

Gathering User Reviews for an Opinion Dictionary

Jun Kikuchi*, Vitaly Klyuev*

*Software Engineering Lab, University of Aizu, Tsuruga, Ikki-machi, Aizu-Wakamatsu City, Fukushima, 965-8580 Japan
s1190017@u-aizu.ac.jp, vklyuev@u-aizu.ac.jp

Abstract— Nowadays, people purchase a lot of products from online shopping sites. To support customers in decision making, some sites collect and provide user reviews on products. However, contents of the user reviews are too abundant for customers to analyze them in a short period of time. The automatic analysis of reviews is important to provide users with valuable information about goods of any category. The objective of this research is to improve the usefulness of reviews for consumers. This research focuses on an opinion dictionary as a collection of specific keywords and key phrases. This opinion dictionary models a standardized better review to extract patterns of trustworthy reviews. In this study, a simple corpus of three different categories of goods is composed. It consists of noun and adjective keywords. This research is successful to obtain essential features and relations among three different categories in the opinion dictionary. Moreover, this opinion dictionary will be applied to supervised learning methods, such as a support vector machine to create a review evaluation system. The findings from this study can contribute to assist users' decisions to evaluate reliable and useful reviews.

Keywords— Information Filtering, Pattern, Text Mining, Opinion Extraction, Opinion Corpus

I. INTRODUCTION

Nowadays, users may connect to the Internet for a lot of different purposes. According to the news [1], an online market worldwide is still expanding, and internet retail sales will reach nearly \$2.5 trillion by 2018. As we know from this statistic, purchasing from online shopping sites is a current trend nowadays. The difference between online shopping and shopping from an ordinary store is that people who use an online store cannot touch, see, and smell a product. At the online store, images, explanations, and feedback reviews are very much important for customers. It means that reviews from users may be crucial for consumers to make decisions about purchasing products.

To provide this helpful information, there are a lot of websites and web applications with functions to post and read user reviews. However, there is one problem with these reviews. This issue is that readers have to consider whether a review is trustful or useless. The latter does not have serious opinions. Some people describe well benefits or disadvantages of products in details. On the other hand, some other people mention meaningless information or unrelated information in their reviews. Even more, another important issue is about duplicated reviews and spam reviews. Malicious people post reviews to raise or lose product reputation deliberately.

Anyone can write and post a review, so it is necessary to improve the reliability of customer reviews.

As a solution of the aforementioned problem, the goal of this research is to create an opinion dictionary, which is a collection of specific keywords. This dictionary is a basis to analyze reliability of reviews. The dictionary is composed of words with part of speech tags. These words are from gathered reviews. This opinion dictionary collects features and patterns of useful reviews. It can be applied to a review evaluation system. A collection of keywords helps measure a value of reviews, it adapts to a support vector machine to analyze reviews automatically.

Opinion mining applied to the review analysis deals with types of reviews and various problems in this area. Reviews have many kinds of writing styles and a large vocabulary, so opinion mining concentrates on the automatic analysis of the reviews in the natural language.

II. RELATED WORK

Generally, opinion mining aims to determine and obtain entities, aspects, and opinion orientations. The entity is a target object such as a product that is evaluated by someone. The aspect mentions a component or a sub-component of the entity. The opinion orientation expresses positive, negative or neutral attitude in an opinion sentence. These three components are essential to analyze any kind of reviews. These basic components have been used to solve a lot of problems of the review analysis such as sentence subjectivity, sentiment classification, lexical expansion, etc. [2] [3].

There is a big stream of publications on opinion mining. Detailed surveys [4], [5] and [6] examine approaches widely used in different applications. Generally, approaches use a language dependent. However, trends in the area are common for western and Asian languages. In our short review of the latest achievements, we selected studies which influenced our research.

There is a research which studies significant elements for helpful online reviews [7]. Its goal is in mining customer requirements for considering as a useful review. In a mobile phone market, some statistical data show what kind of attributes is concentrated on by customers. The analysis of tendency of customer reviews is applied to investigate opinion dictionaries in detail.

To predict the helpfulness of online reviews automatically, a machine learning approach is used in some studies. The SVM, one of superior supervised machine learning algorithms, is applied to evaluate and predict helpfulness of reviews.

Datasets for the SVM are extracted features from reviews, reviewers, and metadata. The datasets included semantic features, structural features, and part of speech (POS) features such as nouns, verbs, adjectives, and adverbs. We refer to these elements of the dataset for analysis of customer reviews [8].

Another research studies a support vector regression (SVR) based on outlier detectors which designed by collecting reviews from the Internet [9]. The SVR detector predicts a score of hotels to improve reliability of hotel ranking. This research did morphological analysis to compile an opinion dictionary using POS patterns. This dictionary focuses on patterns with adverbs of degree and intransitive adverbs. Therefore, a feature value assigned to each opinion word in this dictionary. As a result, the SVR detector which trained using the opinion dictionary also has a feature of TF-IDF. These features help analyze reviews on the Internet and measure their reliability.

This research can contribute to keep a better quality of reliability for consumers and providers. Moreover, this research has another application to analyze product's evaluation from social networking sites (SNS). With keywords matching in an opinion dictionary, providers of products, such as companies or sellers, could gather many opinions and reviews from postings on SNS.

III. METHODS AND DATA COLLECTION

In this study, we analyzed several kinds of reviews on the Internet, and investigated critical and important information that a better review has. We aimed to make a basic opinion corpus. To obtain user reviews, we used Amazon.co.jp because this shopping site has two benefits for our study. One is that Amazon offers many well categorized reviews on each category. Another benefit is that Amazon has a feedback system toward user reviews. This feedback function help evaluate customer reviews by readers. We used this function to gather only better reviews. Every collected review must be voted more than 5 better feedback by readers because this kind of better review has more valuable and helpful contents for this investigation.

With Amazon APIs, this study randomly collected more than 450 customer reviews on products of three categories (Book, Electronic Devices, and Sporting Goods). A number of reviews, we collected for each category, was more than 150 to see tendencies and compile simple dictionaries as a beginning stage of this study. The reason to choose these three different categories is in extracting wide scooping customer reviews and in importance to compare simple sentences with complicated sentences and technical sentences with expressive sentences.

To analyze patterns of better reviews, we focused on review titles, contents, and review ratings. All reviews are only in Japanese. We also divided review contents into two types either an opinion category or a non-opinion category. We took opinion sentences from reviews with the following characteristics:

- containing targets and evaluation expressions;

- containing aspects and evaluation expressions;
- containing only evaluation expressions.

Targets and aspects mean that expressions are related to a product, for instance, a product, service, event, organization, or topic. An evaluation expression characterizes the customer's opinion, for instance, a positive or negative sentiment, attitude, emotion, etc. All reviews were classified into two types: positive or negative. We analyzed only opinion sentences with MeCab and TermExtract [10] in order to look at POSs features from collected reviews. Based on this POS tagging, we investigated reliable and useful reviews, especially we concentrated on noun and adjective words.

An opinion sentence contains many significant clues, for example, opinion entity, opinion, orientation of the opinion, opinion holder, and so on. We considered two significant items, an opinion entity and an opinion, as a core of our opinion dictionaries. An opinion entity, which is a target or aspect, is frequently expressed by noun words. On the other hand, an opinion, which is an evaluation expression, is expressed by adjective words. Therefore, these two POSs were selected for a simple dictionary to extract better review sentences. As a result, we organized six opinion dictionaries in total; three noun dictionaries and three adjective dictionaries. Opinion dictionaries of nouns have wide and varieties of words except for Sporting Goods. Nouns of Sporting Goods have simple and significant words for review such as related to a price, appearance, valuation and etc. On the other hand, opinion dictionaries for adjective are not various. These contain many common words and similar meaning of words.

IV. RESULTS

As we mentioned before, we automatically collected more than 450 reviews on three categories. Table 1 shows statistics on reviews in each category. All reviews were scanned to detect opinion sentences manually. Table 2 shows total numbers of sentences and total numbers of opinion sentences. All opinion sentences were morphologically analysed to obtain clues for the opinion dictionary. We extracted noun words and adjective words from POS tagging files. Figure 1 and Figure 2 show numbers of words extracted for each category and numbers of common words in the reviews on

TABLE 1. Gathered Reviews

Category	Book	Electronic Devices	Sporting Goods
Number of Reviews	160	167	169

TABLE 2. Number of Opinion Sentences in each Category

Category	Book	Electronic Devices	Sporting Goods
Number of All Sentences	1493	2238	839
Number of Opinion Sentences	771	1186	532

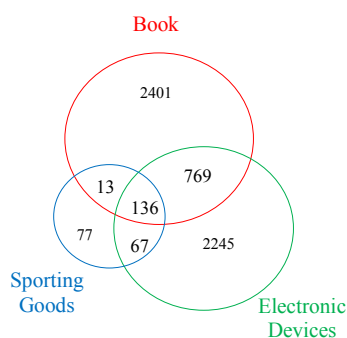


Figure 1. Number of Noun Words in each Category

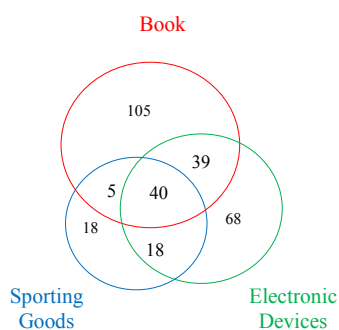


Figure 2. Number of Adjective Words in each Category

Table 3. Number of Opinion Sentences Detected using dictionaries

Category	Book	Electronic Devices	Sporting Goods
Number of All Opinion Sentence	771	1186	532
Books: Nouns	597	936	453
Electronics Devices: Nouns	591	968	468
Sporting Goods: Nouns	491	860	434
Books: Adjectives	504	736	380
Electronics Devices: Adjectives	478	777	394
Sporting Goods: Adjectives	215	443	284

Books, Electronic Devices, and Sporting Goods. These two figures also show statistics on words contained in each an introductory opinion dictionary. Using lists of nouns and adjectives, we tried to detect opinion sentences from all reviews. The purpose of this detection is to check that these opinion corpora extract features of opinion sentences correctly, and we measured how these compiled lists are useful. Table 3 shows a number of detected opinion sentences for each list of nouns and adjectives. Each introductory dictionary can find certain numbers of opinion sentences, but it is still challenging to distinguish a non-opinion sentence from an opinion sentence precisely. Table 4 illustrates how accurate this detection is. These numbers mean how accurate this detection

Table 4. Accuracy of Detection for the Opinion Dictionary

Category	Book	Electronic Devices	Sporting Goods
Books: Nouns	40.3%	43.2%	57.1%
Electronics Devices: Nouns	40.4%	43.6%	57.6%
Sporting Goods: Nouns	42.3%	45.1%	58.2%
Books: Adjectives	44.4%	49.6%	62.7%
Electronics Devices: Adjectives	44.4%	50.7%	63.5%
Sporting Goods: Adjectives	48.6%	61.9%	66.5%

find opinion sentences from all mixed sentences. This accuracy is calculated as follows: a number of correct opinion sentences is divided by the total number of detected sentences. The preliminary examination shows quite a number of false sentences instead of opinion sentences.

V. DISCUSSION

We focused on two POSs, nouns and adjectives because these two types of words have significant characters especially in opinion sentences. According to Figure 1 and Figure 2, we could point out two large word lists in the Books category and Electronic Devices category. In Books category, opinion dictionaries contained complex and difficult words. In contrast, noun and adjective words of Electronic Devices are composed of terms or technical words. Each category has its own unique word set to express opinions. On the other hand, more than 70% of nouns and adjectives have common words in the Sporting Goods category. It means that noun and adjective lists in Sporting Goods are possible to contain more basic words for all categories. To compile a core corpus, it is necessary to deeply investigate Sporting Goods word lists.

Another interesting point in this study is about extraction of opinion sentences for each simple corpus. As can be seen, results of extraction of nouns and adjectives for Books are very similar to results of Electronic Devices. However, Books and Electronic Devices have many different words in the noun corpus and adjective corpus. In addition, the number of extracted sentences for the noun corpus for Sporting Goods results in nearly Books and Electronics Devices. This fact shows that these corpora did not work well. To improve the accuracy of detection, noun corpora of Books and Electronic Devices are needed to be inspected to remove insignificant keywords.

Adjective corpora for each category can detect opinion sentences with the higher accuracy than noun corpora, but we pointed out detected numbers of correct opinion sentences. All three adjective corpora did not detect opinion sentences accurately.

VI. CONCLUSION AND FUTURE WORK

In this research, we compiled simple opinion corpora based on the gathered collections. We found important differences

between Sporting Goods and two other categories. According to extraction results, nouns and adjectives from the corpus on Sporting Goods contain basic keywords for opinion sentences for all three categories. However, two other noun corpora are quite big and too wide, so we need to polish word selection to improve accuracies.

This research also showed that adjectives are important for an opinion corpus according to detection results. To increase the number of correct by detected opinion sentences, it is necessary to inspect each adjective corpus.

This study showed good results for compiling an opinion corpus. We investigated different characteristics and relations between three categories. To achieve better opinion corpora in our future work, we would like to analyze contents of each corpus deeper, for example, selection critical words or combination nouns with adjectives.

REFERENCES

- [1] Retails Sales Worldwide Will Top \$22Trillion This Year. eMarketer. [Online]. Dec. 2014. Available: <http://www.emarketer.com/Article/Retail-Sales-Worldwide-Will-Top-22-Trillion-This-Year/1011765>
- [2] Liu, B. and Zhang, L. A survey of opinion mining and sentiment analysis. *Mining Text Data*, 415-463. 2012.
- [3] Hsieh, H. Klyuev, V. Zhao, Q. and Wu, S. SVR-based outlier detection and its application to hotel ranking. In *Proc. of the 2014 IEEE 6th International Conference on Awareness Science and Technology (iCAST)*, 1-6. 2014.
- [4] Zohreh Madhoushi, Abdul Razak Hamdan, Suhaila Zainudin. Sentiment analysis techniques in recent works. 2015 *Science and Information Conference (SAI)*, 288-291. 2015.
- [5] Erik Cambria, Bjorn Schuller, Yunqing Xia, Catherine Havasi. *New Avenues in Opinion Mining and Sentiment Analysis*. IEEE Intelligent Systems 2013 vol. 28. 13-21. 2013.
- [6] Basant Agarwal, Namita Mittal. Semantic Feature Clustering for Sentiment Analysis of English Reviews. *IETE Journal of Research* 60, 414-422. Nov. 2014.
- [7] Zhang, Z. Qi, J. and Zhu, G. Mining Customer Requirement from Helpful Online Reviews. *Enterprise Systems Conference (ES)*, 2014. 249-254. Aug. 2014.
- [8] Zhang, Y. and Zhang D. Automatically predicting the helpfulness of online reviews. *Information Reuse and Integration (IRI)*, 2014 IEEE 15th International Conference, 662-668. Aug. 2014.
- [9] Pang, B. and Lee, L. Opinion mining and sentiment analysis. *Now the essence of knowledge*. Vol.2, 1-135. 2008.
- [10] Miyashita, M. and Klyuev, V. TermExtract: Accuracy of Compound Noun Detection in Japanese. *Future Information Technology*, 189-194. 2014.



Jun Kikuchi is a bachelor student in the Department of Computer Science and Engineering at the University of Aizu, Japan. His research areas are text mining and opinion mining.



Vitaly Klyuev is a professor at the University of Aizu, Japan. He received a Ph.D. degree in Physics and Mathematics from St. Petersburg State University, Russia in 1983. His research domain includes information retrieval, software engineering and analysis of computer algorithms. He has more than 100 publications in referred journals and conference proceedings, three co-authored and eight co-edited books. Dr. Klyuev is a member of editorial board of several academic journals and a program committee member of more than 20 conferences sponsored by ACM, FTRA, IEEE, ISCA and IARIA.