Q-WordNet: Extracting Polarity from WordNet Senses

Rodrigo Agerri*, Ana García-Serrano*

* GSI Group, Universidad Politécnica de Madrid (UPM)
and
HSLT Group, Vicomtech Research Centre
Mikeletegi Pasealekua 57, 20009 Donostia-San Sebastián (Spain)
ragerri@vicomtech.org

* NLP & IR Group UNED
Juan del Rosal 16, 28040 Madrid (Spain)
agarcia@lsi.uned.es

Abstract

This paper presents Q-WordNet, a lexical resource consisting of WordNet senses automatically annotated by positive and negative polarity. Polarity classification amounts to decide whether a text (sense, sentence, etc.) may be associated to *positive* or *negative* connotations. Polarity classification is becoming important for applications such as Opinion Mining and Sentiment Analysis, which facilitates the extraction and analysis of opinions about commercial products, on companies reputation management, brand monitoring, or to track attitudes by mining online forums, blogs, etc. Inspired by work on classification of word senses by polarity (e.g., SentiWordNet), and taking WordNet as a starting point, we build Q-WordNet. Instead of applying external tools such as supervised classifiers to annotated WordNet synsets by polarity, we try to effectively maximize the linguistic information contained in WordNet, thereby taking advantage of the human effort put by lexicographers and annotators. The resulting resource is a subset of WordNet senses classified as positive or negative. In this approach, *neutral polarity* is seen as the absence of positive or negative polarity. The evaluation of Q-WordNet shows an improvement with respect to previous approaches. We believe that Q-WordNet can be used as a starting point for data-driven approaches in sentiment analysis.

1. Introduction

Opinion Mining or Sentiment Analysis are becoming important for determining opinions about commercial products, on companies reputation management, brand monitoring, or to track attitudes by mining online forums, blogs, etc. Given the explosion of information produced and contained via the Internet, it is not possible to keep up with the constant flow of new information by manual methods. A typical application of Sentiment Analysis would be tracking what bloggers are saying about brands like Apple, Microsoft, Ford, etc. (Symposium, 2010)

This paper is focused on polarity classification at word level, namely, it is interested in the connection between lexical semantics and positive and negative connotations associated to a word sense (Strapparava and Mihalcea, 2007). Given the appropriate context, almost every word can potentially convey affective meaning. Every word, even those that are apparently neutral, can be associated with positive or negative experiences by virtue of their semantic relation with emotional concepts or categories. While some acquire polarity connotations in a specific context, there are many others that are just part of a stereotypical common sense knowledge (e.g., 'cancer', 'war', 'mum', 'angel', etc.). The stereotypical meaning associated to certain words if of particular interest for Sentiment Analysis, as it allows to study the use of words in textual productions (Strapparava and Mihalcea, 2007). In this case, the stereotypical or general domain of language use is just seen as another domain, just as the medical domain or the cultural domain, etc.

This paper presents Q-WordNet[1], a freely available lexical resource consisting of WordNet senses automatically annotated by *polarity*. Polarity classification amounts to decide whether a text (sense, sentence, etc.) may be associated to *positive* or *negative* connotations. Inspired by work on classification of word senses by polarity in SentiWordNet (Esuli and Sebastiani, 2006), and taking WordNet as a starting point, we create Q-WordNet. Instead of applying external tools such as supervised classifiers to annotated WordNet synsets by polarity, we try to effectively maximize the linguistic information contained in WordNet, thereby taking advantage of the human effort put by lexicographers and annotators.

The evaluation of Q-WordNet as a binary classification task shows good improvements with respect to SentiWordNet. However, we would like to stress that Q-WordNet is not a finished resource. Although it can be used in its current form for data-driven sentiment analysis (Pang et al., 2002; Pang and Lee, 2004; Kim and Hovy, 2004; Popescu and Etzioni, 2005; Su and Markert, 2008; Su and Markert, 2009; Danescu-Niculescu-Mizil et al., 2009), or for lexical sentiment analysis tasks (Strapparava and Mihalcea, 2007; Su and Markert, 2009), it could also be used as a training set for supervised classifiers that would subsequently be applied for the improvement of Q-WordNet.

Next section reviews previous related work on Sentiment Analysis and, in particular, on polarity classification. Section 3. describes the approach used to automatically annotated WordNet senses by positive and negative polarity. The resulting lexical resource is evaluated in section 4. and we

---

[1]http://nlp.uned.es/semantics/qwn/qwn.html

finish with some concluding remarks and future work in section 5.

2.   Previous Related Work

There is a large amount of work on Opinion Mining at document level focusing on the automatic analysis of commercial and cultural products (Pang and Lee, 2004; Danescu-Niculescu-Mizil et al., 2009). Other previous approaches aim at determining the subjectivity of sentences by means of *terms* that are markers of opinionated content (Kim and Hovy, 2004; Takamura et al., 2005). These approaches also classify subjective words by their polarity. Note that they classify *words* instead of senses, which means that they are not able to capture the fact that a word may contain various senses of which some of them could have different polarities.

There seems to be an assumption in these works that polarity classification (to determine whether the opinion expressed by a word sense, sentence or text is positive or negative) actually depends on subjectivity detection. In other words, that prior to the task of assigning polarity we need to determine whether it is objective (factual) or subjective (opinion). From this point of view, only those expressions deemed to be subjective are classified by polarity (Kim and Hovy, 2004; Takamura et al., 2005). However, it has been argued that this would leave out all those expressions that are suitable to be classified as *objectively* positive or negative (Su and Markert, 2008). Clear cases are those denoting illnesses, such as 'cancer', 'tuberculosis', etc., which stereotypically carry negative associations. Note that, as it was mentioned in the introduction, these senses may not be negative in a medical domain. We therefore take the "general domain" to be a stereotypical domain, which means that our classification may need to be refined for its use in other domains.

There are also other approaches to subjectivity recognition that work at sense level: Su and Markert (2008; 2009) and Wiebe and Milhacea (2006) annotate subjectivity and objectivity of word senses without assuming previous subjectivity classification.

Closer to our work, Esuli and Sebastiani (2006) annotate WordNet 2.0 senses by using a ternary polarity classification (positive, negative and objective) in which each polarity value is assigned a numerical score in such a way that the sum of the three scores is 1.0. The final resource provides the *positive* and *negative* scores for each synset from which the *objective* score is then obtained by calculating *1 - (PosScore + NegScore)*. In order to build Senti-WordNet, they start by selecting the synsets of 14 paradigmatic positive and negative terms used as seed (Turney and Littman, 2003). They are then iteratively extended by means of lexical relations as defined in WordNet, following the construction of WordNet-Affect (Strapparava and Valitutti, 2004). After hand-collecting a number of labelled terms from other resources, they iteratively add the the synsets reachable by navigating the relations of *direct antonymy, similarity, derived-from, pertains-to, attribute, also-see*. In SentiWordNet, the *objective* synsets are those that do to the *positive* or *negative* synsets, and that contain terms which are not marked as either *positive* or *neg-*

ative by Stone et al. (1966). Every synset is then given a VSM (Salton, 1983) representation (cosine-normalized $tf * idf$) to its gloss, which is taken as the textual representation of its meaning. The vectorial representations are fed to a standard supervised learner. Finally, the tokens that are in both *positive* and *negative* categories are classified as *objective*. They trained 7 supervised classifiers using this method which are used to assign *polarity* and *objectivity* scores to WordNet senses.

An evaluation of their classification is provided as an estimation of its Mean Squared Error (Esuli, 2008). They assume that those senses that are not classified as either *positive* or *negative* are in fact *objective*, namely, expressing factual content. In other words, every word that is not associated with a positive or negative connotation is not expressing an opinion. As mentioned earlier, we do not agree with this assumption as there might be word senses objectively associated with positive or negative connotations. For a word to carry polarity does not need to be subjective. Thus, the work presented in this paper focuses on the classification of WordNet senses by their polarity regardless of whether they express subjective opinions or factual information. Unlike SentiWordNet (Esuli and Sebastiani, 2006), which is built using supervised classifiers to annotate Word-Net senses, we exploit the linguistic information (provided by human annotators) contained in WordNet itself and build our resource using an unsupervised method.

3.   Extracting Polarity from WordNet senses

Our approach to classifying WordNet senses by polarity is based on the view of polarity as an *association of a positive or a negative quality to something or to someone*. The idea is to:

1. Link a sense to an attribute of a quality (e.g., positive or negative).

2. Devise, if needed, a procedure to quantify such association, by using either similarity measures (Agirre et al., 2009), or confidence scores by establishing a ranking in our classification (Esuli and Sebastiani, 2007).

The rest of the paper focuses on the first point, namely, on building a lexical resource based on WordNet with its senses classified by polarity.

3.1.   Rationale

Assigning polarity to a word sense can formulated as a binary classification task. This differs from SentiWordNet in the sense that positive or negative associations are assigned to word senses at a certain level by means of a graded classification.

Although it is reasonable to acknowledge the fact that polarity classification is somewhat dependent on context by providing a graded classification, it is also important to bear in mind the fact that polarity classification should be practical, operative and easy to use in a given domain. Our own experience at trying to use SentiWordNet is that the graded classification needs to be collapsed in absolute categories prior to using the resource. In other cases, it is actually difficult to do such simplification. For example, table 1 shows the SentiwordNet classification of the synset *good#a#15*.

Synset	Pos	Neg	Obj
good#a#15	0.25	0.375	0.375

Table 1: good#a#15 SentiWN scores.

This example shows just how difficult using such classification can be. It also shows that, as we argued in the introduction, a synset can be *objectively positive*. This is made even clearer if we look at the lemma names, gloss and examples of synset good#a#15 listed in table 2.

Synset	Lemmas	Gloss
good#a#15	'good' 'well'	'resulting favorably'
Example 1: it's a good thing that I wasn't there		
Example 2: it is good that you stayed		
Example 3: it is well that no one saw you		
Example 4: all's well that ends well		

Table 2: good#a#15 gloss and examples.

Our aim would therefore be to associate such synsets with the attribute of a quality, instead of assigning them a numerical graded score. In our approach those synsets that are not *positive* or *negative* will be considered *neutral* (as opposed to objective or subjective). Furthermore, those synsets that are classified by our method as both positive *and* negative will be discarded, as opposed to SentiWordNet, that consider them *objectives*.

The discarded synsets represent a shortcoming of our approach. In other words, our approach will need to be improved in order to better classify those synsets into positives or negatives (or to ignore them if that is not possible).

In addition to the qualitative aspect of our approach, our objective is also to maximize the human effort employed in building WordNet, and see how far we can go by *walking* WordNet collecting positive and negative senses as we pass by. Of course, the issue here is where to start *walking*. Instead of starting with a list of manually collected seed terms, we will just rely on the linguistic information contained in WordNet. Given that polarity is seen as the attribute of a quality associated to synsets, we will simply start from the *quality synset* contained in WordNet.

There are five noun quality senses in WordNet, two of which contain attribute relations (to adjectives). From the synset *quality#n#01* the attribute relation takes us to *positive#a#01*, *negative#a#01*, *good#a#01* and *bad#a#01*. *quality#n#02* leads to the attributes *superior#a#01* and *inferior#a#02*. Starting by the attributes of *quality* fits well with our intuition that assigning polarity to a linguistic expression can be seen as associating it with an attribute of a quality.

We therefore take these six synsets, expressing positive and negative qualities, to be the departure of our algorithm. The following step is to algorithmically walk through every WordNet relation collecting (i.e., annotating) those synsets that are accessible from the seeds. The resulting resource, is Q-WordNet (Q from the quality synset), the set of WordNet synsets classified by positive or negative polarity as they are accessible through WordNet's relations.

Before going into specific details, it is quite clear that if we

want Q-WordNet to be proportionate in terms of part-of-speech (POS), then we may not only need to walk through WordNet senses, but also to jump from one POS to the other. Otherwise, as our seeds are adjectives (attributes of nouns are adjectives), we risk to obtain a classification consisting mainly of adjectives plus few nouns, verbs and adverbs. This the reason why every WordNet relation is walked, not just those that preserve the affective content (Strapparava and Valitutti, 2004).

Our approach has been applied to WordNet versions 1.6, 1.7, 2.0 and 3.0. Henceforth, when mentioning WordNet we will be referring indistinctly to any of the four versions, unless it is specified otherwise.

3.2. Walking WordNet

Walking through every relation will allow us to jump from adjectives to other POS. Table 3 lists those relations that allow us to extract synsets of different POS starting from the initial set of adjectives which are attributes of the quality synset. The directionality of the relation is indicated by $\leftarrow$ and $\rightarrow$.

input POS	relation	output POS
adj	$\leftarrow$ attribute $\rightarrow$	noun
adj	derived-from $\rightarrow$	noun & verb
noun	pertainym $\rightarrow$	adj
noun	derived-from $\rightarrow$	adj & verb & noun
verb	derived-from $\rightarrow$	adj & noun
adv & adj	pertainym $\rightarrow$	adv

Table 3: Jumping POS in WordNet.

Using every available WordNet relation and their glosses in the way described above is bound to cause a lot of noise in the classified data which will translate in large numbers of false positives and false negatives. In order to prevent this, every synset that appear at both positive and negative categories, at every step in the algorithm, are filtered. Even though this method will discard a large number of synsets, we want Q-WordNet to be as *clean* as possible. Filtering at step from relation to relation will allow to minimize the number of false positives and negatives present in the resource.

The algorithm to build Q-WordNet starts at the attributes of *quality#n#01* and *quality#n#02*, which are adjectives. We perform 10 iterations over the *also_see* relation (our experiments showed that more than 10 iterations creates too much noise). We then go to *similar-to*. From all the collected adjectives, we get their attributes, which allows us to move to nouns. We obtain nouns and verbs from adjectives through the lexical relation *derived-from*. We then use *hyperonymy*, *hyponymy*, *pertainyms*, *derived-from*, *verb-group* and *cause* (plus the *antonym* relation at every step to filter false positives and negatives) to nouns and verbs.

At this stage, it is easy to see that no adverbs are extracted using our methodology. This is because the only relation linking adverbs to other POS in WordNet is through the *pertainym* lexical relation via adjectives. However, the *pertainym* relation is directional from adverbs to adjectives so we cannot start from the already extracted adjectives. In-

stead, the adverbs in Q-WordNet are extracted from a reverse application of the *pertainym* relation to adjectives (adverbs are pertainyms of adjectives). We proceed by extracting the lemmas of every adverb that is matched to the intersection of every lemma's adjective already extracted by our algorithm and the *pertainyms* of the adverb's lemmas (which are adjectives' lemmas). Applying this to WordNet 1.6, 1.7, 2.0 and 3.0 we obtain the following figures:

WN version	Positive	Negative
Q-WordNet 1.6	2240	1772
Q-WordNet 1.7	2271	1792
Q-WordNet 2.0	2884	2100
Q-WordNet 3.0	**7402**	**8108**
SentiWordNet 1.0 (WN 2.0)	**35049**	
WordNet-Affect 1.0 (WN 1.6)	2874	

Table 4: Total number of synsets classified by sentiment.

We obtain 7402 positive and 8108 negative synsets from the application of our method to WordNet 3.0, whereas for WordNet 2.0 we obtain 2884 as positive and 2100 as negative. As a comparison, WordNet-Affect consists of 2874 synsets (admittedly, fine-grained affective annotation is much harder than polarity), and SentiWordNet contains 35409 synsets labelled as either *positive* or *negative* at any *level* (0.125-1.0), even though a synset might be labelled as objective, positive and negative at the same time. This means that we cannot straightforwardly compare Q-WordNet to SentiWordNet (built from WordNet 2.0) because their classification is graded at a specific level. In the following section we will see that as the polarity confidence scores get higher in SentiWordNet, the number of polarity-classified synsets shrinks significantly. In other words, as the polarity graded scores are lower, the number of polarity-classified synsets grows quite large (35049 synsets) but at the cost of a huge increase in false positives and negatives. We discuss this and other issues in the next section, which describes the results of the evaluation of Q-WordNet and how it compares to SentiWordNet 1.0.

4. Evaluation

In order to evaluate Q-WordNet and compare it with SentiWordNet, we use MicroWnOp as testset (Esuli and Sebastiani, 2006). The corpus has originally been annotated by the providers (Esuli and Sebastiani, 2007) with scores for posi- tive, negative and objective/no polarity, thus a mixture of sub- jectivity and polarity annotation. As MicroWnOp is annotated using SentiWordNet's graded three score method, some modifications are required in order to be used for Q-WordNet's evaluation.

4.1. Preparing Testset

First, all objective synsets are removed, as we want to evaluate positive and negative cases in a binary classification task. This results in a final testset of 737 synsets at the lowest polarity score (0.125). As it is the case with SentiWordNet 1.0, the higher the confidence score, the lower the number of synsets on that confidence level. For example, at the 0.375 level, the number of synsets classified by polarity

is 527. Second, some parts of MicroWnOp were annotated by more than one annotator. In order to deal with these cases, the polarity scores were averaged. As SentiWordNet 1.0 is uses WordNet 2.0, we will compare it with the version of Q-WordNet version built on WordNet 2.0. Finally, as both the testset and SentiWordNet 1.0 are graded, we will be performing a comparison between Q-WordNet and SentiWordNet for every score of the scale. This means that to evaluate the systems using the testset containing all synsets annotated by polarity (727), we will be using SentiWordNet 1.0 with polarity synset at 0.125 level, etc. As Q-WordNet uses a qualitative binary classification, it remains unchanged. After several experiments this method of preparing the evaluation is the one for which SentiWordNet obtains the best results. It should also be clear that we do not penalize (as false positive or negative) the fact that sometimes SentiWordNet gives both positive and negative scores to the same synset.

The systems are evaluated in a binary classification task in terms of precision (P), recall (R) and F1 measure. We also measure accuracy (A). This evaluation method means that systems will get lower results as the number of false positives and negatives increases. Results are compared using the McNemar significance test at the 0.05 significance level.

4.2. Results

Table 5 shows the results for each resource. Furthermore, the right-hand size columns show the polarity graded scores at which both SentiWordNet and the testset are considered and the number total synsets annotated as either positive or negative having each graded score as threshold. For example, SentiWordNet 1.0 consists of 4923 synsets with the minimum polarity score of 0.625. As Q-WordNet does not provide graded scores it always consists of 4984 synsets.

Positives							
Q-WordNet			SentiWN 1.0				
P	R	F1	P	R	F1	p@	syn
.84	.95	.89	.66	.68	.67	.125	35049
.89	.95	.92	.77	.76	.76	.25	22855
.94	.97	.95	.84	.87	.85	.375	14611
.94	.97	.96	.88	.93	.90	.50	9445
.96	.96	.96	.88	.90	.89	.625	4923
.96	.97	.97	.96	1	.96	.75	**2027**
.96	1	.98	1	1	1	.875	466
.98	1	.99	1	1	1	1	14
Negatives							
Q-WordNet			SentiWN 1.0				
P	R	F1	P	R	F1	p@	syn
.89	.67	.76	.63	.60	.62	.125	35049
.88	.77	.83	.70	.71	.70	.25	22855
.93	.86	.89	.83	.78	.80	.375	14611
.92	.88	.90	.88	.80	.84	.50	9445
.91	.91	.91	.84	.80	.82	.625	4923
.94	.91	.92	1	.84	.91	.75	**2027**
1	.93	.97	1	1	1	.875	466
1	.96	.98	1	1	1	1	14

Table 5: Results for Positive and Negative Classes.

The results of table 5 show that Q-WordNet clearly outperforms SentiWordNet 1.0 up to the 0.75 level (the differences are statistically significant) where there are not clear differences between the approaches. It should be noted however that, at polarity scores 0.75, SentiWordNet (2027) is less than half the size with respect to Q-WordNet (4984); at 0.875 level SentiWordNet is ten times smaller. When SentiWordNet is of similar size to Q-WordNet (at 0.625 polarity score) or larger, the results are significantly better for Q-WordNet. Moreover, the high scores of both resources at the highest polarity confidence levels is probably due to the reduced size of the testset, just 238 and 195 synsets respectively. As Q-WordNet does not at this state quantify or rank the polarity annotation, it gets evaluated on all 4984 synsets at every level, which in principle should have made it more vulnerable to false positive and negatives for the higher confidence scores. However, the differences between SentiWordNet and Q-WordNet at 0.875 and 1.0 levels are not statistically significant. The differences (or lack thereof) are clearer in figure 1. Another issue worth mentioning is the fact that results are lower for the negative class, especially in terms of recall. These results are confirmed by the Accuracy scores shown in table 6.

Accuracy Pos/Neg			
Q-WN	SentiWN 1.0	SentiWN Synsets	polarity@
.85	.65	35049	.125
.89	.74	22855	.25
.93	.83	14611	.375
.94	.88	9445	.5
.95	.86	4923	.625
.95	.94	**2027**	.75
.98	1	466	.875
.99	1	14	1

Table 6: Measuring Accuracy.

Figure 1 represents the evolution in accuracy of both Q-WordNet and SentiWordNet 1.0. The square pointed line represents SentiWordNet whereas the diamond one refers to Q-WordNet. The results clearly show that as the level of confidence decreases, Q-WordNet's results do not degrade as much as those of SentiWordNet, which means that we have a lower number of false positives and negatives. Even for a evaluation on a small corpus such as MicroWnOp, we believe that the results are encouragin, and show the efectiveness of the strict filtering of synsets during the annotation process.

Given that Q-WordNet is by no means a finished resource, we believe that these results show excellent potential to carry on enriching it with a linguistic processing of glosses (perhaps using disambiguated glosses). It can also be used as training data for data-driven approaches to Sentiment Analysis or to build classifiers which could be later be deployed on WordNet for a larger and richer polarity annotated resource. Our procedure is also suitable to be combined with similarity and ranking algorithms to offer graded polarity as required by any particular applications. Finally, note that as it entirely depends on WordNet structure, the algorithm is directly applicable to any other WordNets avail-
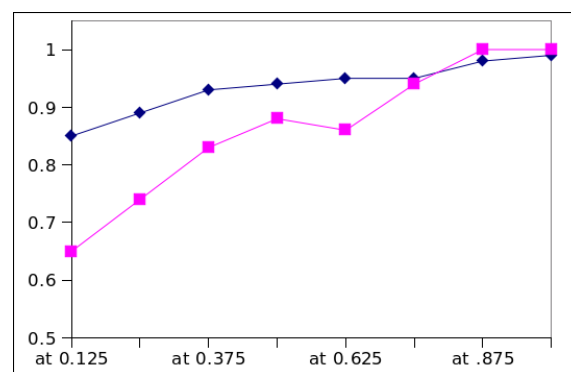

Figure 1: Accuracy Trends on MicroWnOp Corpus.

able in other languages.

5. Concluding Remarks

This paper presents Q-WordNet as a resource consisting of a subset of WordNet synsets annotated by positive and negative polarity. The algorithm used to create Q-WordNet has been applied to WordNet versions 1.6, 1.7, 2.0 and 3.0. We have also offered an evaluation of the version based on WordNet 2.0 for comparison purposes with SentiWordNet 1.0. Results have shown that Q-WordNet in general obtains better results. Further work will evaluate WordNet 3.0 version (consisting of three times more synsets than the WordNet 2.0 version).

Ongoing work towards version 1.0 of Q-WordNet will be focused on the use of WordNet glosses in order to add any synset whose gloss contains a synset already classified by polarity in Q-WordNet. This work can be done using third-party approaches for disambiguation (Agirre and Soroa, 2009) or by exploiting the presence of synset's lemmas in the glosses. For example, we can extract those synsets in WordNet's glosses for which at least one of the lemmas are part of a synset in Q-WordNet. This is done by parsing WordNet's glosses looking for those synsets' lemmas we have already collected. The parsing is performed by an assembled pipeline consisting of the following free available tools: C&C tokenizer (Clark and Curran, v10) and the Stanford POS tagger and Dependency parser (Toutanova et al., 2003; de Marneffe et al., 2006). The general idea is that if a positive synset is matched in a gloss, then the synset whose gloss we have analyzed is also annotated as positive. The parsing is particularly important to lemmatize the glosses but also to make sure that any matched lemma is not under the scope of a negation. If that is the case, the synset is classified as the opposite of the matched lemma.

6. References

Eneko Agirre and Aitor Soroa. 2009. Personalizing pagerank for word sense disambiguation. In *Proceedings of*

the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL-2009), Athens, Greece.

Eneko Agirre, Enrique Alfonseca, Keith Hall, Jana Kravalova, Marius Pasca, and Aitor Soroa. 2009. A study on similarity and relatedness using distributional and wordnet-based approaches. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 19–27, Boulder, Colorado, June. Association for Computational Linguistics.

Stephen Clark and James Curran. v1.0. C&C tools, http://svn.ask.it.usyd.edu.au/trac/candc.

Cristian Danescu-Niculescu-Mizil, Gueorgi Kossinets, Jon Kleinberg, and Lillian Lee. 2009. How opinions are received by online communities: A case study on Amazon.com helpfulness votes. In *Proceedings of the WWW*, pages 141–150.

Marie-Catherine de Marneffe, Bill MacCartney, and Christopher Manning. 2006. Generating typed dependency parses from phrase structure parses. In *Proceedings of Language Resources and Evaluation Conference (LREC)*.

A. Esuli and F. Sebastiani. 2006. Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of the 5th Conference on Language Resources and Evaluation (LREC 2006)*, pages 417–422, Genoa, Italy, May.

Andrea Esuli and Fabrizio Sebastiani. 2007. Pageranking wordnet synsets: An application to opinion mining. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 424–431, Prague, Czech Republic, June. Association for Computational Linguistics.

Andrea Esuli. 2008. *Automatic Generation of Lexical Resources for Opinion Mining: Models, Algorithms and Applications*. Ph.D. thesis, Università di Pisa.

Soo-Min Kim and Eduard Hovy. 2004. Determining the sentiment of opinions. In *Proceedings of Coling 2004*, pages 1367–1373, Geneva, Switzerland, Aug 23–Aug 27. COLING.

Bo Pang and Lillian Lee. 2004. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL'04), Main Volume*, pages 271–278, Barcelona, Spain, July.

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing*, pages 79–86. Association for Computational Linguistics, July.

Ana-Maria Popescu and Oren Etzioni. 2005. Extracting product features and opinions from reviews. In *Proceedings of the International Conference on Empirical Methods in Natural Language Processing (EMNLP'05)*.

G. Salton. 1983. *Introduction to Modern Information Retrieval*. McGraw-Hill.

P. Stone, D. Dunphy, M. Smith, and D. Ogilvie. 1966. *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge (MA): MIT Press.

Carlo Strapparava and Rada Mihalcea. 2007. Semeval-2007 task 14: Affective text. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 70–74, Prague, Czech Republic, June. Association for Computational Linguistics.

Carlo Strapparava and Alessandro Valitutti. 2004. Wordnet-affect: an affective extension of wordnet. In *Proceedings of the 4th International Conference on Languages Resources and Evaluation (LREC 2004)*, pages 1083–1086, Lisbon, May.

Fangzhong Su and Katja Markert. 2008. From words to senses: A case study of subjectivity recognition. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, pages 825–832, Manchester, UK, August.

Fangzhong Su and Katja Markert. 2009. Subjectivity recognition on word senses via semi-supervised mincuts. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 1–9, Boulder, Colorado, June. Association for Computational Linguistics.

Sentiment Analysis Symposium. 2010. http://www.sentimentsymposium.com.

Hiroya Takamura, Takashi Inui, and Manabu Okumura. 2005. Extracting semantic orientations of words using spin model. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*, pages 133–140, Ann Arbor, Michigan, June. Association for Computational Linguistics.

Kristina Toutanova, Dan Klein, Christopher Manning, and Yoram Singer. 2003. Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. In *Proceedings of HLT-NAACL*, pages 252–259.

P. Turney and M. Littman. 2003. Measuring praise and criticism: Inference of semantic oreintation from association. *ACM Transaction on Information Systems*, 21(4):315–346.

Janyce Wiebe and Rada Mihalcea. 2006. Word sense and subjectivity. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, pages 1065–1072, Sydney, Australia, July. Association for Computational Linguistics.