

OpAL: Applying Opinion Mining Techniques for the Disambiguation of Sentiment Ambiguous Adjectives in SemEval-2 Task 18

Alexandra Balahur
University of Alicante
Department of Software and
Computing Systems
abalahur@dlsi.ua.es

Andrés Montoyo
University of Alicante
Department of Software and
Computing Systems
montoyo@dlsi.ua.es

Abstract

The task of extracting the opinion expressed in text is challenging due to different reasons. One of them is that the same word (in particular, adjectives) can have different polarities depending on the context. This paper presents the experiments carried out by the OpAL team for the participation in the SemEval 2010 Task 18 – *Disambiguation of Sentiment Ambiguous Adjectives*. Our approach is based on three different strategies: a) the evaluation of the polarity of the whole context using an opinion mining system; b) the assessment of the polarity of the local context, given by the combinations between the closest nouns and the adjective to be classified; c) rules aiming at refining the local semantics through the spotting of modifiers. The final decision for classification is taken according to the output of the majority of these three approaches. The method used yielded good results, the OpAL system run ranking fifth among 16 in micro accuracy and sixth in macro accuracy.

1 Credits

This research has been supported by Ministerio de Ciencia e Innovación - Spanish Government (grant no. TIN2009-13391-C04-01), and Conselleria d'Educació-Generalitat Valenciana (grant no. PROMETEO/2009/119 and ACOMP/2010/288).

2 Introduction

Recent years have marked the beginning and expansion of the Social Web, in which people freely express and respond to opinion on a whole

variety of topics. Moreover, at the time of taking a decision, more and more people search for information and opinions expressed on the Web on their matter of interest and base their final decision on the information found (Pang and Lee, 2008). Nevertheless, the high quantity of data that has to be analysed imposed the development of specialized Natural Language Processing (NLP) systems that automatically extract, classify and summarize the opinions available on the web on different topics. Research in this field, of *opinion mining* (sentiment analysis), has addressed the problem of extracting and classifying opinions from different perspectives and at different levels, depending on various factors. While determining the overall opinion on a movie is sufficient for taking the decision to watch it or not, when buying a product, people are interested in the individual opinions on the different product characteristics. Especially in this context, opinion mining systems are confronted with a difficult problem: the fact that the adjectives used to express opinion have different polarities depending on the characteristic they are mentioned with. For example, “high price” is negative, while “high resolution” is positive. Therefore, specialized methods have to be employed to correctly determine the contextual polarity of such words and thus accurately assign polarity to the opinion.

This is the aim of the SemEval 2010 Task 18 – *Disambiguation of Sentiment Ambiguous Adjectives* (Wu and Jin, 2010). In the following sections, we first present state-of-the art approaches towards polarity classification of opinions, subsequently describing our approach in the SemEval task. Finally, we present the results we obtained in the evaluation and our plans for future work.

3 State of the Art

Subjectivity analysis is defined by (Wiebe, 1994) as the “linguistic expression of somebody’s opinions, sentiments, emotions, evaluations, beliefs and speculations”. *Sentiment analysis*, on the other hand, is defined as the task of extracting, from a text, the opinion expressed on an object (product, person, topic etc.) and classifying it as positive, negative or neutral. The task of sentiment analysis, considered a step further to subjectivity analysis, is more complex than the latter, because it involves an extra step: the classification of the retrieved opinion words according to their polarity. There are a series of techniques that were used to obtain lexicons of subjective words – e.g. the Opinion Finder lexicon (Wilson et al., 2005) and opinion words with associated polarity. (Hu and Liu, 2004) start with a set of seed adjectives (“good” and “bad”) and apply synonymy and antonymy relations in WordNet. A similar approach was used in building WordNet Affect (Strapparava and Valitutti, 2004), starting from a larger set of seed affective words, classified according to the six basic categories of emotion (joy, sadness, fear, surprise, anger and disgust) and expanding the lexicon using paths in WordNet. Another related method was used in the creation of SentiWordNet (Esuli and Sebastiani, 2005), using a set of seed words whose polarity was known and expanded using gloss similarity. The collection of appraisal terms in (Whitelaw et al., 2005), the terms also have polarity assigned. MicroWNOp (Cerini et al., 2007), another lexicon containing opinion words with their associated polarity, was built on the basis of a set of terms extracted from the General Inquirer lexicon and subsequently adding all the synsets in WordNet where these words appear. Other methods built sentiment lexicons using the local context of words. (Pang et al., 2002) built a lexicon of sentiment words with associated polarity value, starting with a set of classified seed adjectives and using conjunctions (“and”) disjunctions (“or”, “but”) to deduce orientation of new words in a corpus. (Turney, 2002) classifies words according to their polarity on the basis of the idea that terms with similar orientation tend to co-occur in documents. Thus, the author computes the Pointwise Mutual Information score between seed words and new words on the basis of the number of AltaVista hits returned when querying the seed word and the word to be classified with the “NEAR” operator. In our work in (Balahur and Montoyo, 2008a), we compute the polarity

of new words using “polarity anchors” (words whose polarity is known beforehand) and Normalized Google Distance (Cilibrasi and Vitanyi, 2006) scores. Another approach that uses the polarity of the local context for computing word polarity is (Popescu and Etzioni, 2005), who use a weighting function of the words around the context to be classified.

4 The OpAL system at SemEval 2010 Task 18

In the SemEval 2010 Task 18, the participants were given a set of contexts in Chinese, in which 14 dynamic sentiment ambiguous adjectives are selected. They are: 大|big, 小|small, 多|many, 少|few, 高|high, 低|low, 厚|thick, 薄|thin, 深|deep, 浅|shallow, 重|heavy, 轻|light, 巨大|huge, 重大|grave. The task was to automatically classify the polarity of these adjectives, i.e. to detect whether their sense in the context is positive or negative. The contexts were given in two forms: as plain text, in which the adjective to be classified was marked; in the second for, the text was tokenized and the tokens were tagged with part of speech (POS). There was no training set provided.

Our approach uses a set of opinion mining resources and an opinion mining system that is implemented to work for English. This is why, the first step we took in our approach was to translate the given contexts into English using the Google Translator¹. In order to perform this task, we first split the initial file into 10 smaller files, using a specialized program – GSplit32. The OpAL adjective polarity disambiguation system combines supervised methods with unsupervised ones. In order to judge the polarity of the adjectives, it uses three types of judgments. The first one is the general polarity of the context, determined by our in-house opinion mining system - based on SVM machine learning on the NTCIR data and the EmotiBlog (Boldrini et al., 2009) annotations and different subjectivity, opinion and emotion lexica (Opinion Finder, MicroWordNet Opinion, General Inquirer, WordNet Affect, emotion triggers (Balahur and Montoyo, 2008b)). The second one is the local polarity, given by the highest number of results obtained when issuing queries containing the closest noun with the adjective to be disambiguated followed by the conjunction “AND” and a predefined set of 6 adjectives whose polarity is non-

¹ <http://translate.google.com/>

² www.gdgsoft.com/gsplit/

ambiguous – 3 positive - “positive”, “beautiful”, “good” and 3 negative – “negative”, “ugly”, “bad”. An example of such queries is “price high and good”. The third component is made up of rules, depending on the presence of specific modifiers in a window of 4 words before the adjective. The final verdict is given based on the vote given by the majority of the three components, explained in detail in the next sections:

4.1 The OpAL opinion mining component

First, we process each context using Minipar³. We compute, for each word in a sentence, a series of features, computed from the NTCIR 7 data and the EmotiBlog annotations. These words are used to compute vectors of features for each of the individual contexts:

- the part of speech (POS)
- opinionatedness/intensity - if the word is annotated as opinion word, its polarity, i.e. 1 and -1 if the word is positive or negative, respectively and 0 if it is not an opinion word, its intensity (1, 2 or 3) and 0 if it is not a subjective word
- syntactic relatedness with other opinion word – if it is directly dependent of an opinion word or modifier (0 or 1), plus the polarity/intensity and emotion of this word (0 for all the components otherwise)
- role in 2-word, 3-word, 4-word and sentence annotations: opinionatedness, intensity and emotion of the other words contained in the annotation, direct dependency relations with them if they exist and 0 otherwise.

We add to the opinion words annotated in EmotiBlog the list of opinion words found in the Opinion Finder, Opinion Finder, MicroWordNet Opinion, General Inquirer, WordNet Affect, emotion triggers lexical resources. We train the model using the SVM SMO implementation in Weka⁴.

4.2 Assessing local polarity using Google queries

This approach aimed at determining the polarity of the context immediately surrounding the adjective to be classified. To that aim, we constructed queries using the noun found before the adjective in the context given, and issued six different queries on Google, together with six pre-defined adjectives whose polarity is known (3

positive - “positive”, “beautiful”, “good” and 3 negative – “negative”, “ugly”, “bad”). The form of the queries was “noun+adjective+AND+pre-defined adjective”. The local polarity was considered as the one for which the query issued the highest number of total results (total number of results for the 3 queries corresponding to the positive adjectives or to the negative adjectives, respectively).

4.3 Modifier rules for contextual polarity

This rule accounts for the original, most frequently used polarity of the given adjectives (e.g. *high* is *positive*, *low* is *negative*). For each of them, we define its default polarity. Subsequently, we determine whether in the window of 4 words around the adjective there are any modifiers (valence shifters). If this is the case, and they have an opposite value of polarity, the adjective is assigned a polarity value opposite from its default one (e.g. *too high* is *negative*). We employ a list of 82 positive and 87 negative valence shifters.

5 Evaluation

Table 1 and Table 2 present the results obtained by the OpAL system in the SemEval 2010 Task 18 competition. The system ranked fifth, with a Micro accuracy of 0.76037 and sixth, with a Macro accuracy of 0.7037.

System name	Micro accuracy
98-35_result	0.942064
437-381_HITSZ_CITYU_Task18_Run1.key	0.936236
437-380_HITSZ_CITYU_Task18_Run2.key	0.93315
53-211_dsaa	0.880699
186-325_OpAL_results.txt	0.76037
291-389_submission4.txt	0.724717
291-388_submission3.txt	0.715461
437-382_HITSZ_CITYU_Task18_Run3	0.665752

Table 1: Results - top 8 runs (micro accuracy)

System name	Macro accuracy
437-380_HITSZ_CITYU_Task18_Run2.key	0.957881
437-381_HITSZ_CITYU_Task18_Run1.key	0.953238
98-35_result	0.929308
53-211_dsaa	0.861964

³ <http://webdocs.cs.ualberta.ca/~lindek/minipar.htm>

⁴ <http://www.cs.waikato.ac.nz/ml/weka/>

291-388_submission3.txt	0.755387
186-325_OpAL_results.txt	0.703777
291-389_submission4.txt	0.698037
460383_New_Task18_Chinese_test_pos_QiuLikun_R.rar	0.695448

Table 2: Results – top 8 runs (macro accuracy)

Since the gold standard was not provided, we were not able to perform an exhaustive analysis of the errors. However, from a random inspection of the system results, we could see that a large number of errors was due to the translation – through which modifiers are placed far from the word they determine or the words are not translated with their best equivalent.

6 Conclusions and future work

In this article we presented our approach towards the disambiguation of polarity ambiguous adjectives depending on the context in which they appear. The OpAL system’s run was based on three subcomponents working in English – one assessing the overall polarity of the context using an opinion mining system, the second assessing the local polarity using Google queries formed by expressions containing the noun present in the context before the adjective to be classified and the third one evaluating contextual polarity based on the adjective’s default value and the modifiers around it. The final output is based on the vote given by the majority of the three components. The approach had a good performance, the OpAL system run ranking fifth among 16 runs. Future work includes the separate evaluation of the three components and their combination in a unique approach, using machine learning, as well as a thorough assessment of errors that are due to translation.

References

- Balahur, A. and Montoyo, A. 2008a. *A feature-driven approach to opinion mining and classification*. In Proceedings of the NLPKE 2008.
- Balahur, A. and Montoyo, A. 2008b. *Applying a culture dependent emotion triggers database for text valence and emotion classification*. *Procesamiento del Lenguaje Natural*, 40(40).
- Boldrini, E., Balahur, A., Martínez-Barco, P., and Montoyo, A. 2009. *EmotiBlog: an annotation scheme for emotion detection and analysis in non-traditional textual genres*. In Proceedings of the 5th International Conference on Data Mining (DMIN 2009).
- Cerini, S., Compagnoni, V., Demontis, A., Formentelli, M., and Gandini, G. 2007. *Micro-WNOP: A gold standard for the evaluation of automatically compiled lexical resources for opinion mining*.
- Cilibrasi, D. and Vitanyi, P. 2006. *Automatic Meaning Discovery Using Google*. *IEEE Journal of Transactions on Knowledge and Data Engineering*.
- Esuli, A. and Sebastiani, F. 2006. *SentiWordNet: a publicly available resource for opinion mining*. In Proceedings of the 6th International Conference on Language Resources and Evaluation.
- Hu, M. and Liu, B. 2004. *Mining Opinion Features in Customer Reviews*. In Proceedings of Nineteenth National Conference on Artificial Intelligence AAAI-2004.
- Pang, B. and Lee, L. 2008. *Opinion mining and sentiment analysis*. *Foundations and Trends in Information Retrieval* 2(1-2), pp. 1–135, 2008
- Pang, B., Lee, L., and Vaithyanathan, S. 2002. *Thumbs up? Sentiment classification using machine learning techniques*. In Proceedings of EMNLP-02, the Conference on Empirical Methods in Natural Language Processing.
- Popescu, A. M. and Etzioni, O. 2005. *Extracting product features and opinions from reviews*. In Proceedings of HLTEMNLP 2005.
- Stone, P., Dumphy, D. C., Smith, M. S., and Ogilvie, D. M. 1966. *The General Inquirer: A Computer Approach to Content Analysis*. The MIT Press.
- Strapparava, C. and Valitutti, A. 2004. *WordNet-Affect: an affective extension of WordNet*. In Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004).
- Turney, P. 2002. *Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews*. In Proceedings 40th Annual Meeting of the Association for Computational Linguistics.
- Whitelaw, C., Garg, N., and Argamon, S. 2005. *Using appraisal groups for sentiment analysis*. In Proceedings of the CIKM 2005.
- Wiebe, J. (1994). *Tracking point of view in narrative*. *Computational Linguistics*, 20.
- Wilson, T., Wiebe, J., and Hoffmann, P. 2005. *Recognizing contextual polarity in phrase-level sentiment analysis*. In Proceedings of HLT-EMNLP 2005.
- Wu, Y., Jin, P. 2010. *SemEval-2010 Task 18: Disambiguating Sentiment Ambiguous Adjectives*. In Proceedings of the SemEval 2010 Workshop, ACL 2010.