

SentiSys at SemEval-2016 Task 5: Opinion Target Extraction and Sentiment Polarity Detection

Hussam Hamdan

Aix-Marseille University
hamdan.hussam@gmail.com

Abstract

This paper describes our contribution in Opinion Target Extraction and Sentiment Polarity sub-tasks of SemEval 2016 ABSA task. A Conditional Random Field model has been adopted for opinion target extraction. A Logistic Regression model with a weighting schema of positive and negative labels has been used for sentiment polarity. Our submission for opinion target extraction is ranked second among the constrained systems which do not use additional resources and sixth over 19 submissions among the constrained and unconstrained systems in English restaurant reviews. Our submission for Sentiment Polarity is ranked eighth over 22 submissions on the laptop reviews.

1 Introduction

Classifying opinion texts at document or sentence levels is not sufficient for applications which need to identify the opinion targets. Even if the document is about one entity, many applications need to determine the opinion about each aspect of the entity. A user may express a positive opinion towards a restaurant, but he may have a negative opinion towards some aspects as the ambiance. Therefore, we need to identify the aspects and determine whether the sentiment is positive, negative or neutral towards each one. This task is called Aspect-Based Sentiment Analysis or Feature-Based opinion mining as called in the early work (Hu and Liu, 2004). Aspect-Based Sentiment Analysis is composed of four subtasks:

1. Opinion Target Expression Extraction

Opinion Target Expression is a linguistic expression used in a given text to refer to an aspect of the reviewed entity. This subtask aims at identifying all the aspect terms present in a given set of sentences with pre-identified entities such as restaurants, laptops. An opinion target names a particular aspect of the target entity. For example:

”I liked the service and the staff, but not the
food”
”The hard disk is very noisy”

The *service*, *staff* and *food* are opinion target expressions. *hard disk* is multi-word opinion target expression which will be treated as a single term.

2. Aspect Sentiment Detection

Each identified opinion target has to be assigned to one of the following polarity labels: positive, negative or neutral.

For example:

Input: ”I hated their fajitas, but their salads
were great”
Output: {fajitas: negative, salads: positive}

3. Aspect Category Detection

This subtask aims at identifying the aspect categories discussed in a given sentence from a predefined set of aspect categories such as

price and food. Aspect categories are typically coarser than the opinion targets, and they do not necessarily occur as terms in a given sentence.

For example, given the set of aspect categories of restaurant entity {food, service, price, ambiance}:

”The restaurant was too expensive”

The aspect category is {price}.

”The restaurant was expensive, but the menu was great”

The aspect categories are {price, food}.

4. Aspect Category Polarity

Given a set of pre-identified aspect categories such as {food, price}, this subtask aims at determining the polarity of each aspect category. For example:

”The restaurant was too expensive”

{price: negative}

”The restaurant was expensive, but the menu was great”

{price: negative, food: positive}

In this paper, we focus on Opinion Target Extraction (OTE) and Sentiment Polarity towards a target or a category. The description of each subtask is provided by ABSA organizers (Pontiki et al., 2016). For OTE, a CRF model is proposed with several groups of features including syntactic and lexical features. For polarity detection, a logistic regression classifier is trained with the weighting schema for positive and negative labels and several groups of features are extracted including lexical, syntactic, semantic and sentiment lexicon features.

The rest of this paper is organised as follows. Section 2 outlines existing work in target extraction and polarity detection. Section 3 describes our system for opinion target extraction. Polarity detection is presented in Section 4. Section 5 shows the conclusion and the future work.

2 Related Work

Aspect-Based Sentiment Analysis consists of several subtasks. Some studies have proposed different methods for aspect detection and sentiment polarity analysis, others have proposed joint models in order to obtain the aspect and their polarities from the same model, these last models are generally unsupervised.

The early work on opinion target detection from on-line reviews presented by (Hu and Liu, 2004) used association rule mining based on Apriori algorithm (Agrawal and Srikant, 1994) to extract frequent noun phrases as product features. For polarity detection, they used two seed sets of 30 positive and negative adjectives, then WordNet has been used to find and add the synonyms of the seed words. Infrequent product features or opinion targets had been processed by finding the noun related to an opinionated word.

Opinion Digger (Moghaddam and Ester, 2010) also used the Apriori algorithm to extract the frequent opinion targets. The kNN algorithm is applied to estimate the aspect rating scaling from 1 to 5 stands for (Excellent, Good, Average, Poor, Terrible).

Supervised methods use normally Conditional Random Fields (CRF) or Hidden Markov models (HMM). (Jin and Ho, 2009) applied a HMM model to extract opinion targets using the words and their part-of-speech tags in order to learn a model, then unsupervised algorithm for determining the opinion targets polarity using the nearest opinion word to the opinion target and taking into account the polarity reversal words (such as not).

A CRF model was used by (Jakob and Gurevych, 2010) with the following features: tokens, POS tags, syntactic dependency (if the opinion target has a relation with the opinionated word), word distance (the distance between the word in the closest noun phrase and the opinionated word), and opinion sentences (each token in the sentence containing an opinionated expression is labeled by this feature), the input of this method is also the opinionated expressions, they use these expressions for predicting the opinion target polarity using dependency parsing for retrieving the pair target-expression from the training set. (Hamdan et al., 2014; Hamdan et al.,

2015a) also applied a CRF model with different features .

Unsupervised methods based on LDA (Latent Dirichlet allocation) have been proposed. (Brody and Elhadad, 2010) used LDA to figure out the opinion targets, determined the number of topics by applying a clustering method, then they used a similar method proposed by (Hatzivassiloglou and McKeown, 1997) to extract the conjunctive adjectives, but not the disjunctive due to the specificity of the domain.

(Lin et al., 2012) proposed Joint model of Sentiment and Topic (JST) which extends the state-of-the-art topic model (LDA) by adding a sentiment layer, this model is fully unsupervised and it can detect sentiment and topic simultaneously.

(Wei and Gulla, 2010) modeled the hierarchical relation between product aspects. They defined Sentiment Ontology Tree (SOT) to formulate the knowledge of hierarchical relationships among product attributes and tackled the problem of sentiment analysis as a hierarchical classification problem. Unsupervised hierarchical aspect Sentiment model (HASM) was proposed by (Kim et al., 2013) to discover a hierarchical structure of aspect-based sentiments from unlabeled online reviews.

3 Opinion Target Expression (OTE)

The objective of opinion target extraction is to extract all opinion target expressions in a restaurant review, opinion target could be a word or multiple words. This extraction consists of the following steps:

1. Review Segmentation

This step segments each review into sentences. In restaurant dataset, we already have the sentences.

2. Sentence Tokenizing

Each sentence is tokenized to get the terms. One can consider the spaces as separators or use a more complex tokenizer. We tokenize each sentence using NTLK tokenizer¹ which extracts the words, numbers and punctuations.

¹<http://www.nltk.org/api/nltk.tokenize.html>

3. Sentence Tagging

Each term in the sentence should be tagged in order to be presented to a tagging classifier. We choose the IOB notation for representing each sentence in the review. Therefore, we distinguish the terms at the Beginning, the Inside and the Outside of opinion target expression. For example, for the following review sentence:

"But the staff was so horrible to us."

Where *staff* is opinion target. The tag of each term will be:

But:O the:O staff:B was:O so:O horrible:O
to:O us:O.

4. Feature Extraction

This is the main step of opinion target extraction. For representing each term, we extract the following features:

- The term itself.
- POS: We use NLTK parser² to attach a part of speech tag to each term.
- word shape: the shape of each character in the word (capital letter, small letter, digit, punctuation, other symbol)
- word type: the type of the word (uppercase, digit, symbol, combination)
- Prefixes (all prefixes having length between one to four).
- Suffixes (all suffixes having length between one to four).
- Stop word: if the word is a stop word or not.

In addition to the previous features, we extract for each term the following features:

- The two preceding and three subsequent terms of the actual term.
- The value of each two successive features in the the range -2,2 (the previous and subsequent two terms of actual term) for the following features: term, word POS, word

²<http://www.nltk.org/book/ch05.html>

shape, word type. For example, for POS feature we extract:

```
pos[-2]—pos[-1]=DT—JJ
pos[-1]—pos[0]=JJ—NN
pos[0]—pos[1]=NN—,
pos[1]—pos[2]=,—C
```

- We also extract the value of each three successive features in the the range -1,1 for the two features: term POS and term. For example, for the feature *term* we extract:

```
term[-2]—term[-1]—term[0]=a—good—place
term[-1]—term[0]—term[1]=good—place—,
```

5. Training Method

we have used a Conditional Random Fields (CRF) which receives the feature representation of each term in each sentence and builds a tagging model in order to use it for predicting the tags of the new sentences.

3.1 Experiments

The data set is extracted from restaurant reviews, provided by SemEval 2016 ABSA organisers (Pontiki et al., 2016) where each review is composed of several sentences and each sentence may contain several OTEs. CRFsuite tool is used for this experiment. This tool is fast in training and tagging (Okazaki, 2007).

Our submission is ranked second among the constrained systems and sixth over all 19 systems with the F1 score. Table 2 shows the result of our system.

Experiment	F1 Score
Our System	66.545
Baseline	44.071

Table 1. The results of OTE slot.

4 Sentiment Polarity

For a given set of aspect terms within a sentence, we determine whether the polarity of each aspect term is positive, negative or neutral. For example, the system should extract the polarity of *fajitas* and *salads* in the following sentence: *"I hated their fajitas, but their salads were great"*, *fajitas*: negative and *salads*: positive.

This sub-task can be seen as sentence level or phrase level sentiment Analysis. We should determine the polarity, which could be positive, negative, neutral. We use the system described by (Hamdan et al., 2015b). Thus, We propose to use a logistic regression classifier with weighting schema of positive and negative labels with the following features:

- Word n-grams Features

Unigrams and bigrams are extracted for each word in the text without any stemming or stop-word removing, all terms with occurrence less than 3 are removed from the feature space.

- Sentiment Lexicon-based Features

The system extracts four features from the manually constructed lexicons (Bing Liu Lexicon (Hu and Liu, 2004) and MPQA subjectivity Lexicon (Wilson et al., 2005)) and six features from the automatic ones (NRC Hashtag Sentiment Lexicon (Mohammad, 2012), Sentiment Lexicon (Hamdan et al., 2015b)). For each text the number of positive words, the number of negative ones, the number of positive words divided by the number of negative ones and the polarity of the last word are extracted from manual constructed lexicons. In addition to the sum of the positive scores and the sum of the negative scores from the automatic constructed lexicons.

- Negation Features

The rule-based algorithm presented in Christopher Potts' Sentiment Symposium Tutorial is implemented. This algorithm appends a negation suffix to all words that appear within a negation scope which is determined by the negation key and a certain punctuation. All these words are added to the feature space.

- Brown Cluster Features

Each word in the text is mapped to its cluster in Brown clusters, 1000 features are added to feature space where each feature represents the number of words in the text mapped to each cluster. The 1000 clusters is provided in Twitter Word Clusters of CMU ARK group which were constructed from approximately 56 million tweets.

- Category Feature

We also added the category of each OTE as a feature to the feature space.

4.1 Experiments

We trained a L1-regularized Logistic regression classifier implemented in LIBLINEAR, which has given good results in several papers (Hamdan et al., 2015b) (Hamdan et al., 2015a). The classifier is trained on the training data set using the previous features with the three polarities (positive, negative, and neutral) as labels. A weighting schema is adapted for each class, we use the weighting option $-w_i$ which enables a use of different cost parameter C for different classes. Since the training data is unbalanced, this weighting schema adjusts the probability of each label. Thus we tuned the classifier in adjusting the cost parameter C of Logistic Regression, weight w_{pos} of positive class and weight w_{neg} of negative class.

We used the 1/10 of training data set for tuning the three parameters in Laptop reviews, all combinations of C in range 0.1 to 4 by step 0.1, w_{pos} in range 1 to 8 by step 0.1, w_{neg} in range 1 to 8 by step 0.1 are tested. The combination $C=0.3$, $w_{pos}=1.2$, $w_{neg}=1.9$ have been chosen. Table 2 shows the results of our system on the Laptop data set. Our system is ranked eighth over 22 submissions.

Experiment	Accuracy
Laptops	
Our system	74.282
Baseline	44.071

Table 2. Results of sentiment polarity in laptops reviews.

5 Conclusion

We have built two systems for opinion target extraction of restaurant data set, and sentiment polarity analysis for laptops. We have used supervised tagger for OTE, trained a CRF model with several features. A Logistic regression classifier is used for sentiment polarity where we adopted a weighting schema.

References

- Agrawal, R. and Srikant, R. (1994). Fast Algorithms for Mining Association Rules in Large Databases. In *Proceedings of the 20th International Conference on Very Large Data Bases, VLDB '94*, pages 487–499, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.
- Brody, S. and Elhadad, N. (2010). An Unsupervised Aspect-sentiment Model for Online Reviews. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, HLT '10*, pages 804–812, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Hamdan, H., Bellot, P., and Bechet, F. (2014). Supervised Methods for Aspect-Based Sentiment Analysis. In *Proceedings of the Eighth International Workshop on Semantic Evaluation (SemEval 2014)*.
- Hamdan, H., Bellot, P., and Bechet, F. (2015a). Lsislif: CRF and Logistic Regression for Opinion Target Extraction and Sentiment Polarity Analysis. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 753–758, Denver, Colorado. Association for Computational Linguistics.
- Hamdan, H., Bellot, P., and Bechet, F. (2015b). Sentiment Lexicon-Based Features for Sentiment Analysis in Short Text. In *Proceeding of the 16th International Conference on Intelligent Text Processing and Computational Linguistics*, Cairo, Egypt.
- Hatzivassiloglou, V. and McKeown, K. R. (1997). Predicting the Semantic Orientation of Adjectives. In *Proceedings of the Eighth Conference on European Chapter of the Association for Computational Linguistics, EACL '97*, pages 174–181, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Hu, M. and Liu, B. (2004). Mining and Summarizing Customer Reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '04*, pages 168–177, New York, NY, USA. ACM.
- Jakob, N. and Gurevych, I. (2010). Extracting Opinion Targets in a Single- and Cross-domain Setting with Conditional Random Fields. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, EMNLP '10*, pages 1035–1045, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Jin, W. and Ho, H. H. (2009). A Novel Lexicalized HMM-based Learning Framework for Web Opinion mining. NOTE FROM ACM: A Joint ACM Conference Committee Has Determined That the Authors of This Article Violated ACM's Publication Policy on Simultaneous Submissions. Therefore ACM Has Shut off

- Access to This Paper. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pages 465–472, New York, NY, USA. ACM.
- Kim, S., Zhang, J., Chen, Z., Oh, A., and Liu, S. (2013). A Hierarchical Aspect-Sentiment Model for Online Reviews. In *Proceedings of The Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13)*. AAAI.
- Lin, C., He, Y., Everson, R., and Ruger, S. (2012). Weakly Supervised Joint Sentiment-Topic Detection from Text. *IEEE Trans. on Knowl. and Data Eng.*, 24(6):1134–1145.
- Moghaddam, S. and Ester, M. (2010). Opinion Digger: An Unsupervised Opinion Miner from Unstructured Product Reviews. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management, CIKM '10*, pages 1825–1828, New York, NY, USA. ACM.
- Mohammad, S. (2012). #Emotional Tweets. In **SEM 2012: The First Joint Conference on Lexical and Computational Semantics – Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012)*, pages 246–255, Montréal, Canada. Association for Computational Linguistics.
- Okazaki, N. (2007). *CRFsuite: a fast implementation of Conditional Random Fields (CRFs)*.
- Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., Clercq, O. D., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., Jimenez-Zafra, S. M., and Eryigit, G. (2016). SemEval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation, SemEval '16*, San Diego, California. Association for Computational Linguistics.
- Wei, W. and Gulla, J. A. (2010). Sentiment learning on product reviews via sentiment ontology tree. In *Proceedings of the 48th Annual Meeting of the ACL*, pages 404–413. ACL.
- Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J., Choi, Y., Cardie, C., Riloff, E., and Patwardhan, S. (2005). OpinionFinder: A System for Subjectivity Analysis. In *Proceedings of HLT/EMNLP on Interactive Demonstrations, HLT-Demo '05*, pages 34–35, Stroudsburg, PA, USA. Association for Computational Linguistics.