number of link of each adjective in the selected training and test set.
- Higher α creates subset of A such that more adjectives are connected to each other.

Clustering Results

α	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

Classified as positive:

bold decisive disturbing **generous good honest important large mature patient peaceful positive proud sound stimulating straightforward** strange **talented vigorous witty**

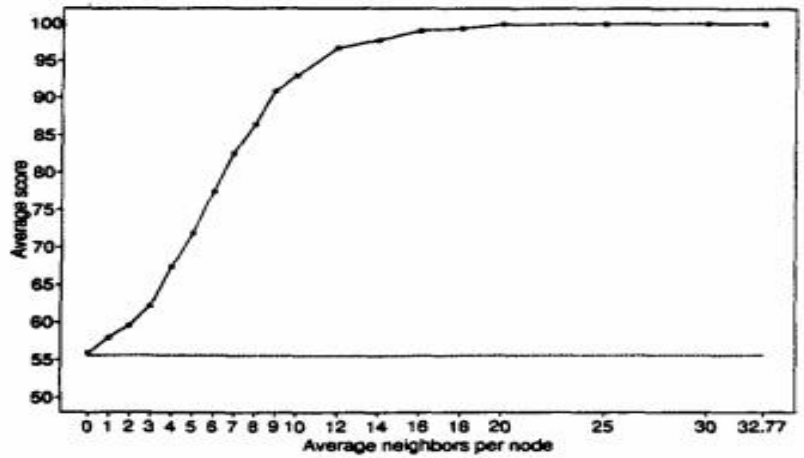
Classified as negative:

ambiguous cautious **cynical evasive harmful hypocritical inefficient insecure irrational irresponsible minor** outspoken pleasant **reckless risky selfish tedious unsupported vulnerable wasteful**

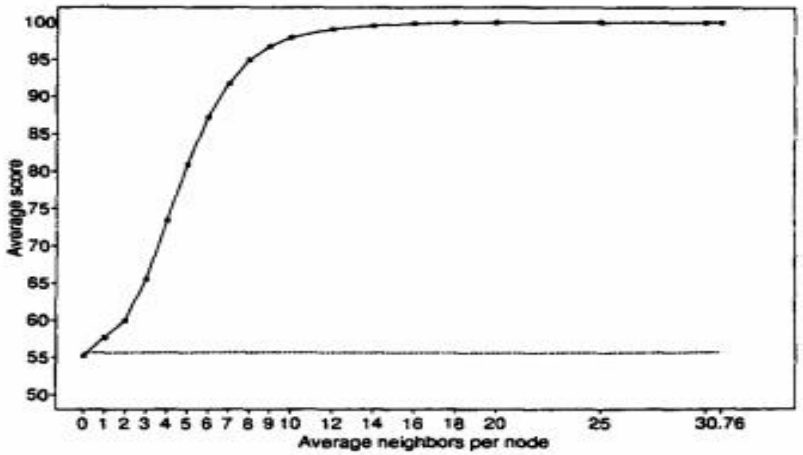
Performance

- 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

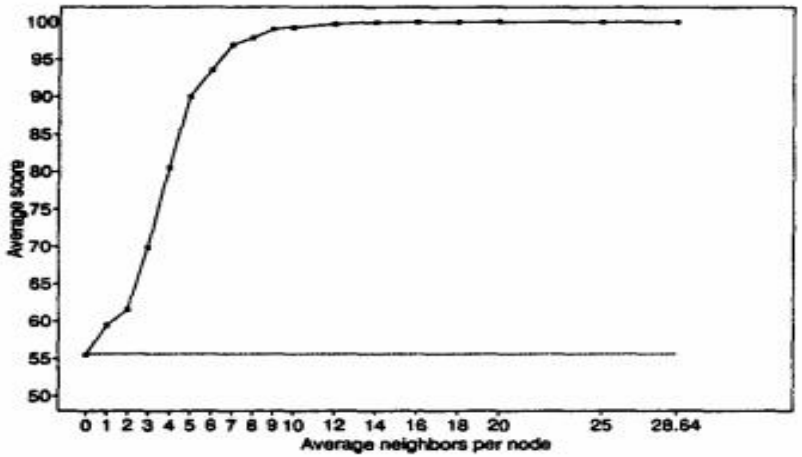
Results



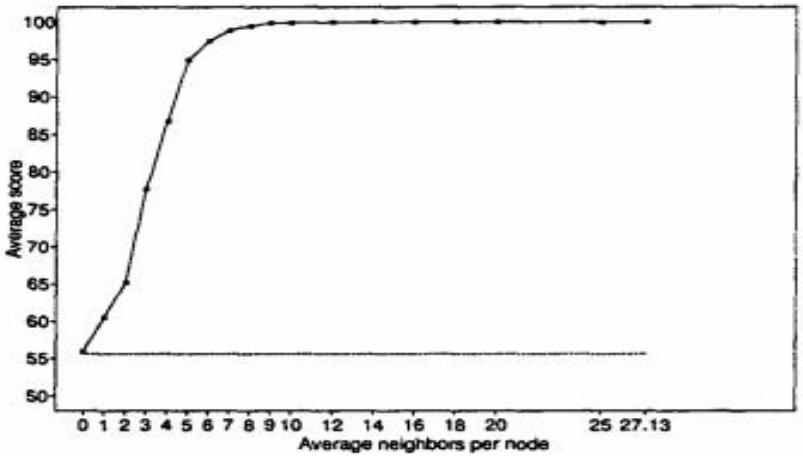
(a) $P = 0.75$



(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!