Ontology-Enhanced Aspect Detection for Sentiment Analysis

Kim Schouten, Flavius Frasincar, and Franciska de Jong
Erasmus University Rotterdam
P.O. Box 1738, NL-3000 DR Rotterdam, the Netherlands
schouten@ese.eur.nl, frasincar@ese.eur.nl, f.m.g.dejong@eshcc.eur.nl

1 Introduction

Aspect-based sentiment analysis aims to find the topics, or aspects, that are discussed in a text, and compute the sentiment that is expressed about those aspects. While the current trend is to use more and more data, there are benefits of using less data. Hence, we propose a knowledge-driven approach that needs only a fraction of the training data because of the common domain knowledge available to the classifier.

2 Method and Evaluation

For the task of aspect detection, we train a binary linear Support Vector Machine (SVM) for each aspect category label that can appear, so it is possible to predict zero or more aspects per sentence. The features consist of lemmatized words (base), WordNet synsets (+S), and ontology concepts (+O).

The ontology consists of three main parts: a domain part, which models the aspects in the given domain, a sentiment part, which models the sentiment values (i.e., positive and negative), and a *SentimentExpression* class which links the domain concepts to sentiment words and sentiment values. The ontology is lexicalized by means of data property added to each individual. Since we want to use both the subclass inference and object properties, we choose to model all domain concepts as having both the class and the individual role. In Fig. 1, the sentiment expression for "cold beer" is shown with its related concepts. Note that the ellipse around the *Beer* class and the *Beer* individual denotes the fact that those are two roles of the same concept. This ontology design allows us to infer the presence of a domain concept if we encounter a sentiment word that has that domain concept as its target. For example, upon finding the word "delicious" we can infer the presence of *Sustenance* (i.e., food and drinks).

We compare the performance of different variants of the method, averaging the performance over 10 runs of 10-fold cross-validation with randomly assigned folds. The cross-validated results are on the official training data of SemEval-2015. Using the training data and official test data, we also show the in-sample and out-of-sample F_1 scores in the last two columns. To get comparable results with the baseline, our method needs only 20% of the training data.

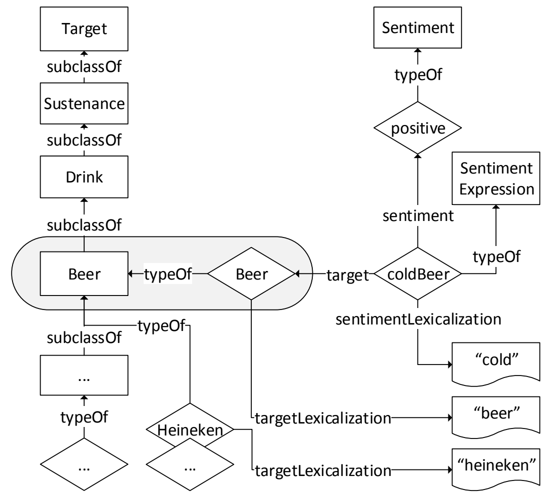

Figure 1: Snippet of the ontology showing a sentiment expression and its related concepts

Table 1: The performance on the aspect detection task

	avg. F_1	st.dev.	in-F_1	out-F_1
base	0.5749	0.0057	0.803	0.5392
+S	0.6317	0.0039	0.896	0.5728
+O	0.6870	0.0026	0.858	0.6125
+SO	0.6981	0.0040	0.920	0.6281

3 Conclusion

In this paper we presented an aspect detection method for sentiment analysis that uses ontologies to limit its dependence on training data. The experiments show that ontologies are indeed useful for aspect-based sentiment analysis, with an improvement of over 10% when comparing to a standard bag-of-words approach. Future work includes expanding the ontology, for example by automatically extracting facts from the Web.

Acknowledgments

This work has been published as "Kim Schouten, Flavius Frasincar and Franciska de Jong. Ontology-Enhanced Aspect-Based Sentiment Analysis. In *17th International Conference on Web Engineering (ICWE 2017)*, volume 10360 of *Lecture Notes in Computer Science*, pages 302-320. Springer, 2017."