

iTac: Aspect Based Sentiment Analysis using Sentiment Trees and Dictionaries

Fritjof Bornebusch¹, Glaucia Cancino¹, Melanie Diepenbeck¹, Rolf Drechsler^{1,2}, Smith Djomkam¹, Alvine Nzeungang Fansou¹, Maryam Jalali¹, Marc Michael¹, Jamal Mohsen¹, Max Nitze¹, Christina Plump¹, Mathias Soeken^{1,2}, Fred Tchambo¹, Toni¹, Henning Ziegler¹

¹ Faculty of Mathematics and Computer Science, University of Bremen, Germany

² Cyber-Physical Systems, DFKI GmbH, Bremen, Germany

itac@cs.uni-bremen.de

Abstract

This paper describes our approach for the fourth task of the *SemEval 2014 challenge: Aspect Based Sentiment Analysis*. Our system is designed to solve all four subtasks: (i) identifying aspect terms, (ii) determining the polarity of an aspect term, (iii) detecting aspect categories, and (iv) determining the polarity of a predefined aspect category. Our system is based on the Stanford sentiment tree.

1 Introduction

Online reviewing, rating, and recommendation have become quite popular nowadays. Based on online reviews and rating, people may decide whether to buy a certain product or visit a certain place (restaurant, shop, etc.). Due to the increasing number of reviews, an automatic system is needed that can evaluate these reviews as positive, negative, or neutral.

In this paper, we propose a system for the fourth task of the SemEval 2014 challenge (*Aspect Based Sentiment Analysis*). The target is to identify aspects of given target entities and to determine the sentiment that is expressed towards each aspect in terms of a polarity. The problem has been divided into four different subtasks: (i) extracting aspects from a given sentence, (ii) determining the polarity of each aspect, (iii) matching suitable aspect categories and (iv) identifying the polarity of these categories.

2 Related Work

There are several different approaches to perform sentiment analysis on a given sentence. References Turney (2002) and Pang et al. (2002) started

This work is licensed under a Creative Commons Attribution 4.0 International License. Page numbers and proceedings footer are added by the organizers. License details: <http://creativecommons.org/licenses/by/4.0/>

to classify a given sentence to be either *positive* or *negative*. Dave et al. (2003) continued to include the *neutral* semantic orientation to his work. These approaches perform sentiment analysis on a whole sentence and use phrases such as adjectives and adverbs to get a polarity. They collect all these phrases and determine their polarity (e.g. positive, neutral, or negative). Hence, it differs from our work that performs sentiment analysis based on each aspect term.

Another approach by Snyder and Barzilay (2007) tries to perform aspect based sentiment analysis, which performs sentiment analysis for various aspects for a given restaurant. Our work differs from their approach and is more closely related to Hu and Liu (2004). Individual parts of the sentence are classified separately since different parts can express different polarities. But the authors only consider product features instead of aspect terms. Aspect terms can be product features but they can also include conditions such as *ambience* that influences an opinion which have not been addressed in Hu and Liu (2004).

3 Preliminaries

Our system is based on *Natural Language Processing (NLP)* libraries such as the *Stanford CoreNLP*.¹ The system is heavily based on the Stanford sentiment tree.

3.1 Stanford sentiment tree

The *sentiment treebank* introduced by Socher et al. (2013) was developed at the University of Stanford to predict the sentiment of movie reviews. It contains approximately 12,000 sentiment annotated parse trees of movie reviews. The sentiment prediction can determine five sentiment classes (very negative, negative, neutral, positive, very positive) using a recursive neural tensor network trained on

¹<http://nlp.stanford.edu/software/corenlp.shtml>

Table 1: Removed word categories with examples.

Category	Example
person	husband, wife, mother, boyfriend
time	date, year, month, Monday-Sunday
location	NYC, Manhattan, street, Avenue
misstaged	everything, something, none, some, any

the sentiment treebank. We aggregate the sentiment classes into three classes (negative, neutral, positive).

4 Implementation

Our system is divided into four subsystems that are described separately in the following section. Although described separately, some subtasks depend on each other (e.g. Aspect Category Extraction and Aspect Category Polarity).

4.1 Aspect term extraction

The aim of this subtask is to find aspect terms that are discussed in a given sentence. Our approach follows an idea presented by Hu and Liu (2004). A word in a given sentence is considered to be an aspect term if it satisfies the following three conditions.

C1.1 It is tagged as a noun (tagged with NN, NNS, NNP, or NNPS).

C1.2 It is one of the 20% most common nouns of all given sentences.

C1.3 It does not belong to a forbidden word category (listed in Table 1).

Following this extraction, adjacent aspect terms are combined to multi-word aspect terms.

Example 1 “My wife bought it and was very happy, especially with the hard drives and battery life.” The result of the rule application is shown in Table 2. When multi-word aspect terms are considered, battery and life are combined to a single term. The row indicated by terms shows the extracted aspect terms of the sentence. In the last row gold terms are compared to actual aspect terms given by the training data.

The results of our system are shown in Table 3. These results could be improved by using typed dependencies. The use of the *adjectival modifier (amod)* and the *noun compound modifier (nn)* relations can help to improve finding multi-word aspect terms.

Table 2: Rule-satisfaction for example.

Rule	Result			
found nouns	wife	drives	battery	life
frequent noun?	✓	✓	✓	✓
non-forbidden?	x	✓	✓	✓
terms		drives	battery life	
gold terms		hard drives	battery life	

Table 3: Results for term extraction.

Domain	Precision	Recall	F-measure
Laptop	0.23	0.25	0.24
Restaurant	0.37	0.40	0.38

4.2 Aspect term polarity

After extracting the aspect term from the sentence the next task is to predict its polarity. For this task we are using the Stanford sentiment tree.

The sentiment tree is designed to predict the sentiment of a whole sentence. Because the sentiment tree contains polarities for every node of the parse tree it is reasonable to use it for aspect sentiment prediction.

Our algorithm examines the sentiment tree nodes to predict the polarity of an aspect. The following outlines the basic steps for aspect sentiment prediction.

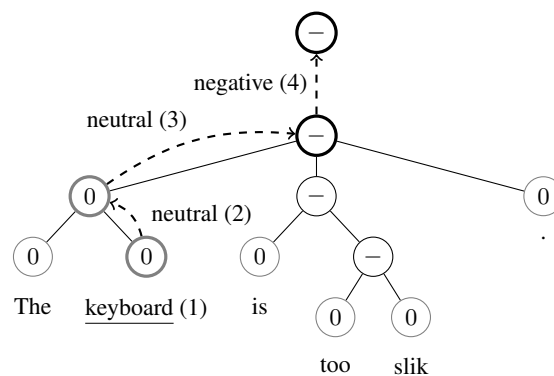


Figure 1: Example of the sentiment tree algorithm for the sentence “The keyboard is too slik.”.

1. Create the sentiment tree for the sentence and fetch the node of the aspect term stem.
2. Traverse the tree from that node up to the root. The first non-neutral polarity on the path from the node to the root node is chosen.
3. If the algorithm reaches the root node without finding a non-neutral polarity, the aspect term is predicted as neutral.

Table 4: Results for term polarity.

Domain	Prec.	Rec.	F-measure	Accuracy
Laptop				0.52
- negative	0.31	0.79	0.45	
- neutral	0.33	0.09	0.15	
- positive	0.79	0.65	0.72	
Restaurant				0.62
- negative	0.35	0.78	0.48	
- neutral	0.25	0.05	0.08	
- positive	0.83	0.75	0.79	

Example 2 Figure 1 illustrates the algorithm for the sentence “The keyboard is too slick.”. The aspect term keyboard is underlined. The algorithm starts at the keyboard node (denoted with 1) and examines the parent node (2). Since the parent node has a neutral polarity, the root node needs to be examined (3). Due to the negative polarity of the root node, the aspect term keyboard is negative (4).

The results of the algorithm with the test data set are shown in Table 4. We got quite good results for negative and positive aspect terms. But there are problems to predict neutral aspect terms, due to the fact that the sentiment tree rarely predicts neutral polarities. Overall our accuracy is nearly 10 percent points above the ABSA baselines.

4.3 Aspect category detection

This section describes the approach for the third subtask that identifies aspect categories discussed in a given sentence, using a predefined set of aspect categories, such as *food*, *service*, *ambience*, *price*, and *anecdotes/miscellaneous* as a neutral category. Our approach is twofold, depending on whether the sentence contains aspect terms or not.

Sentences with aspect terms. We illustrate our approach with the following example sentence.

Example 3 Consider the sentence “Even though it is good seafood, the prices are too high.” with the predefined aspects terms *seafood* and *price*.

1. If the aspect term is a category, it can be directly assigned as a category. In this example the category *price* is present and will be assigned.
2. Dishes are very challenging to detect as an aspect term. For that problem we added a list of dishes scraped from Wikipedia to detect them. If a noun is not part of the list we search DuckDuckGo² for the description of that noun

²<https://duckduckgo.com>

Table 5: Result for category extraction.

Domain	Precision	Recall	F-measure
Restaurant	0.63	0.52	0.59

and check whether it is a dish. If it is a dish, then the category food is assigned.

3. For unassigned aspect terms, the similarity between aspect terms and all categories will be calculated. For this purpose, *RiTa.WordNet similarity* has been used. If the path length is smaller than 0.4 (with the help of the training data we experimentally determined the best comparison value) the aspect term is assigned to the category. In our example *seafood* is similar to *food* and therefore the category is *food*.
4. If no aspect category could be found, the category is *anecdotes/miscellaneous*.

Sentences without aspect term. The third step from the previous approach is executed for all nouns in the sentence. But the threshold is decreased to 0.19 to reduce the number of recognized categories. If no similarity falls below the threshold, the category is *anecdotes/miscellaneous*.

The results of the third subtask are presented in Table 5. Although the presented results are moderately good, there exist some issues worth to be considered here: Using WordNet (Miller, 1995), it is only possible to find the similarity between two concepts and not a group of concepts. For example *Japanese Tapas* with *food* would not work. Furthermore, WordNet only recognizes the similarity between words of the same part of speech, it means many possible relations between verbs and nouns, and also adjectives and nouns are missing. Also, we were not able to calculate the similarity between a term and the default category.

4.4 Category polarity

This section describes the last subtask which aims to find the polarity of an aspect category for a given sentence. For the given aspect category which can be *food*, *service*, *ambience*, *price*, or *anecdotes/miscellaneous*, the task is to find its polarity. This subtask is applied only for the topic *restaurant*. The second and third subtask must have been solved since their evaluations are required to classify which aspect term belongs to which aspect category. In the third subtask all aspect terms are

grouped in categories and in the second one the aspect terms are set with their polarities, which we use to calculate how many times a specific polarity is chosen under the same aspect category. Then we can assign a polarity to a specific aspect category. In order to find the polarities of an aspect category we carefully analyzed the training data and defined a set of rules to find all possible cases. We will discuss these rules in the following.

R4.1 If the aspect term polarities of the same category are equal, then their polarity is tagged as the category polarity.

Example 4 “Prices are higher to dine in and their chicken tikka marsala is quite good.” The found aspect terms in this sentence are Prices which is negative and chicken tikka marsala which is positive. Both aspect terms belong to different categories. The category food (chicken tikka marsala) is positive and the category price (Prices) is negative.

R4.2 If one of the aspects of a specified category is neutral, it has no influence on the polarity of a category, as long as at least one other polarity exists. The polarities of all other aspect terms will determine the polarity of a specific category.

Example 5 “Our server checked on us maybe twice during the entire meal.” In this sentence the following aspect terms are found: server as negative and meal as neutral. Both aspect terms belong to the same category service, so the category service has the value negative.

R4.3 If the aspect term polarities under a same category are both *positive* and *negative*, then the category polarity is tagged as *conflict*.

Example 6 As an example consider the sentence: “The sweet lassi was excellent as was the lamb chettinad and the garlic naan but the rasamalai was forgettable.” Here four aspect terms were found: sweet lassi, lamb chettinad, and garlic naan with positive polarities but rasamalai has a negative polarity. This results in a conflict polarity for the category food.

R4.4 If the found category was annotated as *anecdotes/miscellaneous* but no aspect term was found in the second subtask, then we use the sentiment tree. It generates a specific polarity for the entire sentence which we define as the category’s polarity.

Example 7 The sentence: “A guaranteed delight!” has no aspect term. Using the sentiment tree the

Table 6: Results for category polarity.

Domain	Prec.	Rec.	F-measure	Accuracy
Restaurant				0.63
- conflict	0.08	0.10	0.09	
- negative	0.45	0.73	0.56	
- neutral	0.24	0.17	0.20	
- positive	0.86	0.70	0.77	

polarity for the category anecdotes/miscellaneous is positive.

We applied our approach on the training data. The results are shown in Table 6. We achieved an F-measure of 0.85 for the *positive* polarity. Our accuracy is 0.56 which is not a good achievement in comparison to other submissions in this subtask. The possible reason for this result could be that the first subtask also did not reach good accuracy measures.

5 Conclusion & future works

This paper describes our system to solve the individual subtasks by using the Stanford CoreNLP, RiTa.WordNet (Guerini et al., 2013) and a food database developed by ourselves. These libraries offer methods to classify sentences and determine the polarities.

Through the usage of the library based methods, it is not possible to take effect to the result. At this point other libraries such as NLTK³ could help to increase it. They offer the possibility to train several classifiers with own data. But the classifier are not domain independent, because they need to be trained with sentences that belong to a specific domain, e.g. laptop or restaurant, in order to get the right polarity.

Our approach is more domain independent, because we do not need any domain to calculate the right polarities. That’s why we can use our tool to process sentences of any domain, without further changing the algorithms.

In the future, we expect progress towards the following directions. First, we want to improve the identification of aspect terms which consist of more than two consecutive nouns. Second, we want to identify aspect terms which are not available as a part of the sentence. Finally, improvements to determine polarity of sentences with unclear context (i.e. the absence of adjectives).

³<http://www.nltk.org/>

References

- Kushal Dave, Steve Lawrence, and David M. Pennock. 2003. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *WWW*, pages 519–528.
- Marco Guerini, Lorenzo Gatti, and Marco Turchi. 2013. Sentiment analysis: How to derive prior polarities from SentiWordNet. In *EMNLP*, pages 1259–1269.
- Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *KDD*, pages 168–177.
- George A. Miller. 1995. WordNet: A lexical database for english. *Commun. ACM*, 38(11):39–41.
- Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up? sentiment classification using machine learning techniques. In *EMNLP*, pages 79–86.
- Benjamin Snyder and Regina Barzilay. 2007. Multiple aspect ranking using the Good Grief algorithm. In *HLT-NAACL*, pages 300–307.
- Richard Socher, Alex Perelygin, Jean Y. Wu, Jason Chuang, Christopher D. Manning, Andrew Y. Ng, and Christopher Potts. 2013. Recursive deep models for semantic compositionality over a Sentiment Treebank. In *EMNLP*, pages 1631–1642.
- Peter D. Turney. 2002. Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews. In *ACL*, pages 417–424.