J. Chosun Natural Sci.
Vol. 14, No. 2 (2021) pp. 35 – 40
https://doi.org/10.13160/ricns.2021.14.2.35

Modeling Topic Extraction-based Sentiment Analysis Based on User Reviews

Tae-Yeun Kim†
National Program of Excellence in Software center, Chosun University, Gwangju, Korea

Abstract

In this paper, we proposed a multi-subject-level sentiment analysis model for user reviews using the Latent Dirichlet Allocation (LDA) method targeting user-generated content (UGC). Data were collected from users' online reviews of hotels in major tourist cities in the world, and 30 hotel-related topics were extracted using the entire user reviews through the LDA technique. Six major hotel-related themes (Cleanliness, Location, Rooms, Service, Sleep Quality, and Value) were selected from the extracted themes, and emotions were evaluated for sentences corresponding to six themes in each user review in the proposed sentiment analysis model. Sentiment was analyzed using a dictionary. In addition, the performance of the proposed sentiment analysis model was evaluated by comparing the emotional values for each subject in the user reviews and the detailed scores evaluated by the user directly for each hotel attribute. As a result of analyzing the values of accuracy and recall of the proposed sentiment analysis model, it was analyzed that the efficiency was high.

Keywords : LDI, user-generated content

(Received June 17, 2021; Revised June 17, 2021; Accepted June 24, 2021)

1. Introduction

Along with the rapid growth of social networks, Internet users are able to search and collect travel and tourism information alone. As a result, numerous tourism information about tourist attractions, hotels, and restaurants are created in the form of online user-generated contents (UCG)[1,2]. The user-generated contents found in websites of various social media include consumer experience, user feedback, and product review. Also, searching the travel-related information is one of the most famous online activities[3].

In response, natural language processing and data mining methods have been hugely spotlighted for developing the analysis method which precisely extracts the opinions of people from large amount of user-generated contents[4]. The purpose of sentiment analysis and opinion mining is to explore the subjective information such as opinion, attitude, and sentiment expressed in the user-generated contents automatically.

Topic-based sentiment analysis is one of the basic frameworks of sentiment analysis[5]. There have been numerous researches on topic-based sentiment analysis. The sentiment analysis was conducted by selecting a topic by extracting a noun from the sentence of document, analyzing the sentiment with a sentiment lexicon based on the topic, and defining topic concept in the level of aspect for the topic extraction[6]. For selecting the extracted aspect, the aspect was selected with experts[7]. In addition, there were numbers of researches that suggested new research methods by combining topic models and sentiment analysis[8].

The topic-based sentiment analysis researches through topic modeling mainly analyze the sentiment by creating multiple topics for all reviews posted by users. In other words, it analyzes the overall sentiment toward the

†Corresponding author : tykim@chosun.ac.kr

topic extract by combining the reviews of all users. However, as there might be different points of sentiment in one review, it is necessary to analyze the sentiment for each topic of the reviews. In response, this study seeks to develop a model for analyzing the sentiment of each topic of each customer review. The suggested sentiment analysis model is capable of inter-comparing various sentiments of each user for multiple topics.

In the sentiment analysis model, the data of user review on hotel is collected for sentiment analysis for each topic. First, LDA model is applied on the online review to extract the topic that users are most interested in. The extracted topics show the types of topics that hotel guests are interested in. Here, topics are used to analyze the sentiment of the users in each topic of the user reviews. To evaluate the performance of the suggested sentiment analysis model, this study compared the hotel user's evaluation on each attribute of hotel and the results of sentiment analysis conducted by using the 6 attributes in the user reviews. In this model, hotel administrators are capable of reviewing sentiment values for various topics in user reviews and identifying the results of sentiment analysis for each detailed attribute of the hotel.

2. Related Research

2.1. Topic model

Recently, big data is showing a rapid growth. Compared to numerical data, text data produced through the social media used by people in real-time is getting much more important.

In regard to researches about text-based documents, there have been numerous researches on various analysis methods including topic model method. Topic model is a method of exploring different topics holding significance in collection of documents and the model is used for extracting topics. Each topic is composed of similar terms holding significance. The purpose of topic model is to explore the significant topics in mass data. According to Blei and Lafferty, topic model-based information system is capable of interpreting documents automatically and extracting useful structures[9].

In addition, there have been numerous researches on the topic model for finding the topics of text data[10]. Blei et al. applied LDA model to extract topics from documents[11]. LDA model which is often used for topic extraction assumes that extracted topics have different shares in each document. LDA model regards topic-mixed parameter as the variable taken from Dirichlet distribution to expand pLSA model[12].

2.2. LDA-based sentiment analysis

There have been active topic model researches based on LDA algorithm. Lin and He developed a new LDA-based probabilistic modeling framework[13]. This research model is called a joint sentiment/topic (JST) model and it is capable of analyzing and carrying out sentiment classification and topic extraction at the same time. Xianghua et al. applied LDA, found out multiple global topics in social reviews, and demonstrated that local topic and sentiment are extractable based on the sliding window context[4]. Marrese-Taylor et al. upgraded the method suggested by Liu and applied on the tourism field. This study developed natural language processing (NLP) rules related to sentiment classification and more complicated aspect-level subjectivity and verified the effectiveness of the suggested method[7,14].

3. Sentiment Analysis Modeling

Sentiment analysis modeling has a research framework as indicated in Figure 1. In the stage 1, data is collected by crawling user-created contents mainly from websites. This study used a python program to collect data from online reviews posted on a travel platform (Tripadvisor). After collecting the data, cases with missing value were deleted. The user reviews were posted in English and all reviews were processed by eliminating stop words and converting all characters into small letters.

In step 2, the optimum number of topic had to be set before classifying the topics. The review data collected depending on the number of topic was applied with LDA model to extract important topics. Based on the general meaning of the extracted keywords, the topics extracted by referring to documents and interpreting the classified topics are matched with 6 attributes of the hotel. Then, the final keyword is determined by referring to the document based on the keyword for the matching topic. Like this, the sentence for each topic is obtained by matching the sentences that include keyword for each topic.

Lastly, the sentiment value for each topic was figured

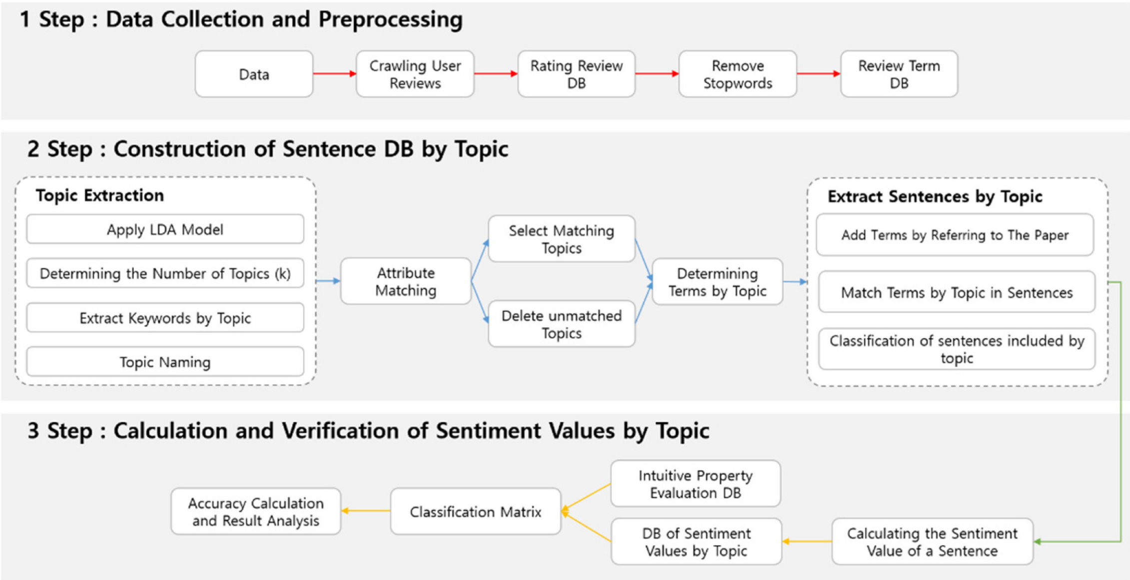

Fig. 1. System configuration.

out through sentiment analysis based on the sentiment lexicon for sentences of each topic. The performance of the suggested model was evaluated by comparing the sentiment values for each topic of the reviews calculated by using the suggested sentiment analysis model and customer evaluation value of 6 attributes in online reviews. Also, classification matrix was used to evaluate whether the extracted sentiment value for each topic is significant or not. In addition, each measurement result demonstrates the meaning of the suggested model.

4. Experiments and Results Analysis

4.1. Data collection

This study targeted on Tripadvisor platform and collected online review data on hotels located in 7 major cities (London, Paris, Bangkok, Beijing, Hong Kong, Seoul, Singapore). The collected data includes the total evaluation grade, detailed evaluation grade for each attribute, and evaluation review for hotels that users experienced from January 2019 to December 2019. Total of 34,679 hotel data cases were collected while 11,700 of them were effective. Table 1 is the summary of the collected data. In online reviews, the evaluation grade of 4 and 5 were regarded as positive sentiment while evaluation grade of 1 and 2 were regarded as negative sentiment.

4.2. Experiment design

The experiment of the suggested sentiment analysis model used a python program. The processes of analysis using the collected data were as follows. In the data

Table 1. Collection data

Rating	Overall	Cleanliness	Location	Rooms	Service	Sleep Quality	Value
1	622	296	148	463	337	242	438
2	998	460	332	745	423	344	703
3	3592	1734	1810	3251	1541	1353	3234
4	11259	5629	4943	7899	3986	3717	8298
5	18208	14399	10840	13788	8817	6844	12221
NA	0	12161	16606	8533	19575	22179	9785

preprocessing process, stop words that are not necessary for the analysis were processed while items such as numbers and special letters that are hard to analyze were eliminated. After data preprocessing, LDA was used to extract topics. In LDA model-based experiment, parameters were designated as follows. The alpha (α) value was set by the number of topic (K) divided by 50 while beta (β) was set to 0.01. Then, LDA was used to extract the adequate topics for the collected hotel reviews. In addition Bing Liu sentiment lexicon was used for the sentiment analysis on the sentences for each extracted topic. Bing Liu sentiment lexicon is composed of 2,006 positive words and 4,683 negative words. This study analyzed the positive cases (sentiment value>0) and negative cases (sentiment value<0) without considering the case where sentiment value is 0.

To evaluate the performance of sentiment analysis constructed based on the suggested model, this study compared the evaluation grade for each attribute and sentiment value obtained through the suggested model. For the performance evaluation criteria, this study used accuracy, precision, recall, and F-measure.

4.3. Results analysis

In the LDA model, K value of the topic and alpha (α) and beta (β) of Dirichlet distribution are to be determined by the researcher. In response, this study combined 5-fold cross-validation to figure out the optimum K valuae with perplexity as the evaluation index. LDA model showed a higher generalization ability in text analysis when perplexity was lower[15]. This study combined perplexity together with cross-validation to figure out the most optimum number of topic. The formula for perplexity is calculated as following Equation 1[16].

$$\text{Perplexity } (D_{test}) = \exp\left\{-\frac{\Sigma_{d=1}^{M}\log P(W_d)}{\Sigma_{d=1}^{M} M_d}\right\} \quad (1)$$

D_{text} indicates the test set and D_{test} includes M number of documents. N_d is the length of the document d while $P(W_d)$ is the probability of document d created by the model.

Table 2 indicates the results of extracting 30 top keywords and 10 related keywords with high probability.

The final keywords related to 6 top attributes were derived by referring to documents based on the keywords extracted by using LDA. Table 3 indicates the sentiment analysis results for sentences extracted by matching to sentences of reviews including keywords. “-1” means negative sentiment while “1” means positive sentiment. When the user review didn’t have any sentence related to the topic, it was indicated as “NA”.

To evaluate the performance of the sentiment analysis constructed based on the suggested model, this study compared the evaluation grade for 6 hotel attributes and sentiment value obtained from the suggested model. For the performance evaluation criteria, this study used recall, accuracy, and F-measure. Accuracy is the percentage of classifying positive and negative cases accurately. Precision is the percentage of actual positive cases among the cases that were assumed to be positive. Recall is the percentage of cases assumed to be positive among groups of correct answers.

Table 2. Topic extraction using LDA

Topic Number	Dimension	Top 10 Words
Topic 1	Service	asked, booking, call, called, card, desk, guests, left, reception, told
Topic 2	Rooms	bath, bathroom, bed, coffee, large, shower, toilet, towels, TV, water
Topic 3	Sleep Quality	air, door, floor, hear, night, noise, open, room, sleep, window
…	…	…
Topic 30	Cleanliness/Location/Service	clean, friendly, great, helpful, highly, location, perfect, recommend, staff, stay

Table 3. Sentiment value matrix for each review-topic

	Cleanliness	Location	Rooms	Service	Sleep Quality	Value
Review1	Positive	Positive	Positive	Negative	NA	Negative
Review2	Negative	NA	Negative	Negative	Positive	Negative
…	…	…	…	…	…	…
Review 11700	NA	Positive	Positive	Positive	NA	NA

Table 4. Confusion matrix

	Actual Positive	Actual Negative
Predicted Positive	TF (True Positive)	FP (False Positive)
Predicted Negative	FN (False Negative)	TN (True Negative)

Table 5. Results of sentiment analysis by topic. (%)

	Accuracy	Precision	Recall	F-measure
Cleanliness	91.61	94.32	93.32	93.62
Location	89.48	90.95	94.12	92.51
Rooms	89.50	92.65	92.26	92.46
Service	92.15	93.85	92.90	93.37
Sleep Quality	78.69	80.69	92.34	86.14
Value	85.13	89.30	90.77	90.03
Average	87.76	90.29	92.62	91.36

$$\text{Precison}=\frac{TP}{TP+FP} \tag{2}$$

$$\text{Recall}=\frac{TP}{TP+FN} \tag{3}$$

$$\text{F-measure}=\frac{2\times(\text{Precison}\times\text{Recall})}{(\text{Precison}\times\text{Recall})} \tag{4}$$

Table 4 is the classification matrix and Equation 2, 3, 4 shows the related recall, precision, and F-measure. Also, Table 5 shows the results of evaluating the sentiment analysis model performance based on recall, precision, and F-measure.

The experiment results were as follows. In “Sleep Quality”, accuracy and precision showed the lowest performance while recall showed relatively higher performance. Also, accuracy for “Room”, “Service”, “Location” and “Cleanliness” was over 85% except “Sleep Quality” and “Value” while the minimum value for F-measure criteria was 86.14%. In the measurement of recall, all of the 6 top topics showed recall over 90%. This demonstrated the effectiveness in classification of sentiment for each topic of the review.

4. Conclusion

This study used LDA to extract topics from the online evaluation review data on hotels that users personally experienced. This study analyzed the topics by using 10 keywords that were frequently used in each topic. By analyzing the extracted topics, this study could explore the hotel topics that users were most interested in.

This study also used LDA-based model to predict the user’s sentiment toward toe topic depending on interest. The topics extracted for evaluating the performance of the suggested sentiment analysis model were used for sentiment analysis on each topic of the reviews and applied with performance evaluation criteria. The suggested model is expected to be applicable not only in hotel fields but also in other fields, too.

The limitations were as follows. The sematic sentiment lexicon holds a limit of depending on the general meaning without understanding the situation or environment where the language is used. Also, there needs to be researches on a sentiment model that includes “Neutral” opinions in sentiment analysis. Furthermore, further sentiment analysis on different topics and fields and researches for enhancing model accuracy are necessary.

References

[1] Shan, S., Liu, X., Wei, Y., Xu, L., Zhang, B., and Yu, L., “A new emergency management dynamic value assessment model based on social media data: a multiphase decision-making perspective”, *Enterprise Information Systems*, vol. 14, no. 5, pp. 680-709, 2020.

[2] Yunxian, C., Renjie, L., Shuliang, Z., and Fenghua, G., “Correction: Measuring multi-spatiotemporal

scale tourist destination popularity based on text granular computing", *PloS one*, vol. 15, no. 5, e0233068, 2020.

[3] Gretzel, U., Zarezadeh, Z., Li, Y., and Xiang, Z., "The evolution of travel information search research: a perspective article", *Tourism Review*, vol. 75, no. 1, pp. 319-323, 2019.

[4] Xianghua, F., Guo, L., Yanyan, G., and hiqiang, W., "Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and HowNet lexicon", *Knowledge-Based Systems*, vol. 37, pp. 186-195, 2013.

[5] Reyes-Menendez, A., Saura, J. R., and Alvarez-Alonso, C., "Understanding# WorldEnvironment-Day user opinions in Twitter: A topic-based sentiment analysis approach", *International journal of environmental research and public health*, vol. 15, no. 11, pp. 2537, 2018.

[6] Nguyen, T. H., and Shirai, K., "Topic Modeling based Sentiment Analysis on Social Media for Stock Market Prediction", *In ACL*, vol. 1, pp. 1354-1364, 2015.

[7] Marrese-Taylor, E., Velásquez, J. D., and Bravo-Marquez, F., "A novel deterministic approach for aspect-based opinion mining in tourism products reviews", *Expert Systems with Applications*, vol. 41, pp. 7764-7775, 2014.

[8] Bauman, K., Liu, B., and Tuzhilin, A., "Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews", *In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 717-725, 2017.

[9] Blei, D. M., and Lafferty, J. D., "A correlated topic model of science", *The annals of applied statistics*, vol. 1, no. 1, pp. 17-35, 2007.

[10] Shams, M., and Baraani-Dastjerdi, A, "Enriched LDA (ELDA): Combination of latent Dirichlet allocation with word co-occurrence analysis for aspect extraction", *Expert Systems with Applications*, vol. 80, pp. 136-146, 2017.

[11] Blei, D. M., Ng, A. Y., and Jordan, M. I., "Latent dirichlet allocation", *Journal of machine Learning research*, vol. 3, pp. 993-1022, 2003.

[12] Lienou, M., Maitre, H., and Datcu, M., "Semantic annotation of satellite images using latent Dirichlet allocation", *IEEE Geoscience and Remote Sensing Letters*, vol. 7, pp. 28-32, 2010.

[13] Lin, C., and He, Y, "Joint sentiment/topic model for sentiment analysis", *In Proceedings of the 18th ACM conference on Information and knowledge management*, pp. 375-384, 2009.

[14] Liu, B., "Sentiment analysis and opinion mining", *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1-167, 2012.

[15] Wang, X., Zhang, B., Huang, M., Chang, F., and Zhou, Z., "Improved LDA Model for Credibility Evaluation of Online Product Reviews", *IEICE TRANSACTIONS on Information and Systems*, vol. 102, no. 11, pp. 2148-2158, 2019.

[16] Liu, B., "Sentiment analysis: Mining opinions, sentiments, and emotions", *Cambridge university press*, 2020.