# Old Wine or Warm Beer:
# Target-Specific Sentiment Analysis of Adjectives

**Angela Fahrni & Manfred Klenner**[1]

**Abstract.** In this paper, we focus on the target-specific polarity determination of adjectives. A domain-specific noun, the target noun, is modified by a qualifying adjective. Rather than having a prior polarity, adjectives are often bearing a *target-specific* polarity. In some cases, a single adjective even switches polarity depending on the accompanying noun. In order to realise such a 'sentiment disambiguation', a two stage model is proposed: Identification of domain-specific targets and the construction of a target-specific polarity adjective lexicon. We use Wikipedia for automatic target detection, and a bootstrapping approach to determine the target-specific adjective polarity. It can be shown that our approach outperforms a baseline system that is based on a prior adjective lexicon derived from SentiWordNet.

## 1 INTRODUCTION

Approaches to sentiment analysis range from counting the (prior) polarity of words [9] to systems that do a full compositional semantics analysis of sentence affect [5]. Specific resources have been developed, e.g. adjective lists [4], SentiWordNet [2] or WordNet-Affect [8], that compile the prior polarity of words. It has been noted, however, that the polarity of words is not in any case domain-independent [9]. An 'unpredictable plot' in the movie domain might be a good thing, but an 'unpredictable boss' surely is not. Moreover, as we would argue, even within a domain, the polarity of adjectives can vary. Take the adjective 'cold'. While a 'cold coke' is positive, a 'cold pizza' is not. 'Coke' and 'pizza' are domain-specific targets. Note that the adjective 'cold' has the same WordNet sense in both contexts (i.e. temperature reading), but the polarities are inverse. A kind of target-specific sentiment disambiguation seems to be necessary.

We propose a two stage model. First, the targets of a domain are identified. We use Wikipedia's and Wikionary's category system to get a comprehensive and moreover dynamic (since both resources are growing and growing) target list. In a second step, the target-specific polarity of adjectives is determined in a corpus-driven manner by searching for combinations of a target-specific adjective with adjectives that have a known prior polarity (e.g. good, excellent etc.). In order to evaluate our approach, we have derived an adjective lexicon with prior polarities from SentiWordNet. It serves as a baseline in our experiments carried out with 3891 automatically extracted and – by two independently working annotators[2] – manually classified (positive, negative, neutral) noun phrases.

The domain of fast food restaurants was chosen as a test bed for our approach. We have downloaded about 1600 (manually classified) texts from epinions.com, a website with a huge amount of customer opinions concerning a broad spectrum of topics (holiday resorts, cars, credit cards, lawyers ..). Although these texts are manually classified along five categories: from very bad (one star) to really good (five stars), they do not establish a gold standard for our task, which is NP polarity detection. However, the ultimative goal of our work is to do sentiment detection on the sentence and eventually on the text level.

## 2 WIKIPEDIA-BASED TARGET DETECTION

Wikipedia's category system[3] is used to organise the stock of Wikipedia articles. It is hierarchical, but it does not constitute a genuine taxonomy, since it is based on pragmatic rather than ontological considerations. Although Wikipedia's hierarchy might be questionable, it actually does identify crucial domain-specific concepts. Moreover, the category tree also specifies named entities such as product names, proper names and brand names. This is a big advantage, since these items most often are the targets we are interested in. Adapting to a new domain boils down to identify the crucial Wikipedia categories (on an appropriate hierarchical level).

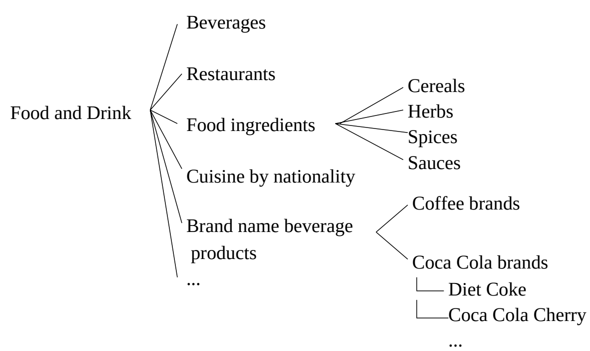


**Figure 1.** Wikipedia category 'Food and Drink'

In the fast food domain, /Food and drink/ is the most interesting category[4], it identifies 46807 targets. See Fig. 1 for a fragment of the category tree. Rather than using a flattened list of these targets, we keep the hierarchy in order to propagate polarities. For example, if it is known that 'cold coca cola' is positive then 'cold coca cola cherry'

[1] Institute of Computational Linguistics, University of Zurich, Switzerland, email: angela.fahrni@swissonline.ch, klenner@cl.uzh.ch

[2] Annotation mismatches have been resolved afterwards.

[3] http://en.wikipedia.org/wiki/Portal:Contents/Categorical_index

[4] /Furniture/ and /Service/ might as well provide additional targets.

also is. Note that in the literature, e.g. [7], what we call targets, is sometimes called *features* and *attributes*.

## 3 TARGET-SPECIFIC OR PRIOR POLARITY

We argue that only a few adjectives do have a prior positive or negative polarity. Especially vague adjectives such as 'big', 'young', 'large', 'deep' are best understood as bearing neutral prior polarity. They adapt, however, to the context by either acting as intensifiers of the intrinsic positive or negative polarity of a target noun (e.g. 'deep insight', 'deep disappointment') or they combine with a neutral noun to form a non-neutral polarity (e.g. 'old bread'). Extreme examples such as 'cold pizza' and 'cold coke' where a single neutral adjective yields positive or negative polarity depending on the (neutral) noun are best explained as the violation ('cold pizza') and affirmation ('cold coke') of intrinsic or common sense properties of target objects (pizza, coke). In the absence of common sense reasoning, we propose a corpus-driven approach to determine such target-specific polarities.

Nevertheless, we sometimes need prior polarities, since they help us to explain certain compositional effects. Take the adjective 'lost' which has a (prior) negative polarity. A 'lost virtue' (virtue=positive) is negative, 'lost glasses' (glasses=neutral) is negative, but 'lost anger' (anger=negative) is positive. If 'lost' had no prior (negative) polarity, it could not combine with a negative word to form a positive noun phrase ('lost anger'). It is also not a simple valency shifter, otherwise we could not explain how the combination with a neutral noun forms a negatively qualified noun phrase ('lost glasses').

In the literature, adjectives with a clear prior polarity have been used as a seed list in order to identify the polarity of additional adjectives, e.g. [9]. The assumption of these approaches was that the augmented list again establishes a set of adjectives having a prior polarity. Contradicting polarities of an adjective encountered in a corpus were interpreted as a kind of noise and are resolved to one (predominant) polarity using statistical measures. We argue that often it is not noise what is encountered but target-specific sentiment ambiguity.

While we are relying on the same methods to identify the polarity of non-seed adjectives discussed in the literature, namely contextual pattern such as coordination, we aim at building a target-specific adjective lexicon instead of a domain-independent lexicon. Our seed adjective lexicon consists of 120 negative and 80 positive adjectives, where the polarity is supposed to be domain- and target-independent[5]. A few examples of positive polarity adjectives are: wonderful, tasteful, superb, positive, perfect, nice, ideal, great, excellent, delightful, delicious, good, beautiful.

## 4 TARGET-SPECIFIC POLARITY LEXICON

To get a target-specific polarity lexicon, two different corpora are being used. The one previously described (1600 texts from epinions.com, henceforth corpus I) and the world wide web (corpus II). Corpus I is tagged and all targets are identified. The most frequent targets from corpus I are used to find new texts in corpus II. Corpus I acts as a kind of reference corpus: we know that these texts are fast food ratings and thus we know that the adjectives used there and the targets from our (Wikipedia derived) target list (which might be noisy) that occur in these texts actually are relevant for the task at hand: the construction of a target-specific adjective lexicon. Corpus II is the pool used to identify the polarity of the non-seed adjectives from corpus I with respect to specific targets.

After we have identified the adjectives and targets we are interested in, we proceed as follows: We search both corpora for tag sequences that relate a target and at least two adjectives. It must hold that:

- the noun or noun sequence is a target
- at least one of the adjectives is from the seed list
- at least one of the adjectives comes from the stock of target-relevant adjectives

Currently, two sequence patterns are considered:

- adjective coordination (incl. modifying adverbs) e.g. 'good and tasteful burger'
- copula constructions, e.g. NP BE Adj Adj+ e.g. 'the french fries are soggy and rather tasteless'

It is assumed that adjectives in such constructions share the polarity[6]. As already mentioned, this kind of pattern directed polarity determination of adjectives is not new. However, in contrast to previous approaches, we require a *target* to be present relative to which the sentiment disambiguation is done. Moreover, the adjective and target must be of *interest* according to a reference corpus (corpus I).

**Table 1.** Examples of polarity tagged noun phrases

| ADJ | Target | OWN-Pol. | SWN-Pol. |
|---|---|---|---|
| hot | burger | 1 | 0 |
| cheap | burger | 0.949 | -1 |
| fresh | fruit | 0.75 | 0 |
| mouth-watering | burger | 0.975 | 1 |
| sized | sandwich | 1 | 0 |
| supersonic | burger | 0.95 | 0 |

OWN System SWN SentiWordNet

Table 1 gives some examples of polarity tagged pairs generated by our systems. All noun phrases receive positive polarity from our system (OWN-Pol.) but quite different polarities from our baseline system that relies on an adjective list derived from SentiWordNet (last column, SWN-Pol., see section 5). The *polarity values* are gradual, ranging from -1 (very bad) to 1 (very good); 0 means neutral.

The polarity values of the seed adjectives are manually set, in some cases we took information from SentiWordNet into account. Polarity values of non-seed adjectives are given as the mean of the polarity values of their peers (where a peer is e.g. a seed adjective that occurs together with it in a coordination)[7].

All those adjectives that have a single polarity with all of its targets receive a *domain-specific* polarity. These and only these adjectives are combined with the seed list to form an augmented seed list. If the polarity of an adjective depends on the target, an adjective-target pair is added to the polarity-specific lexicon. Those adjectives from corpus I that have not received a polarity in the first cycle (since they never occurred e.g. in a coordination with a seed adjective) get a second (third and fourth) change. They might get a polarity in another round, on the basis of the incrementally augmented seed list. Currently, we run four such incremental cycles.

[5] Of course, figurative language readily produces counterexamples.

[6] There are, of course, exceptions, e.g. 'rich and poor people'.

[7] We have also experimented with a *confidence value* of a polarity classification, which is meant to tell us how strong a decision was.

## 5 A PRIOR-POLARITY LEXICON DERIVED FROM SentiWordNet

[2] introduce a semi-automatic approach to derive a version of WordNet where word senses are bearing polarities. The resource is called SentiWordNet and is freely available for research purposes. The developers rely on the same idea as described above, namely a seed of paradigm words with a clear polarity.

**Table 2.** SentiWordNet: ’unpredictable’

| POS | synset | pos. | neg. | word | sense |
|---|---|---|---|---|---|
| a | 1781371 | 0.0 | 0.625 | unpredictable#a | #1 |
| a | 708935 | 0.0 | 0.0 | unpredictable#a | #2 |
| a | 566807 | 0.0 | 0.25 | unpredictable#a | #3 |

Table 2 shows the entry of ’unpredictable’. The numbers below the polarity tags (pos., neg.) indicate the polarity strength (1 indicates maximal strength). Word sense 1 and 3 of ’unpredictable’ have negative polarity, while word sense 2 is neutral.

In SentiWordNet, the adjective ’hot’ has 22 senses, 7 of them have neutral, 5 have negative and 10 have positive polarity. Since we are only interested in the polarities, we merge the positive, negative and neutral senses into one polarity entry, respectively. Each entry receives as its polarity weight the weighted sum of its SentiWordNet scores, e.g.

$$weight('hot' = pos) = \frac{\sum_{i \in swn\_pol('hot'=pos)} swn\_score(i)}{\mid synsets('hot') \mid}$$

where $swn\_pol('hot' = pos)$ denotes the set of synsets of ’hot’ bearing positive polarity and $swn\_score(i)$ is the value of the SentiWordNet entry of word sense $i$ of ’hot’.

This way, the adjective ’hot’ in its neutral reading gets a weight of 0.6, while positively interpreted it receives 0.28, leaving a 0.12 score to the remaining negative case. Applying this strategy to SentiWordNet, we have generated an adjective lexicon with prior polarities that blends the numerical weights of an adjective’s SentiWordNet entry into three discrete polarity classes. Altogether, 21194 adjective entries has been derived. Note that some of them has received three (’e.g. ’hot’), some two (e.g. ’unpredictable’) and other only one polarity entry (e.g. ’good’).

## 6 EVALUATION

We have carried out an evaluation of our system on the basis of 3891 manually classified noun phrases[8]. The resulting gold standard comprises 1832 positive, 415 negative and 1644 neutral instances.

Three different experimental settings are distinguished. First, we compared the polarity decisions of SentiWordNet (our baseline system) and our system for the whole data set ($all$). Second, we took only those classifications that received different polarities from the two systems ($conflict$). Third, only the instances where both systems agreed in their polarity assignment are taken ($agree$).

Table 3 shows the accuracy under these conditions. Given the whole data set (3891 NPs), our system outperforms SentiWordNet by 6.8%. This setting is the ’realistic’ one, so, given domain-specific texts, a substantial improvement can be achieved with the methodology we propose. If we (only) evaluate the conflicting classifications (1937 NPs), our system shows its strength. Here an improvement of 13.7% was achieved. The evaluation of those cases where both systems assign the same polarity (1954 NPs shows that we can design a high-accuracy system by combining both resources, SentiWordNet and our system.

[8] which corresponds to 2426 NP types

**Table 3.** Accuracy under 3 experimental settings

| | SWN | OWN |
|---|---|---|
| all | 63.4 % | 70.2% |
| conflict | 39.2 % | 52.9% |
| agree | 87.4 % | 87.4% |

SWN SentiWordNet OWN System

**Table 4.** Evaluation of (all) 3891 noun phrases

| | SWN | | | OWN | | |
|---|---|---|---|---|---|---|
| | prec | rec | f-meas | prec | rec | f-meas |
| pos | 97.5 % | 39.5% | 55.9% | 66.0% | 91.2 % | 76.5% |
| neg | 89.8 % | 34.2% | 49.5% | 82.3% | 37.2 % | 51.1% |
| neut | 53.7 % | 97.4% | 69.3% | 77.3% | 55.2 % | 64.4% |
| ⊘ | | | 58% | | | 64% |

SWN SentiWordNet OWN System ⊘ arithmetic mean

Table 4 shows the results (whole data set) for each single class. We can see that our approach clearly outperforms SentiWordNet with respect to the positive NPs ( 76.5% F-measure compared to 55.9%), but only slightly given the negative NPs (1.6%). Given neutral NPs, SentiWordNet wins (4.9%). A closer look at the data shows that SentiWordNet has a strong bias towards neutral classifications.

**Table 5.** Evaluation of (conflict) 1937 classification conflicts

| | SWN | | | OWN | | |
|---|---|---|---|---|---|---|
| | prec | rec | f-meas | prec | rec | f-meas |
| pos | 37.5 % | 1.2% | 2.4% | 53.1% | 98.4 % | 68.9% |
| neg | 58.9 % | 11.3% | 19.3% | 51.5% | 17.2 % | 25.7% |
| neut | 38.9 % | 95.7% | 55.3% | 50.8% | 4.2 % | 7.8% |
| ⊘ | | | 25.5% | | | 34.2% |

SWN SentiWordNet OWN System ⊘ arithmetic mean

From Table 5 we can see that our system has a bias towards positive classifications, but precision is still reasonable, so a F-measure of 68.9% was achieved. Note that it is the class of conflicting classifications where our system (a target-specific approach) has to prove is advantages over a system with prior polarities. The overall difference in performance is 8.7%. But since normally one is interested in positive and negative polarities rather than neutral, our system improves at the right place. If we look at these two classes, our approach is 70.9% superior to SentiWordNet (to be more precise: our adjective list derived from SentiWordNet). However, we clearly have to come to a more balanced performance.

Finally, from Table 6 ($agree$) we can see that a combination of the two approaches can act as a high-precision system correctly identifying 98.3% of the positive and 99.9% of the negative NPs. Note however that those NPs receiving the same vote from both systems most often include a seed adjective. So they won’t have a target-specific polarity.

**Table 6.** Evaluation of 1953 (agree) decisions.

| | SWN + OWN | | |
|---|---|---|---|
| | prec | rec | f-meas |
| pos | 98.3 % | 83.2% | 90.1% |
| neg | 99.9 % | 56.4% | 72.1% |
| neut | 78.9 % | 98.9% | 87.8% |
| ⊘ | | | 83.3% |

⊘ arithmetic mean

## 7 RELATED WORK

Our approach to the identification of polarity of adjectives is based on the ideas of [3] (among others). However, [3] only identify prior polarity, not contextual.

Work on contextual polarity detection is described in [11]. Here, a (supervised) machine learning approach is used to find the contextual polarity of words. In our (semi-supervised) approach, the notion of a domain-specific target is stressed, while in their approach this is left implicit as a problem to be solved by the machine learning component. Note that [11] are striving to cope with a more challenging domain, namely news texts. Accordingly, the empirical performance reported there is worse than the one reported here. But we can not seriously compare both.

Our approach to target detection is based on Wikipedia's category systems. Others, e.g. [7] have used contextual, e.g. meronymy discriminators such as 'the X of Y' where X is identified as an *attribute* of the *feature* Y. We plan to improve our Wikipedia based approach by also taking Wikipedia articles into account. Then, contextual discriminators but also available tools such as those described in [6] might prove helpful.

There are several approaches to derive polarity tagged adjective lists form WordNet, e.g. [1], [4]. Since we plan to use SentiWordNet [2] also as a source for noun and verb polarity, we have already worked with it to derive a baseline system for adjective polarity detection.

Finally, the interaction between word sense disambiguation and subjectivity has been discussed by [10]. However, in their system sentiment detection helps word sense disambiguation while in our approach a single word sense might even give rise to two inverse target-specific polarities. Our solution to that problem has the side effect that word sense disambiguation becomes superfluous. If an adjective changes polarities depending on the target, both, adjective and target are added to the target-specific lexicon[9].

## 8 CONCLUSION AND FUTURE WORK

We have introduced a semi-supervised approach to NP polarity detection that is based on a target-specific polarity lexicon induced from a seed lexicon and two corpora. This enables our system to assign different polarities to NPs with the same adjective but a different target noun (e.g. 'cold french fries' and 'cold buttermilk'). Our system outperforms a baseline system derived from SentiWordNet. The SentiWordNet baseline establishes a kind of upper bound for approaches that rely on *prior* polarity information only.

Although we have started to experiment with both a measure of polarity strength and a confidence metric, the results have not been sufficiently evaluated and thus are not presented here. Polarity strength tells us how strong a positive or negative evaluation (here NP polarity) is, confidence indicates how reliable a polarity decision is.

Currently, NP polarity depends exclusively on adjective polarity. This is an artefact of the chosen domain where nouns mostly are neutral (food, furniture, employees etc.). But NP polarity often is compositional (as is sentence polarity). For example, a positive adjective and a negative noun ('excellent forgery', 'perfect spy') combine to a negative polarity. Therefore, but also to prove the domain independence of our model, we plan to switch to another domain.

We are also working on a model of sentence-level sentiment analysis. Currently, our main focus lies on the identification of basic dependency structure that reliably indicate 'subject verb object' constellations ('I love this little book'). We then will focus on negation ('never'), intra-sentential valency shifters ('but') and complex compositional phenomena ('this could not fail to get nasty').

[9] However, neither approach is fully satisfying, since there are cases where different senses of an adjective do have inverse polarities when combined with a single noun, e.g. the 'inexpensive' versus 'poor quality' reading of 'cheap food'.

## REFERENCES


[1] Alina Andreevskaia and Sabine Bergler, 'Mining wordnet for a fuzzy sentiment: Sentiment tag extraction from wordnet glosses', in *Proceedings EACL-06, the 11rd Conference of the European Chapter of the Association for Computational Linguistics*, pp. 209–216, Trento, IT, (2006).

[2] Andrea Esuli and Fabrizio Sebastiani, 'SentiWordNet: A publicly available lexical resource for opinion mining', in *Proceedings of LREC-06, the 5th Conference on Language Resources and Evaluation*, Genova, IT, (2006).

[3] Vasileios Hatzivassiloglou and Kathleen R. McKeown, 'Predicting the semantic orientation of adjectives', in *Proceedings of ACL-97, 35th Annual Meeting of the Association for Computational Linguistics*, pp. 174–181, Madrid, ES, (1997). Association for Computational Linguistics.

[4] Jaap Kamps, Robert J. Mokken, Maarten Marx, and Maarten de Rijke, 'Using WordNet to measure semantic orientation of adjectives', in *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*, pp. 1115–1118. European Language Resources Association, Paris, (2004).

[5] Karo Moilanen and Stephen Pulman, 'Sentiment composition', in *Proceedings of the Recent Advances in Natural Language Processing International Conference (RANLP-2007)*, pp. 378–382, Borovets, Bulgaria, (September 27-29 2007).

[6] Simone Paolo Ponzetto and Michael Strube, 'An API for measuring the relatedness of words in Wikipedia', in *Companion Volume to the Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, pp. 49–52, Prague, Czech Republic, (June 2007).

[7] Ana-Maria Popescu and Oren Etzioni, 'Extracting product features and opinions from reviews', in *Proceedings of HLT-EMNLP-05, the Human Language Technology Conference/Conference on Empirical Methods in Natural Language Processing*, pp. 339–346, Vancouver, CA, (2005).

[8] C. Strapparava and A. Valitutti, 'Wordnet-affect: an affective extension of wordnet', in *Proceedings of LREC-04, the 5th Conference on Language Resources*, (2004).

[9] Peter D. Turney and Michael L. Littman, 'Measuring praise and criticism: Inference of semantic orientation from association', *ACM Transactions on Information Systems*, **21**(4), 315–346, (2003).

[10] Janyce Wiebe and Rada Mihalcea, 'Word sense and subjectivity', in *Proceedings of COLING/ACL-06, the 21st Conference on Computational Linguistics / Association for Computational Linguistics*, pp. 1065–1072, Sydney, AUS, (2006). Association for Computational Linguistics.

[11] Theresa Wilson, Janyce Wiebe, and Paul Hoffmann, 'Recognizing contextual polarity in phrase-level sentiment analysis', in *Proceedings of Human Language Technologies Conference/Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005)*, Vancouver, CA, (2005).