Complaint Data Analysis for Recommender System and its Applications

journal or publication title	KGPS review : Kwansei Gakuin policy studies review
number	27
page range	79-86
year	2020-03-31
URL	http://hdl.handle.net/10236/00028649

Complaint Data Analysis for Recommender System and its Applications

Liang Yang

Abstract:

In recent years, recommender systems have become an important tool for helping users to find relevant contents. A majority of the current recommender systems focus on users' positive reviews and analyze users' interests and preferences in order to make suitable recommendations and suggestions to them. However, there are few recommender systems focusing on users' complaint reviews and dissatisfaction. In this paper, in order to address users' dissatisfaction, we present an approach to recommend various information based on complaint reviews analysis and its applications. We analyze the Complaint Dataset provided by Fuman Kaitori Center and focus on the information recommendation of advice and advertising slogan word. We describe the detail of this recommender system that implements the application part providing a query keyword generation method supporting for collecting web pages containing advice; and a word dissatisfaction degree evaluation method supporting for advertising slogan production. We perform questionnaire experiments to demonstrate the effectiveness and usability of the proposed system.

Key word and phrases: Complaint Data, Advice, Advertising Slogan, Recommender System

1. Introduction

In this paper, we use and analyze the Complaint Dataset provided by Fuman Kaitori Center and focus on the dissatisfaction of users. For Fuman kaitori center, they analyze users' complaint reviews and provide hints for products and services improvement to a lot of companies and communities. However, it takes quite a long time until these service or products to be improved. Also, users' dissatisfaction cannot be addressed during this long time. Therefore, in this paper, we analyze complaint reviews and make various information recommendation to users directly in order to address their dissatisfaction in real time. For example, many users post negative reviews about a certain service online. However, it is quite difficult to respond to various user demands for service in real time as the service is provided by the company. In addition, there are almost no direct solutions when users feel dissatisfied with a certain service. In addition, nowadays, advertising slogans are always used to impress consumers and arouse their interests for purchasing. However, some advertising slogans cannot appeal products' strengths and advantages by using inappropriate words. As a result, consumers may be affected by these advertising slogans and they may lose interests in these products. Then it may lead to decrease of sales and leave bad

brand impression to customers. Also, with the growing share of digital advertisements, it is considered as a huge burden for copywriters to create various advertising slogans based on users' different needs and make everyone feels satisfied with.

On the basis of these backgrounds, in this paper, in order to address users' dissatisfaction, we aim to recommend various information based on complaint reviews analysis and its applications. There are 2 applications in this paper: Advice recommendation for users; Advertising slogan word recommendation for copywriters.

2. Overview

2.1 Complaint Dataset

In this paper, we use and analyze a dataset of complaint reviews from the Fuman Kaitori Center, which is provided by Insight Tech Inc. from the National Institute of Informatics. In this paper, we refer to the Fuman Kaitori Center's dataset as the Complaint Dataset. The Fuman Kaitori Center is a website on which users can post their complaints about topics such as products, services, education, work, relationships and so on. Moreover, users get points when they post complaints that they can exchange for coupons for online shopping website. For Fuman Kaitori Center, they can analyze users' complaints and extract their needs as useful development plan for products and services in the future. This dataset contains about 5 million complaint reviews that were posted from 18 March 2015 to 12 March 2017 by around 100,000 users. In addition, there are two characteristics of Complaint Dataset. First is that all of the reviews are negative contents. The other one is that various ways of complaints about one product or service could be found in this dataset.

2.2 Our Approach

In this paper, in order to address users' dissatisfaction, we discuss how we analyze and apply Complaint Dataset for various information recommendation. The overview of our approach is described as Figure 1 shows. Our proposed system can be utilized as 2 applications:

- Advice recommendation for users.
- Advertising slogan word recommendation for copywriters.

For the first application, we propose a recommender system by analyzing complaint data to offer suitable advice to users who post complaints. We extract query keywords from various user complaints about a certain service by calculating the score of each query. Then suitable web pages containing advice are recommended from the results of the query.

For the second application, we propose a recommender system by analyzing complaint data to suggest proper advertising slogan word to support copywriters' work. We calculate the dissatisfaction degree of feature words about a certain product by classifying complaint reviews. Then proper advertising

slogan words are recommended based on the result of dissatisfaction degree.

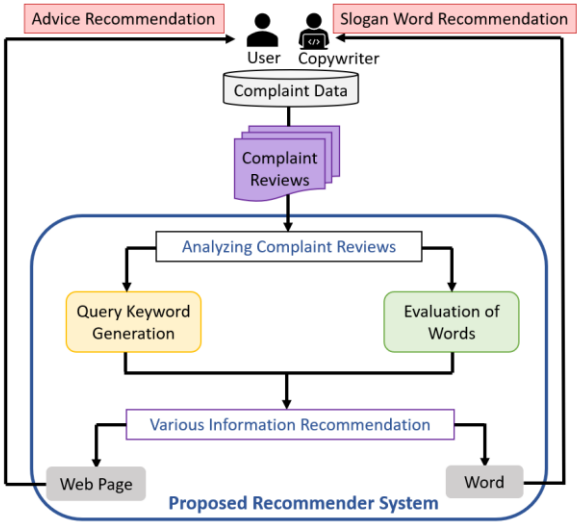

Figure 1 Our approach

3. An Advice Recommender System Based on Complaint Data Analysis

3.1 System Overview

In this work, we propose a method by analyzing complaint data from Fuman Kaitori Center for recommendation of advice in order to address users' dissatisfaction about a ertain service. Figure 2 shows the system flow of our proposed method.

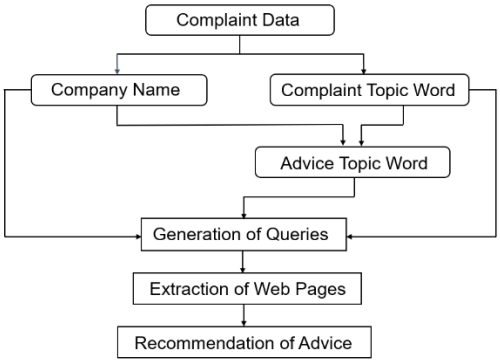

Figure 2 System Flow

To generate query keywords, we first analyze complaint data from the Fuman Kaitori Center and extract company names from the data. Second, from the negative reviews for each company, we extract complaint topic words that represent users' dissatisfaction of each company. Third, we extract advice

topic words that can help users to collect advice according to their various complaints about a certain service. Then, we generate the query based on the company name, complaint topic word, and advice topic word. Finally, suitable web pages containing advice that could address a user's complaint are recommended from the results of the query. This advice could address users' dissatisfaction and respond to their different demands in a comprehensive way. An example of advice recommendation is described in Figure 3.

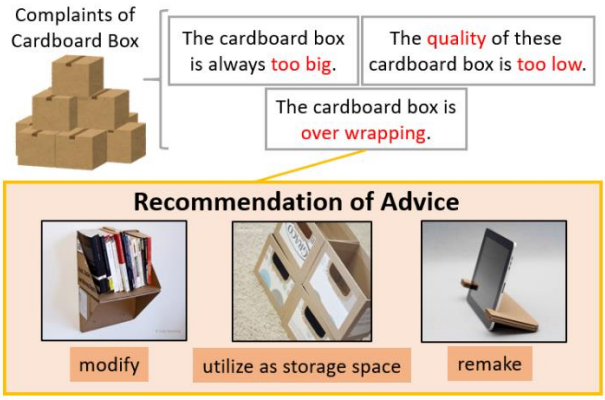

Figure 3 Example of Advice Recommendation

3.2 Experiment and Evaluation

In this study, we conducted an experiment to extract the complaint and advice topic words in order to verify the feasibility of proposed system. For this experiment, we analyzed the subcategory of ``IT web services" of the Complaint Dataset, which is under the category ``industry." We analyzed 1,000 complaint reviews for each of 3 companies.

In this paper, we conducted a questionnaire-based survey to evaluate the usability and effectiveness of the proposed method. The purposes of the questionnaire are to evaluate if the complaint topic words could express the dissatisfaction in the complaint reviews; if by searching with the query keywords could be helpful to collect advice; if are satisfied with the result of advice recommendation. The result showed by using those nouns are with the highest value of importance in that complaint review for one time only, 60% of appropriate complaint topic word can be extracted from the complaint review. And if we search advice by using the query which is with the lowest score for one time only, 70% of appropriate queries can be offered to make web search in order to address the complaints expresses in the complaint reviews. Also, 75% of the answer felt satisfied with the contents of advice they searched with the queries they’ve chosen. The result demonstrated that scoring candidate keyword is effective. In addition, comparing to the traditional search engine, users have to choose from 10 queries, our proposed method offered users only 1 query to choose from. It implied that by using the proposed system can help users to release their burden when searching for advice comparing to traditional search engine. However, we found out that the longer

the candidate keyword was, the lower the score will be when making the candidate keywords. In the future, we plan to develop a method to ensure if the candidate keyword is related to the complaint topic word to better exclude noise in the results.

4. An Advertising Slogan Recommender System Based on Complaint Data Analysis

4.1 System overview

In this study, we propose a system to recommend advertising slogan word on the basis of Complaint Dataset analysis by evaluating words' dissatisfaction degree in order to support advertising slogan production. Figure 4 presents the system flow of our proposed method. First, we extract feature words of product by collecting explanatory texts from its official websites or positive reviews online. Then we calculate the importance of the words. Second, we extract complaint reviews that contain feature words and product name. Next, we analyze these complaint reviews and classify them based on the rules we made. Third, we calculate the dissatisfaction degree of feature word and rank them. Finally, appropriate words that can be used as advertising slogans to address users' dissatisfaction are recommended from the results of the ranking. An example of advertising slogan word recommendation is described in Figure 5.

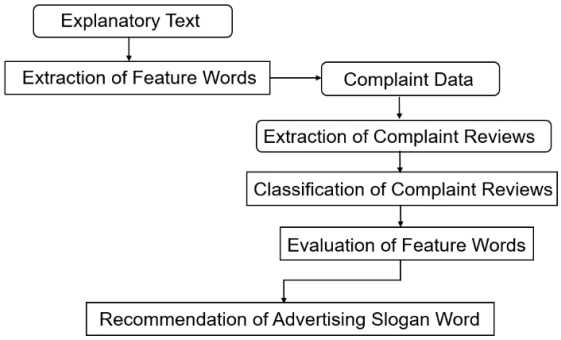

Figure 4 System Flow

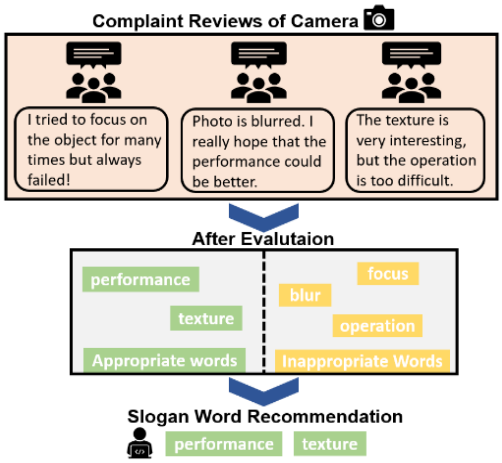

Figure 5 Example of Advertising Slogan word Recommendation

4.2 Experiment and Evaluation

In this study, we conducted an experiment to extract products' feature words and calculate the dissatisfaction degree of feature words in order to verify the feasibility of proposed method. For this experiment, we analyzed 10,865 complaint reviews of the product ``camera'' from the Complaint Dataset.

In this paper, we conducted 2 questionnaire-based surveys to evaluate the effectiveness of the proposed system. The purpose of the questionnaire survey is to evaluate if the feature words with lower dissatisfaction degree are really less likely to be dissatisfied with for users; if the feature words with higher dissatisfaction degree are really more likely to be dissatisfied with for users. We collected 16 advertising slogans of camera and combined them into groups. Each advertising slogan contains one feature word we extracted from the experiment to be evaluated by respondents. Each group had 2 advertising slogans to compare with. For questionnaire 1, we asked respondents to compare them based on the view of their motivation of purchase after reading these advertising slogans. For questionnaire 2, we asked respondents to compare them based on the view of their regrets of purchasing after reading these advertising slogans. For Questionnaire 1, the result showed that 62% of answers chose the advertising slogan with lower dissatisfaction degree. It explained that 62% of answers tend to be more satisfied and willing to purchase after reading those advertising slogans containing words with lower dissatisfaction degree. For Questionnaire 2, the result showed that 72% of answers chose the advertising slogan with higher dissatisfaction degree. It explained that 72% of answers tend to be more likely to be regretted after purchasing based on the impression of those advertising slogans containing words with higher dissatisfaction degree. In addition, we evaluated the attractive degree and regret degree by calculating the selected ratio of each advertising slogan. For the evaluation of attractive degree, it shows that words with lower dissatisfaction degree tend to be more attractive to users; Word with the highest dissatisfaction degree is least attractive to users. For the evaluation of regret degree, it shows that words with higher dissatisfaction degree tend to be more easily to be regretted with to users; Words with the lowest dissatisfaction degree are the most hardly to be regretted with to users. Also, there are several words like “view finder”, “photographic subject”, “shutter”, “model”, they have almost the same dissatisfaction degree and regret degree.

From the result, it is observed that by using the proposed system could rank words to be used as advertising slogan effectively. When producing camera's advertising slogan, word like “performance” should be appealed more often. Word like “camera shake” should be less used because it may be easily dissatisfied with. Moreover, it implied by using the proposed system can help advertising slogan producers to release their burden when making a choice of words during production by recommending the word with lower dissatisfaction degree.

5. Conclusion

In this thesis, in order to address users' dissatisfaction, we proposed a recommender system for

suggesting various information based on complaint review analysis and its applications. We analyzed the Complaint Dataset and focused on the information recommendation of advice and advertising slogan word. We described the detail of this recommender system that implements the application part providing a query keyword generation method supporting for collecting web pages containing advice; and a word dissatisfaction degree evaluation method supporting for advertising slogan production.

To achieve our proposes, for the advice recommendation application, first we analyzed the Complaint Dataset and extracted particular company names from the data. Second, we extracted complaint topic word that can express the dissatisfaction target of users from the complaint reviews for each company by calculating the importance of each word. Third, we extracted advice topic word that can help users to collect advice to address their complaints by calculating the relevancy of the candidate keywords to the Complaint Dataset and score them using morphological and polarity analyses. Next, we generated a query based on the company name, complaint topic word, and advice topic word on the basis of the score of each query. Finally, suitable web pages containing advice are recommended from the results of the query. In addition, we evaluated the effectiveness and usability of the proposed system through a questionnaire survey, and the results showed that the generated query keywords would be useful for collecting advice. Also, the recommendation of advice returned by query keywords could address users' dissatisfaction with a service and respond to different user demands in a comprehensive way. In the future, we plan to evaluate the satisfaction degree of each query and analyze the result. Furthermore, we would like to consider new methods to obtain candidate query keywords which users are hard to associate to enhance the usability of the proposed system.

For the advertising slogan word recommendation application, first we extracted feature words from particular product's explanatory texts by using the TF-WebIDF method. Second, we extracted complaint reviews that contain feature words and product name. Then, we analyzed these complaint reviews and classify them based on the rules we made. Third, we evaluated the dissatisfaction degree of feature word on the basis of complaint review classification result and rank them. Finally, proper words that suggest could be used as advertising slogans are recommended from the results of the ranking. In addition, we evaluated the effectiveness of the proposed method through 2 questionnaire surveys, and the results shows that by using the proposed system could rank words to be used as advertising slogan effectively and support production. Moreover, it implied by using the proposed system can help copywriters to release their burden when making a choice of words during production by recommending the word with lower dissatisfaction degree. In our future work, we plan to analyze other positive reviews to recommend words to be used as advertising slogans that are easily to be satisfied with. Furthermore, we will consider a method to extract compound words to be used as advertising slogans such as “interesting texture” from Complaint Dataset.

**References ***Cited references in this paper. Other references are listed in the master thesis.

[1] Fuman kaitori Center: https://fumankaitori.com/

[2] Mitsumasa Kondo, Tetsushi Morita, Akimichi Tanaka, and Tadasu Uchiyama.“Personalized Query Recommendation Using HITS-Based Wikipedia Ranking Algorithm and User History,”The 19th Data Engineering Workshop(DEWS2008), B2-4. Mar. 2008.