Reputation Analysis Using Key Phrases and Sentiment Scores Extracted from Reviews

Huang Yipu
Graduate School of Science and Engineering, Ibaraki University

Minoru Sasaki
Computer and Information Sciences, Faculty of Engineering, Ibaraki University

Kanako Komiya
Graduate School of Engineering, Tokyo University of Agriculture and Technology

{21nm723s, minoru.sasaki.01}@vc.ibaraki.ac.jp, kkomiya@go.tuat.ac.jp

Abstract

In recent years, reviews of various products by purchasers have been posted on websites. By browsing these reviews, company representatives intend to collect and analyze opinions from the users' side of the products and use them to improve the products. However, existing analysis methods analyze reviews by dividing them into words, and do not analyze the function names that are phrases. Therefore, we propose a reputation analysis method using a decision tree model based on the frequency of key phrases extracted from reviews and sentiment scores. Experimental results showed that the proposed method using key phrases improved the accuracy by 3 points compared to the existing method that analyzes by words.

1 Introduction

When consumers purchase a product or service, they invariably consult reviews. In fact, review analysis has many benefits not only for individuals but also for companies.

For example, the use of the analysis for product planning, analysis of user needs and dissatisfaction, hints for function and service improvement, and verification of the effectiveness of promotions and marketing measures are the effects of corporate review analysis.

In order to improve a product, it is necessary to know what reviewers like about the product and how they evaluate it. Therefore, we use two criteria, the frequency of occurrence of key phrases and the sentiment score, to determine the reviewer's opinion of the product. For example, if the frequency of occurrence of a key phrase is high and the sentiment score is also high, we judge that "everyone is interested in the product, and it has a good reputation" and that the product does not need to be improved immediately.

This time, key phrases (important words) are extracted from a certain review, and a dictionary is created with key phrases as keys and the frequency of occurrence of key phrases as values. Then, we calculate the sentiment score from the review sentences, treat the sentiment score and frequency of key phrases as explanatory variables, and treat whether improvement is necessary or not or priority as objective variables, and create a decision tree model.

The following sections describe key phrase extraction, sentiment score calculation, decision tree model creation, and quality evaluation.

2 Related Research

There have been several studies on review analysis methods. Kobayashi collects and analyzes reviews from web pages, they use TF-IDF to extract keywords, and then extract emotion words, and combine keywords and emotion words (Kobayashi, 2008). However, it is difficult to grasp the whole picture because of the variety of emotion words extracted from various reviews. Abe analyzes product reviews, assigns a score to each evaluation item, and proposes to create an item-by-item dictionary of evaluation expressions related to hotel evaluations, such as keywords, features, and degree (Abe, 2020).

3 Key phrase extraction

3.1 Key Phrase

Key phrase extraction is a technique for extracting phrases that best describe a document. A phrase here is a set of words whose meanings are combined, but in practice, noun phrases are often employed. Many methods already exist, such as PageRank, Text Rank, Single Rank, and Topic Rank (Page, 1999) (Sullivan, 2007).

3.2 Extraction Process

In this case, we will use PKE (Boudin, 2016) to extract key phrases. PKE is produced by Florian Boudin, the author of the MultipartieRank paper.

text	key phrase
バッテリーの方が大きい笑	['バッテリー']
軽量化の為仕方がないのだろう	['軽量化']

Table 1: Examples of Key phrase extraction

Key Phrase Frequency

It is important to calculate the frequency of each key phrase in order to know the customer's level of interest in the product from the reviews (Berger and Mittal, 2000) (Juan-Manuel, 2014). A dictionary is created by adding value values based on the number of times a key phrase appears as shown in Figure 1.

{'バッテリー': 404, '軽量化': 134, '充電器': 501, …}

Figure 1: Example of Dictionary

Next is the key phrase frequency calculation, the definition of the key phrase frequency is as follows.

$$z = \frac{x - \mu}{\sigma},$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x - \mu)^2},$$

where x represents the frequency of key phrase and μ represents the Number of key phrases.

The result of key phrase extraction「トルクは申し分なく握った時の重量配分も問題ない」→「トルク,申し分,重量配分,問題」→「1.7452, -0.923112, 0.4218, 2.9283233」. From this, each review was indicated on the frequency of the key phrase.

4 Emotion Score Calculation

The customer's evaluation is judged to be good or bad based on the review, and the sentiment analysis is used to determine whether the customer's evaluation is good or bad.

4.1 Model

Here we used bert-base-japanese-sentiment from huggingface to evaluate sentiment score as hown in Figure 2.

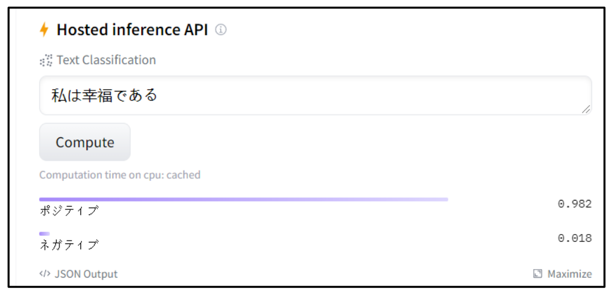

Figure 2: Example of Sentiment Evaluation

4.2 Procedure

In order to effectively perform sentiment analysis here, the sentiment score needs to float between [-1, 1] as shown in Figure 3. In this experiments, Bert-base-Japanese-sentiment model[1] is used to slightly improve the results.

For Positive：

$$score = 2 \times score - 1$$

For Negative：

$$score = 1 - 2 \times score$$

Figure 3: Emotion Scores

4.3 Assignment of Sentiment Scores

One by one, the review sentences are assigned an emotion score. The following are some of the results.

[1] https://huggingface.co/daigo/bert-base-japanese-sentiment (Currently not available)

Text1: そして電池の持ちがいい
Score: 0.9681352376937866
text2: 本体部分が小さいのはびっくりで
score: 0.771661639213562
text3: 多少重いがやむを得ないか
score: -0.9029324054718018
…

5 Decision tree model

5.1 Summary

Decision tree analysis is a data mining method used for “prediction”, “discrimination”, and “classification”. It is an analysis method that finds “explanatory variables” that affect the “dependent variable” of customer information, survey results, etc., and creates a tree-like model.

5.2 Explanatory Variables

The explanatory variables are the objects to be analyzed. Here, the level of interest (frequency of key phrases) and the sentiment score value are done as explanatory variables.

5.3 Objective Variable

The objective variable is the one that displays the results brightly, influenced by the various explanatory variables. Here, the objective variables are set as A, B, C, and D four types as shown in Table 2. The details that each item represents are shown below.

A: Attention and good evaluation

B: Noteworthy and poorly rated

C: Not attracting attention and rated good

D: Not noticed and evaluated poorly

Then, the priority order of improvement for the firms is B D C A. The order of priority for the firms is B D C A.

level of interest	emotion	Objective variable
High	positive	A
	negative	B
low	positive	C
	negative	D

Table 2: Definition of Objective Variables

5.4 Decision Tree Model Construction

If we build a decision tree model, we need explanatory variables and an objective variable. The explanatory variables are the frequency of the key phrases and the sentiment score, previously computed and attached. The objective variable is also an explicitly invisible element. The dataset is divided in a 3:7 ratio between training and test data. The objective variables are then added manually to the training data.

explanatory variable		Objective variable
level of interest	Emotion Score	type
2.923331	0.822314	A
3.22134	-0.92314	B
-0.83167	0.22314	C
1.99913	0.883424	A
…	…	…

Table 3: Decision Tree Model

Now we will analyze the analysis with the CART decision tree (Quinlan, 1987) described in Figure .

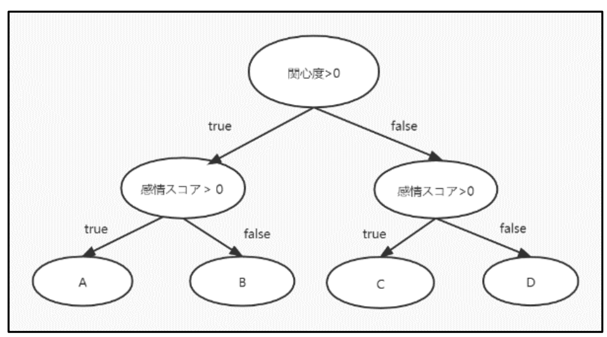

Figure 4 Decision Tree Conceptual Diagram

6 Quality Assessment

6.1 Method

To verify that the model works well, 559 product reviews are used to evaluate the accuracy of the model, which varies with the frequency of key phrases and sentiment scores.

6.2 Results

Using the new 559 reviews, the same key phrase frequency and sentiment scores are transformed, and this model is put in. The results are shown in Table 4. Based on these results, about 83% of the reviews are correct.

	quantity
Results match	468
Difference in results	91

Table 4: Experimental Results

6.3 Analysis and Comparison

The results show that this decision tree model method is effective. Compare the current proposal with the review-analysis methods of Kobayashi et al. and Abe et al. as shown in Table 5.

Our proposed method is a little more accurate than that of Kobayashi et al. because I can grasp the whole picture. Abe et al. used their own evaluation score, but the authenticity of the original evaluation score is still a matter to be examined.

And with our method, we can use the model to clearly show the reviewers and the company's priorities for improvement of the next product. However, there are some areas that need to be improved.

Key word extraction	Emotional Analysis	Model	Accuracy
PKE, Key Phrases	Sentiment score for sentence	Decision Tree Modeling	**83%**
TF-IDF, Keywords	Emotional word extraction	Keyword and Emotional Word	79.9%
Keywords	Review own evaluation score	Target expression dictionary	80.6%

Table 5: Comparison of methods

7 Conclusions

Analysis of reviews is important for companies to improve their products. We propose a CART decision tree with key phrase frequency and sentiment score as explanatory variables to determine the interest and reputation of a product. The four objective variables are used to determine a company's next steps in improving its products. Future work includes improving the key phrase extraction method to increase the accuracy of review analysis and improving the simultaneous decision tree model.

References

Daisuke Kobayashi, Syunichi Nagao, Daisuke Suto, Ushio Inoue. 2008. Proposal of a collection and analysis system of user reviews by text analysis of web pages, DEWS2008 B8-2.

Katsumi Abe and Shinsuke Nakajima. 2020. Hotel recommendation system based on automatic review scoring scheme, DEIM2020 P1-28.

Lawrence Page, Sergey Brin, Rajeev Motwani and Terry Winograd. 1999. The PageRank Citation Ranking: Bringing Order to the Web, The Web Conference.

Danny Sullivan. 2007. What Is Google PageRank? A Guide For Searchers & Webmasters, Search Engine Land.

Juan-Manuel Torres-Moreno. 2014. Automatic Text Summarization. John Wiley & Sons.

Adam L. Berger and Vibhu Mittal. 2000. OCELOT: a system for summarizing Web pages, Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.

Taher H. Haveliwala. 2003. Topic-Sensitive PageRank: A Context-Sensitive Ranking Algorithm for Web Search, IEEE Transactions on Knowledge and Data Engineering, 15(4) pp.784–796.

Florian Boudin. 2016. pke: an open source python-based keyphrase extraction toolkit, Proceedings of the 26th International Conference on Computational Linguistics (COLING2016): System Demonstrations, pp.69-73.

John R. Quinlan. 1987. Induction of decision trees. Machine Learning, 1(1) pp.81-106.