
ASPECT BASED SENTIMENT ANALYSIS

Abhishek Bhatia

bhatiaabhishek8893@gmail.com

Dristanta Nirola

dristantanirola@gmail.com

1 INTRODUCTION

The recent boom of blogs, review sites and especially social networking sites has led to the creation of a large pool of information. This text-based user-generated opinionated content is an invaluable source of information for businesses, enabling them to generate actionable insights into the preferences of customers. Automatically analyzing the large volume of data generated from customer feedback would help businesses tailor products and services to various market segments. Sentiment analysis or opinion mining is a text analysis method to detect polarity (e.g. a positive or negative opinion) within the text. The text can be analyzed at various levels i.e. document level to examine the overall polarity of the document, sentence level in which the document is broken down into individual sentences to obtain a fine grain level of polarity, and aspect level where the polarity of the specific domain-related entities (e.g., laptops, restaurants) and their aspects/attributes (e.g., battery, screen; food, service) are analysed. However, the current approaches focus mainly on detecting the overall polarity of the text [1, 2, 3]. In unstructured product reviews, people tend to express multiple opinions about various features of the product. Depending upon their likes and dislikes some of the opinions expressed about specific features may be positive and some may be negative. In such a case, determining the feature specific opinion will prove more helpful than the overall opinion. Hence, it is vital especially for businesses to analyze product reviews to obtain the sentiment of a particular aspect or a feature of their product which they can improve upon. Aspect extraction and polarity detection are extremely helpful for building recommendation systems, improving relevant product features and catering to customer's specific needs. This forms the basic premise of our project, where our goal is to identify the polarity expressed towards a certain aspect in a user review text.

2 PROBLEM OVERVIEW

Consider the following review text: *"The food was delicious but the service was horrible"*, the customer is expressing positive sentiment about "food" and negative sentiment about "service". An aspect based sentiment classifier would first extract the relevant aspect terms i.e. food, service and classify the sentiment expressed towards each of these aspects i.e. positive, negative. It is, however, important to note that the opinions may be explicitly expressed or implicitly expressed. Likewise, aspects which are facets of target entities can also be categorized into implicit and explicit [4]. Explicit aspect terms are words or phrases that explicitly convey the target entity in the sentence. So, in the above review text, the words "food" and "service" are explicitly mentioned in the sentence. Whereas in the following review text: *"The McSupreme cheeseburger at McDonald's is sooo tasty and extremely cheap as well!!"*, a positive opinion is expressed about cheeseburger and it's price through an implicit aspect clue (IAC) [5].

3 OBJECTIVE

Aspect Based Sentiment Analysis (ABSA) is a branch of sentiment analysis, comprising mainly four broad subtasks (ST); ST1: aspect term extraction, ST2: aspect term polarity, ST3: aspect category detection, and ST4: aspect category polarity. In this report, we focus on the 2nd subtask (ST2) of determining the polarity expressed towards each aspect term in a review text. Formally, the task can be stated as follows:

Subtask 2 : *For a given set of aspect terms within a sentence, determine whether the polarity of each aspect term is positive, negative or neutral.*

For example:

“I loved their fajitas” $\rightarrow \{fajitas : positive\}$

“I hated their fajitas, but their salads were great” $\rightarrow \{fajitas : negative, salads : positive\}$

“The fajitas are their first plate” $\rightarrow \{fajitas : neutral\}$

4 DATA DESCRIPTION

We use the SemEval 2014 dataset [6] consisting of user reviews on two domains: (i) Restaurant reviews, (ii) Laptop. The dataset in both the domains contains over 3000 English sentences each along with annotations for aspect terms (Subtask 1), aspect term polarity (Subtask 2), while the annotations for aspect categories (Subtask 3) and aspect-category specific polarity (Subtask 4) are only available for the restaurant domain. The sentences in the datasets are annotated using XML tags.

5 METHODOLOGY

The following steps briefly describe the methodology we employ:

- **Preprocessing** - Application of text cleaning steps such as stopword, punctuation removal etc
- **Feature engineering** - We experiment with various feature engineering techniques suited for text data such as vectorization techniques like term frequency and document frequency, POS tagging, dependency parsing and the use of pre-trained word embeddings (GloVe, Word2vec)
- **Machine learning models** - We evaluate the performance of various machine learning algorithms like naive bayes, logistic regression, SVM, bagging (Random Forest) and boosting (AdaBoost) algorithms
- **Deep learning models** - We implement different variants of deep RNN based sequence models and discuss their advantages over the simpler machine learning methods in capturing long-range dependencies in textual data

REFERENCES

- [1] Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1–135.
- [2] Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2009. Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational Linguistics*, 35(3):399–433
- [3] Changli Zhang, Daniel Zeng, Jiexun Li, Fei-Yue Wang, and Wanli Zuo. 2009. Sentiment analysis of Chinese documents: From sentence to document level. *Journal of the American Society for Information Science and Technology*, 60(12):2474–2487
- [4] Chen, H.Y. and Chen, H.H., 2016, August. Implicit polarity and implicit aspect recognition in opinion mining. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (pp. 20-25).
- [5] Ivan Omar Cruz-Garcia, Alexander Gelbukh, and Grigori Sidorov. 2014. Implicit aspect indicator extraction for aspect-based opinion mining.
- [6] <http://alt.qcri.org/semeval2014/task4/index.php?id=data-and-tools>.