

A Survey on Opinion Mining Methods from Online Reviews

Vibhuti Patel*, Prof. Mital Panchal

Department of Information Technology, L. D. College of Engineering, Ahmedabad, Gujarat, India

ABSTRACT

Now-a-days e-commerce becomes more mature and number of reviews about product grows as well with it. So, it is difficult for a customer to browse through large number of reviews for item of interest and make a decision about the product. For manufacturer, it's also difficult to keep track of the product reviews. Using opinion mining the process becomes easier for customers as well as manufacturers to keep track of the product reviews. This paper analyzing the existing methods for opinion mining from online customer reviews, and the analysis and conclusion would be helpful in obtaining a method, that could give better opinion mining results and overcoming the existing limitation.

Keywords: opinion mining, e-commerce technology, WTM, POS, Part-of-Speech Tagging, Latent Dirichlet Allocation

I. INTRODUCTION

As e-commerce technology become more and more popular, the convenience of online shopping has attracted more and more people. In order to receive product feedback timely and to update future customer with previous shopping experiences of the same product, it is common for merchants to allow their customers to leave product reviews. As the number of customers increases, the number of reviews about the product grows rapidly. So mining of reviews become important process. Using opinion mining of online reviews customer, who wants to buy the product, can get idea about the product's quality and manufacturer can improve the product on time.

In this paper different kind of opinion mining methods are summarized. In opinion mining some methods give overall mining, i.e. only positive and negative reviews are defined. Some method gives mining result specific to product feature. Some use novel word-based alignment model to extract opinion target and opinion words. Methods are supervised, semi-supervised and unsupervised method.

The purpose of this review paper is to focus on drawbacks and advantages of all these methods and try

to make some robust method. This paper is organised in sections where Section II describes opinion mining, Section III describes the various methods for opinion mining, section IV compares different methods, their advantages and disadvantages, finally we conclude in Section V.

II. METHODS AND MATERIAL

A. Opinion Mining

Opinion mining is extracting people's opinion from the web. It is a field of study that analyses people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. Sometime the word "Sentiment Analysis" is used for opinion mining. [6]

Levels of Opinion Mining

There are three categories for opinion mining task: document level, sentence level and entity and aspect level [6].

Document Level : The task at document-level analysis is to classify whether a whole document express a positive

or negative opinion. This level of opinion mining assumes that each document gives opinion on a single entity (i.e. single product or single topic).

Sentence Level: In this type of analysis the task is to identify whether each sentence gives a positive, negative or neutral sentiment. Neutral usually means no opinion in the sentence. This level of analysis is related to subjective classification, which distinguishes objective sentences that expresses factual information from subjective sentences that expresses subjective views and opinions.

Entity and Aspect Level: Aspect level analysis performs fine-grain analysis. Aspect level is earlier called feature level. Instead of looking at language constructs (documents, sentences, paragraphs, etc.), aspect level directly look for opinion. It is based on idea that an opinion consists of a sentiment (positive or negative) and a target (of opinion).

B. Related Work

Method-1

Gaurav Dubey [1] proposed a simple method for opinion mining. The method exploits the publicly available pool of online product reviews by automatic extracting marketing intelligence from the vast repository of user generated opinions through main three steps: Part-Of-Speech (POS) tagging, Sentiment Analysis through rule mining, and summarizing and displaying the output.

Part-of-Speech Tagging (POST) or lexical set is used to find out the grammatical words in any document or user speech: like noun, verb, adjective etc. This can be done either on basis of definition, e.g. all names are noun like India, or on the basis of context which depends upon the relationship with neighbouring or similar words [7].

Method-2

LiZhen Liu [2] proposed fine-grain approach to extract the product feature. In the mining process first of all LDA (Latent Dirichlet Allocation) model is used to find the features that people are most interested in. After feature generation the strength of the opinion words are calculated. Using this strength, the emotional intensity is divided into six levels: Positive (strength, neutral, weak) and negative (strength, neutral, weak). Then, all features

are ranked according to the frequency of their appearances in the reviews.

Method-3

Kang Liu [3] proposed a novel approach to extract opinion targets based on word based translation model (WTM). In this approach opinion target extraction is composed of two main components: mining association between opinion target and opinion words using word based translation model and opinion words, candidate confidence estimation using Google n-gram corpus. Then, the candidate with higher confidence than the threshold will be extracted as opinion target.

Method-4

A. Jeyapriya [4] proposed supervised learning algorithm for extracting aspects and mining opinions in product reviews. The system uses customer reviews to extract aspect and mine whether given review is positive or negative opinion. First of all, the method removes Stop words. Stop words are words which are most frequently used in English and not useful in text mining. After removal stop words stemming is performed to form root word of a word. Porter stemmer algorithm is used for stemming process. After this POS tagging is done. POS tagging is as explained before. Here Stanford tagger is used for POS tagging. After POS tagging aspect extraction is performed. Nouns are extracted and then its frequency is compared to minimum support count. The word which has higher count than minimum support count is extracted. The Naïve Bayesian algorithm using supervised term counting based approach is used for sentence and aspect orientation. Finally, the system identifies the number of positive and negative opinions of each extracted aspect in customer reviews.

Method-5

Kang Liu [5] proposed method for co-extracting opinion targets and opinion words from online reviews based on the word alignment model. The process of extracting opinion targets/words is viewed as co-ranking process. The opinion relation is identified using word based alignment process. To obtain the optimal alignments and EM based algorithm to train the model. The standard alignment model is usually trained in completely unsupervised manner. So, to improve alignment performance, partial supervision is performed on the

statistic model to incorporate partial alignment links into alignment process.

In next step partial alignment links are obtained by using high-precision syntactic patterns. After that opinion association among words is calculated. After this candidate confidence is estimated using random walking with restart algorithm and penalizing high degree vertices, which may introduce noise. Finally, candidate with higher confidence than minimum threshold is extracted as the opinion target or opinion word.

III. RESULTS AND DISCUSSION

Comparisons of Methods

Here we have studied different methods for opinion mining. Method 1 is simple method and gives overall review mining about the product. This method does not give the review mining around the product feature. Method 1 is rule based method, so it cannot give good results when the reviews are in informal language. Compared to Method 1, Method 2 gives summarization of reviews around the product feature. So, the mining is more useful than Method 1. Though, Method 2 is better than Method 1, it gives summarization result in form of listing all comments. So, user may need to read those comments for better understanding.

Method 4 uses supervised learning, so it has its own importance. Using supervised learning the mining result is more precise than unsupervised methods. It may consume more time than unsupervised methods.

Method 3 and Method 5 use word-based models, Method 3 uses word-based translation model and Method 5 uses word alignment model. Method 3 is simple compared to Method 5. Method 3 is unsupervised method, whereas in Method 5 unsupervised as well as partially supervision is used. Both methods co-extract opinion target and opinion word. Co-extracting is important because we can go through reviews of product for specific feature. In both of the methods, word based model are used, so if the reviews are in informal language then also these methods can give precise results. So Method 3 and Method 5 are more effective methods compared to all the methods sueded here.

IV. CONCLUSION

In this paper we have studied different methods that are used for opinion mining. From section III, it is seen that methods have their own advantages and disadvantages. Method 3 and Method 5 are effective, but in future we can make it more effective using combination of these methods and supervised learning approach.

V. REFERENCES

- [1] Gaurav Dubey, Ajay Rana, Naveen Kumar Shukla, "User Reviews Data Analysis using Opinion Mining on Web", In 2015, IEEE.
- [2] LiZhen Liu, WenTao Wang, HangShi Wang, "Summarizing Customer Reviews Based On Product Feature", In 2012, IEEE.
- [3] Kang Liu, Liheng Xu, Jun Zhao, "Opinion Target Extraction Using Word-Based Translation Model", In 2012, Joint Conference Empirical Methods Natural Language Processing Computer Natural Language Learning
- [4] A.Jeyapriya, C.S.KanimozhiSelvi, "Extracting Aspects and Mining Opinions in Product Reviews using Supervised Learning Algorithm", in 2015, IEEE.
- [5] Kang Liu,Liheng Xu and Jun Zhaao, "Co-Extracting Opinion Targets and Opinion Words from Online Reviews Based on The Word Alignment Model", In 2015, IEEE Transaction.
- [6] Bing Liu, "Sentimental Analysis and Opinion Mining" In 2012, Synthesis Lectures on Human Language Technologies, Morgan & Claypool Publishers.
- [7] Bo Pang, Lillian Lee, "opinion mining and sentiment analysis", In 2008, Foundation and Trends in Information Retrieval, Volume-2.