ISSN 2288-4866 (Print)
ISSN 2288-4882 (Online)
http://www.jiisonline.org
J Intell Inform Syst 2019 March: 25(1): 21~41
http://dx.doi.org/10.13088/jiis.2019.25.1.021

An Analytical Approach Using Topic Mining for Improving the Service Quality of Hotels*

Hyun Sil Moon
School of Business &
AI Management Research Center,
KyungHee University
(*pahunter@khu.ac.kr*)

David Sung
School of Business &
AI Management Research Center,
KyungHee University
(*davidsung88@khu.ac.kr*)

Jae Kyeong Kim
School of Business &
AI Management Research Center,
KyungHee University
(*jaek@khu.ac.kr*)

Thanks to the rapid development of information technologies, the data available on Internet have grown rapidly. In this era of big data, many studies have attempted to offer insights and express the effects of data analysis. In the tourism and hospitality industry, many firms and studies in the era of big data have paid attention to online reviews on social media because of their large influence over customers. As tourism is an information-intensive industry, the effect of these information networks on social media platforms is more remarkable compared to any other types of media. However, there are some limitations to the improvements in service quality that can be made based on opinions on social media platforms. Users on social media platforms represent their opinions as text, images, and so on. Raw data sets from these reviews are unstructured. Moreover, these data sets are too big to extract new information and hidden knowledge by human competences. To use them for business intelligence and analytics applications, proper big data techniques like Natural Language Processing and data mining techniques are needed. This study suggests an analytical approach to directly yield insights from these reviews to improve the service quality of hotels. Our proposed approach consists of topic mining to extract topics contained in the reviews and the decision tree modeling to explain the relationship between topics and ratings. Topic mining refers to a method for finding a group of words from a collection of documents that represents a document. Among several topic mining methods, we adopted the Latent Dirichlet Allocation algorithm, which is considered as the most universal algorithm. However, LDA is not enough to find insights that can improve service quality because it cannot find the relationship between topics and ratings. To overcome this limitation, we also use the Classification and Regression Tree method, which is a kind of decision tree technique. Through the CART method, we can find what topics are related to positive or negative ratings of a hotel and visualize the results. Therefore, this study aims to investigate the representation of an analytical approach for the improvement of hotel service quality from unstructured review data sets. Through experiments for four hotels in Hong Kong, we can find the strengths and weaknesses of services for each hotel and suggest improvements to aid in customer satisfaction. Especially from positive reviews, we find what these hotels should maintain for service quality. For example, compared with the other hotels, a hotel has a good location and room condition which are extracted from positive reviews for it. In contrast, we also find what they should modify in their services from negative reviews. For example, a hotel should improve room condition related to soundproof. These results mean that our approach is useful in finding some insights for the service quality of hotels. That is, from the enormous size of review data, our approach can provide practical suggestions for hotel managers to improve their service quality. In the past, studies for improving service quality relied on surveys or interviews of customers. However, these methods are often costly and time consuming

* This work was supported by a grant from Kyung Hee University in 2013.(KHU-20130544)

and the results may be biased by biased sampling or untrustworthy answers. The proposed approach directly obtains honest feedback from customers' online reviews and draws some insights through a type of big data analysis. So it will be a more useful tool to overcome the limitations of surveys or interviews. Moreover, our approach easily obtains the service quality information of other hotels or services in the tourism industry because it needs only open online reviews and ratings as input data. Furthermore, the performance of our approach will be better if other structured and unstructured data sources are added.

Key Words : service quality, topic mining, decision tree, big data analysis, online review analysis

Received : January 8, 2019 Revised : February 22, 2019 Accepted : March 11, 2019
Publication Type : Conference(Fast-track) Corresponding Author : Jae Kyeong Kim

1. Introduction

Thanks to the rapid development of information technologies, the data available on Internet have grown rapidly. In this era of big data, many studies have attempted to offer insights and express the effects of data. For example, some studies that utilize Google Trends observed the efficiency of data in predicting economic metrics such as unemployment, automobile demand and so on (Askitas and Zimmermann, 2009; Choi and Varian, 2012; Goel et al. 2010). Additionally, Gantz and Reinsel (2011) found that big data technologies are useful in extracting new value and discovering hidden knowledge from an enormous volume of information. The difference between the big data and traditional data environments is not limited to the volume of data. Compared to the traditional data environment, big data analytics need to handle a flexible range of data including semi-structured and unstructured data sets from a variety of sources such as social media and web logs (Zikopoulos and Eaton, 2011).

Especially in the tourism and hospitality industry, many firms and studies have paid attention to online reviews on social media platforms among various other data sources. Social media platforms like blogs and online review sites support customers sharing their experiences (Xiang and Gretzel, 2010). Moreover, other users are easily influenced by these reviewers' opinions because they are considered more credible than firm-created information (Grewal et al., 2003; Herr et al., 1991). As tourism is an information-intensive industry (Werthner and Klein, 1999), the effect of these information networks on social media platforms is more remarkable compared to any other types of media (Xiang and Gretzel, 2010). For example, higher rating scores on social media platforms leads to significantly increased sales for hotels and restaurants (Öğüt and Onur Taş, 2012; Zhang et al., 2010). Additionally, another study indicated the effects of forecasting hotel room sales by investigating reviews on social media platforms (Ye et al., 2011).

However, although many studies and firms are trying to use social media platforms as new information sources, there are some limitations to the improvements in service quality that can be made based on opinions on social media platforms. Users on social media platforms represent their opinions as text, images, and so on. That is, raw data sets from these reviews are unstructured. Moreover, these data sets are too big to extract new value and hidden knowledge by human competences. To use them for business intelligence and analytics applications, proper big data techniques like Natural Language Processing (NLP) and data mining techniques are needed. Among various unstructured data sets, this study utilized text data sets because most reviews are uploaded in the form of textual data. As many studies have been interested in analyzing textual data, text mining techniques, which can offer valuable insights from huge textual datasets, have been developed in many research domains (Cambria et al., 2013; Hagenau et al., 2013; Kusumasondjaja et al., 2012; Salehan and Kim, 2016; Ur-Rahman and Harding, 2012).

In this study, in order to improve the service quality of hotels based on online reviews, we first use topic mining techniques as a form of text mining. Topic mining refers to a method for finding a group of words from a collection of documents (referred to as corpus) that represents a document. Among several topic mining methods, we adopted the Latent Dirichlet Allocation (LDA) algorithm, which is considered the most universal algorithm (Blei et al., 2003). In the LDA algorithm, each review is treated as a mixture of topics, and each word is attributable to one of the review's topics. Therefore, we can find what text was included in each review through the LDA. However, LDA is not enough to find insights that can improve service quality because it cannot find the relationship between topics and ratings. To overcome this limitation, we also use the Classification and Regression Tree (CART) method, which is a kind of decision tree technique (Breiman, 1984). Using a decision tree as one of the classification algorithms in data mining techniques has an advantage for the intuitive visualization of the results; they can be represented as a tree. In particular, CART is a widely utilized decision tree method, in which the tree is built by splitting the data and fitting a prediction model (Loh, 2011). Through the CART method, we can find what topics are related with positive or negative ratings of a hotel and visualize the results. Therefore, this study aims to investigate the representation of an analytical approach for the improvement of hotel service quality from unstructured review data sets. For this purpose, we propose a two-phase analytical procedure based on topic mining and decision tree techniques. Hopefully, this study can provide useful insights to firms and researchers who want to analyze reviews on social media platforms.

2. Research Backgrounds

2.1 Effects of Online Reviews

Currently, many firms try to understand their customers and recognize business trends through the growth of online reviews (Baars and Kemper, 2008). Because online reviews contain the opinions of customers on products or services, they can influence others' decisions and have a tremendous effect on potential sales; therefore, many firms attempt to strategically utilize their online reviews (Dellarocas, 2006; Duan et al., 2008). In other words, online reviews function as one of the factors that determines purchase decisions (Archak et al., 2011; Lee, 2011). Thus, various studies on the practical implications of online reviews have been implemented.

Particularly, earlier studies on online review analysis found that online reviews have effects on the sale of products (Chen and Xie, 2008; Chevalier and Mayzlin, 2006; Dellarocas et al., 2007; Forman et al., 2008). For example, Zhang and Dellarocas (2006) found that online movie review scores could be a significant factor in forecasting box office revenues. In addition, these scores are also an important factor in explaining both aggregate and weekly box office revenues, especially in the early weeks of a film's release (Liu, 2006). These earlier studies found that online reviews influence not only product sales in various fields but also decision-making processes. Therefore, researchers proposed that it is important to find hidden and unexpected knowledge from online reviews. Based on these results, recent studies have focused on extracting useful values from online reviews. Fortunately, most online reviews contain not only textual data but also other types of data such as rating scores expressing their level of satisfaction. Based on these various forms of available data, many studies have utilized online reviews to find some insight for their research domains (Das and Chen, 2007; Mudambi and Schuff, 2010; Salehan and Kim, 2016).

In the tourism and hospitality field, many studies also pay attention to online reviews (Pan et al., 2007; Sparks et al., 2013; Ye et al., 2011; Zhang et al., 2010). Recent customers usually depend on online reviews because social media platforms provide and share significant information with customers anywhere and at any time. Moreover, the growth of tourism-related online review sites such as TripAdvisor, Yelp, and Expedia has accelerated the importance of reviews. Thus, these online communities have been treated as the most important information source in the tourism and hospitality industry (Fang et al., 2016; Liu and Park, 2015; Vermeulen and Seegers, 2009). To be specific, Xiang and Gretzel (2010) pointed out that social media has gradually become the most important information sources for people who plan to travel. Additionally, Sparks and Browning (2011) found that customers' behaviors are reliant on online reviews when customers make decisions on booking. Similarly, the findings in some studies show that a higher rating score leads to significantly increased sales for hotels and restaurants (Öğüt and Onur Taş, 2012; Zhang et al., 2010). In addition, Ye et al. (2011) proved the effects of forecasting room sales in hotels by investigating online reviews. When considering two sentimental sides within a review, positive reviews have facilitated the increased level of reserved hotels and provided an expression of a customer's achievements (Xie et al., 2014). Nonetheless, negative reviews have been considered more reliable than their counterparts (Kusumasondjaja et al., 2012).

In summary, lots of related studies offer insights into customers' decision making and the link to online information in many industries including the tourism and hospitality industry. Based on the above studies, this study specifically provides new findings with a proper analytical approach to online reviews regarding improving the service quality of hotels.

2.2 Topic Mining

Analyzing online reviews as they are requires proper big data techniques such as text mining. Text mining is defined as the process that uncovers hidden knowledge from large amounts of textual data. Although there are many methods in text mining studies, we use topic mining techniques to discover what a review says. The topic mining techniques refer to extracting potential information from textual data based on common issues, and they have been developed and studied in various fields. Among these techniques, we use the Latent Dirichlet Allocation (LDA) algorithm, which has been familiarly used in topic mining studies. For example, based on the LDA algorithm, Zhao et al. (2016) analyze the social trends of a society, and others examine a range of interested topics among the users.

As a kind of unsupervised machine learning method, the LDA algorithm was originally developed as one of the statistical methodologies that considered each topic within an individual document (Park and On, 2017). In the results of the LDA, a topic is represented as a designated group of probability status (Blei et al., 2003).

The following equations are expressed as the probabilistic generative process (Blei and Lafferty, 2009).

(1) For each topic;
  1) Draw the topic distribution over words $\beta_k \sim Dir(\eta)$

(2) For each document
  1) Generate a vector of topic proportions $\theta_d \sim Dir(\alpha)$
  2) For each word
    a) Draw a topic assignment $Z_{d,n} \sim Mult(\theta_d)$, $Z_{d,n} \in \{1, \dots, K\}$
    b) Draw a word $W_{d,n} \sim Mult(\beta_{z_{d,n}})$, $W_{d,n} \in \{1, \dots, V\}$

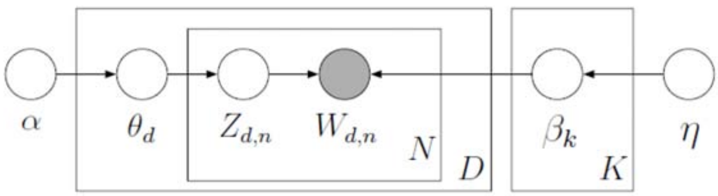
where K is specified as the number of topics, V means the size of the words, and d means a review. As the LDA assumes that each review consists of mixed topics, $Z_{d,n}$ is defined as the topic assignment of the *n-th* word to review d. Additionally, the proportions of parameter α and topic parameter η are hyperparameters to determine the Dirichlet distributions. Then, θ_d and β_k assume the probability of review d is defined as

$$\prod_{n=1}^{N^d} \left(\sum_{k=1}^{K} \mathrm{P}\left(w_{d,n} \middle| z_{d,n} = k, \beta_k\right)\mathrm{P}(z_{d,i} = k \mid \theta_d)\right) = \prod_{n=1}^{N^d} \left(\sum_{k=1}^{K} \beta_{w_{d,n}}^{k} \theta_{d,k}\right) \quad (1)$$

The LDA takes the Bayesian approach and is a complete generative model. Moreover, it assumes that both θ and β follow Dirichlet with hyperparameters α and η. Thus, the generative process of LDA can be denoted by the following joint distribution of the latent variables (topics, topic, proportions, and topic assignments) and observed variables (words) by using Monte Carlo Markov Chain methods based on Bayesian statistics (Griffiths, Steyvers, Blei, & Tenenbaum, 2005).

$$p(w, z, \theta, \beta \mid \alpha, \eta) = \prod_{k=1}^{K} P(\beta_k \mid \eta) \prod_{d=1}^{D} P(\theta_d \mid \alpha) \left(\prod_{n=1}^{N^d} P\left(w_{d,n} \middle| Z_{d,n}, \beta_k\right) P(Z_{d,n} \mid \theta_d)\right) \quad (2)$$

Finally, the estimates of θ and β are obtained by examining the posterior distribution. Figure 1 shows the graphical model for the LDA algorithm.



〈Figure 1〉 The graphical model for the Latent Dirichlet Allocation (Blei & Lafferty, 2009)

As a result, the LDA produces *K* topics which consist of some words. Regarding the determination of the number of topics *K*, this is up to the researchers. However, some studies determined this automatically by the harmonic mean. (Griffiths et al., 2005; Lee and Kim, 2018). Therefore, we also used the harmonic mean to determine the number of topics. Additionally, each topic represents the commonly occurring themes based on the probabilistic distribution for a set of words in documents. Moreover, for each review that consists of topics, a probability value (or weight) for each topic is assigned to each observed topic and the value of sum among participants is to be 1. In other words, the sum of the weighted value is 1 and each document is determined differently based on the designated weight of total number of words. Therefore, it can be stated that this is a proper method in detecting the commonly addressed issues of the overlapping event in this study.

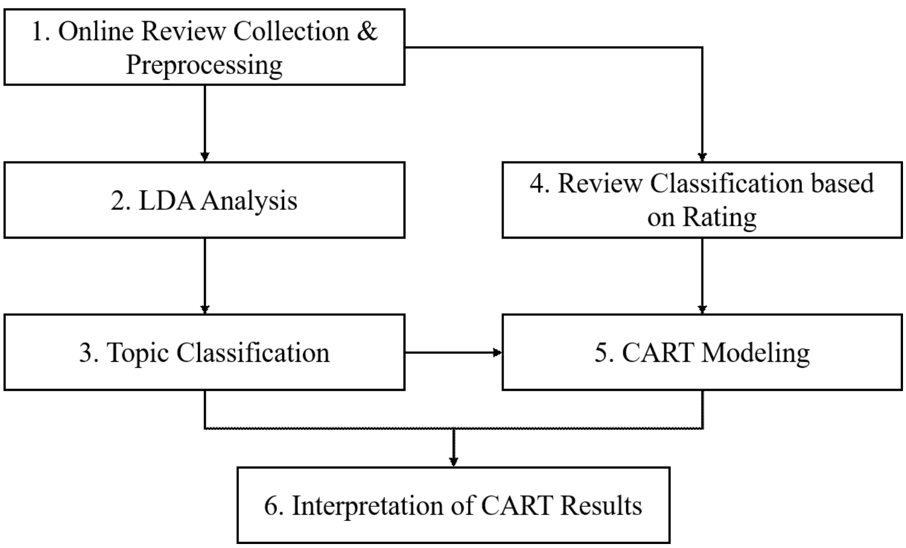
However, this method cannot solely suggest insights for improving the service quality of hotels due to its anonymous nature. To solve this problem, we also use the Classification and Regression Tree (CART), which is a kind of decision tree induction. It constructs trees from data that could show the largest information gain at each node (Breiman, 1984). We expect to find the relation between ratings and topics with CART. Through additional analysis, we can suggest some insights for hotels to improve their service quality.

3. Methodology and Research Design

3.1 Overall Procedure

Figure 2 shows the overall procedure of the proposed analytical approach. First, we collected online review data from an online review site. To obtain customers' opinions about a hotel, we also collected ratings for each review. Then, as the raw review data required preprocessing, which means convertings unclean textual data to clean, we conducted the following steps: removing punctuation, stopping word eliminations, and stemming which is the process of reducing inflected words to their root form.

With the preprocessed review data, we conducted the LDA analysis. Based on harmonic means, the number of topics was automatically selected. The example of topics is illustrated in Table 1. However, as shown in Table 1, initial LDA results just show a collection of words. Therefore, we classified these topics into 12 categories related to service quality: Restaurant, View, Staff, Image of Brand, Event, Information Gain, Room, Facilities, Value, Membership, Location, and Sound Proof. For example, if a topic contains words such as breakfast, buffet, and food,



〈Figure 2〉 Overall Procedure

〈Table 1〉 An Illustrative Example of the LDA analysis

TOPIC	Word-1	Word-2	Word-3	…	Word-n
Topic-1	breakfast	great	food	…	buffet
Topic-2	view	harbour	window	…	room
…					
Topic-n	airport	station	convenient	…	walk

it could be classified into 'Restaurant'.

After the LDA analysis, along with ratings estimated on 5-point scales, we classified reviews into two categories: positive (4 and 5 in rating scores) and negative (1 and 2 in rating scores). Because rating score 3 means a neutral opinion, we exclude them from CART modeling. According to the LDA analysis results, overall reviews are revealed as probabilities of each topic, as shown in Table 2. Each value means the probability of how a topic explains a review. For example, Review-1 in Table 2 could be well explained by Topic-3. We used these values as input variables for CART modeling. Additionally, due to the data imbalance problem, we conducted separately CART modeling for positive and negative reviews. And the target variable for the CART model is whether the target hotel is correct or not. The target hotel refers to a subject to analyze service quality. Therefore, if there are four hotels in our data, there will be 8 CART models which are positive/negative models for each hotel. Finally, based on the proper interpretation of CART results, we can find some insights for each hotel to improve its service quality.

To ensure confidence in our analysis, we divided the review data set into two parts: a training set (70%) for modeling and a testing set (30%) for validation. Then, to evaluate model performance, we estimate the accuracy and the Area Under the ROC(Receiver Operating Characteristics) Curve (simply AUC) values, which have been used to evaluate classification algorithms. Accuracy means how often the model made correct classifications and is calculated as follows:

〈Table 2〉 An Illustrative Example of Each Topic's Proportions for Each Review

REVIEW	Topic-1	Topic-2	Topic-3	…	Topic-n
Review-1	0.05	0.02	0.6	…	0.07
Review-2	0.22	0.13	0.05	…	0.02
…					
Review-n	0.06	0.04	0.03		0.01

$$Accuracy = \frac{\textit{the number of target hotels correctly classified}}{\textit{total number of reviews}} \quad (3)$$

The ROC curve was first used in signal detection studies and represents the relationship between the true positive rate and the false positive rate (Huang & Ling, 2005). The AUC is the area under the ROC curve and provides the performance of an algorithm. For example, a 0.5 value of the AUC equals the random model and 1 value of the AUC, which means the perfect classification model. Therefore, as a higher accuracy and a higher AUC value means a more accurate model, we can experiment with the performance of the CART model.

3.2 Research Design

To discover insights from online reviews, we collected review data from TripAdvisor (www.tripadvisor.com), which is the most popular social media platform in the tourism industry. TripAdvisor provides visitors' reviews for places or hotels with additional information, such as a 5-scale rating. An example of the review data is shown in Figure 3.

Aligned with our purpose, we first selected the city of Hong Kong for our experiments because it is well known for its tourist attractions and various hotels. Moreover, it is an English-speaking area, and reviews are primarily written in English, thereby making it easy to collect the relevant reviews. Then, we selected 4 different famous hotel chains because they provide full service, from restaurants to accommodations. The following are the hotels that were considered in this study: Conrad Hong Kong Hotel (Conrad), InterContinental Hong Kong Hotel (Inter), JW Marriott Hotel Hong Kong (JW), and W Hong Kong Hotel (W). Through examining and comparing the representative hotels, we expect to observe any significant presence of differences in terms of how much each hotel impacted customers' satisfaction.

Raw review data for hotels were collected from

〈Figure 3〉 An Example of Review Data from TripAdvisor

〈Table 3〉 Data Description for Experiments

Hotel	Positive Review	Negative Review	Total
Conrad	1,071	117	1,188
Inter	1,049	57	1,106
JW	689	64	753
W	1,289	50	1,339
Total	4,098	288	4,386

January 2013 to December 2015 (3 years) and consist of 4,576 reviews. However, as mentioned above, we excluded reviews with a rating of 3 because we assume that they did not reveal any opinions. Therefore, we use 4,386 reviews for our experiments. A detailed description of the data set is illustrated in Table 3.

3.3 Experimental Results

To find insights for the improvement of service quality in hotels from online reviews, we first conducted the LDA analysis, which shows the text content of the reviews. As the number of topics are automatically determined by the harmonic mean, we extracted 30 topics from the overall review data set. Then, we classified 30 topics into 12 categories according to the representative words of each topic. Table 4 provides an example of topics with the 5 representative words as well as their designated topic names.

Based on these topics, the overall reviews were revealed with the probability how each topic explains each review. With these values, we conducted CART modeling. To obtain and compare results, we constructed 8 CART models according to hotels and opinions. Their results are shown below.

3.3.1 Results of Positive Reviews

As mentioned above, we used accuracy and AUC to estimate the performance of each model. For positive reviews, the performance results are

〈Table 4〉 An Example of Topic Extraction

Topic	Topic Name	Word-1	Word-2	Word-3	Word-4	Word-5
Topic 1	Restaurant	breakfast	great	food	restaurants	buffet
Topic 2	View	view	harbour	room	victoria	window
Topic 3	Staff	want	just	need	work	things
Topic 4	Image of Brand	service	high	expect	luxurious	standard
Topic 5	Restaurant	lounge	floor	executive	drinks	level

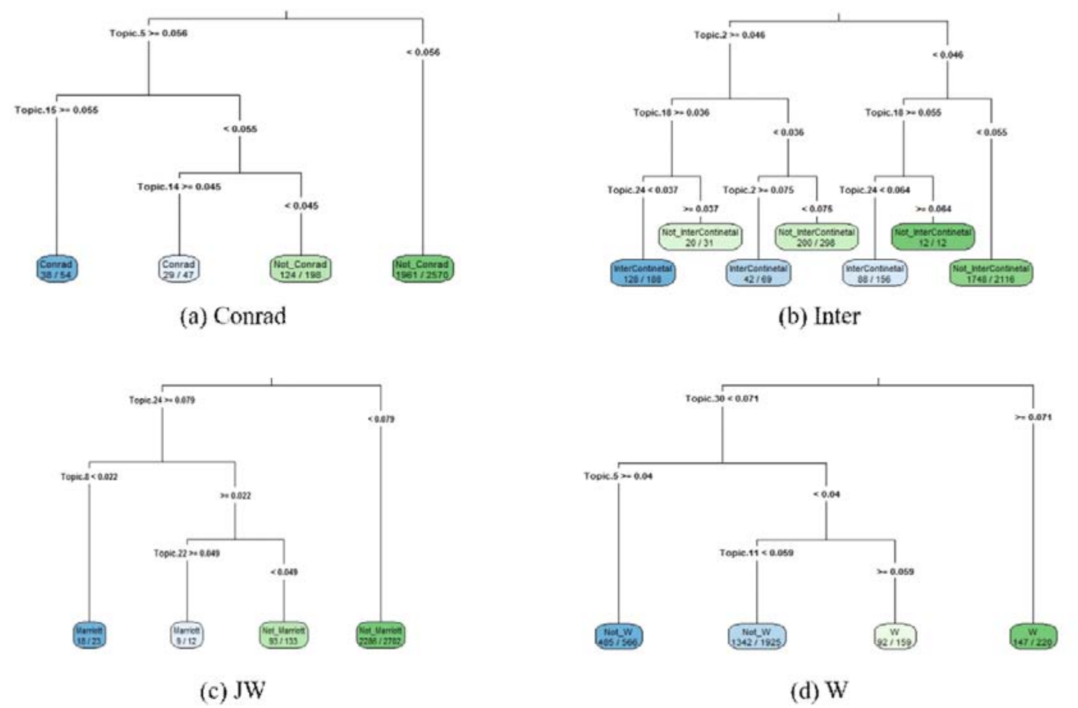
shown in Table 5.

As shown in Table 5, the accuracy of the JW model and the AUC of the Inter model are the highest. Additionally, as overall metrics are above 0.5, they have a higher performance than the random model. Therefore, we could explain what the positive reviews of each hotel said with these models. The results of each model are shown in Figure 4.

Based on these CART results, we extracted main rules to determine the hotel. As these rules could explain what topics are related to the hotel, we are able to yield insights from these results. The main rules are shown in Table 6.

〈Table 5〉 The Performance of Models for Positive Reviews

Hotel	# of Reviews	Accuracy	AUC
Conrad	1,071	0.745	0.567
Inter	1,049	0.775	**0.660**
JW	689	**0.827**	0.530
W	1,289	0.722	0.641



〈Figure 4〉 CART Results for Positive Reviews

〈Table 6〉 Main Rules of Positive Reviews for Each Hotel

Hotel	Rule	
Conrad	(Topic 5 $\geq$ 0.056 & Topic 15 $\geq$ 0.055)	$\Rightarrow$ Conrad
	(Topic 5 $\geq$ 0.056 & Topic 15 $<$ 0.055 & Topic 14 $\geq$ 0.045)	$\Rightarrow$ Conrad
Inter	(Topic 2 $\geq$ 0.046 & Topic 18 $\geq$ 0.036 & Topic 24 $<$ 0.037)	$\Rightarrow$ Inter
	(Topic 2 $\geq$ 0.046 & Topic 18 $<$ 0.036 & Topic 2 $\geq$ 0.075)	$\Rightarrow$ Inter
	(Topic 2 $<$ 0.046 & Topic 18 $\geq$ 0.055 & Topic 24 $<$ 0.064)	$\Rightarrow$ Inter
JW	(Topic 24 $\geq$ 0.079 & Topic 8 $<$ 0.022)	$\Rightarrow$ JW
	(Topic 24 $\geq$ 0.079 & Topic 8 $\geq$ 0.022 & Topic 22 $\geq$ 0.049)	$\Rightarrow$ JW
W	(Topic 30 $\geq$ 0.071)	$\Rightarrow$ W
	(Topic 30 $<$ 0.071 & Topic 5 $<$ 0.04 & Topic 11 $\geq$ 0.059)	$\Rightarrow$ W

Through the interpretation of these results, we can find some insight into what service induces positive opinions in hotel customers. First, in the case of Conrad, topics 5 (Restaurant), 14 (Staff), 15 (Membership) were included in the main rules. To be specific, if a review of the Conrad was written for its restaurant and membership, it would be had a positive rating. Additionally, although it is not related to membership, the score would be positive if it was written for the restaurant and staff. Therefore, we can conclude that the Conrad Hong Kong Hotel provides good services for its restaurant, membership, and staff. Second, the main rules of Inter contained topics 2 (View), 18(View), and 24 (Location). That is, the InterContinental Hong Kong Hotel provides a good view compared to the other hotels. However, topic 24 was always lower than the specific values. This means that a positive review for them should not be about location. In other words, this hotel should provide better access to improve their service quality. Third, the topic 8 (Room), topic 22 (Facilities), and topic 24 (Location) were revealed in the main rules for JW. In particular, the location of JW Marriott Hotel Hong Kong elicits positive opinions from their customers. If a review was written about the room and facilities, it would also be associated with a positive rating. These results mean that the hotel provide good access and facilities. Lastly, the main rules for W could be explained by topic 30 (Location), topic 5 (Restaurant), and topic 11 (Room). Compared with the other hotels, the W Hong Kong Hotel has a good location and room condition but a bad restaurant because Topic 5 was always lower than the specific values.

Consequently, with our proposed analytical approach, we could find some insights for each hotel related with positive reviews. These findings mean that each hotel should maintain their advantages to provide good services for their customers.

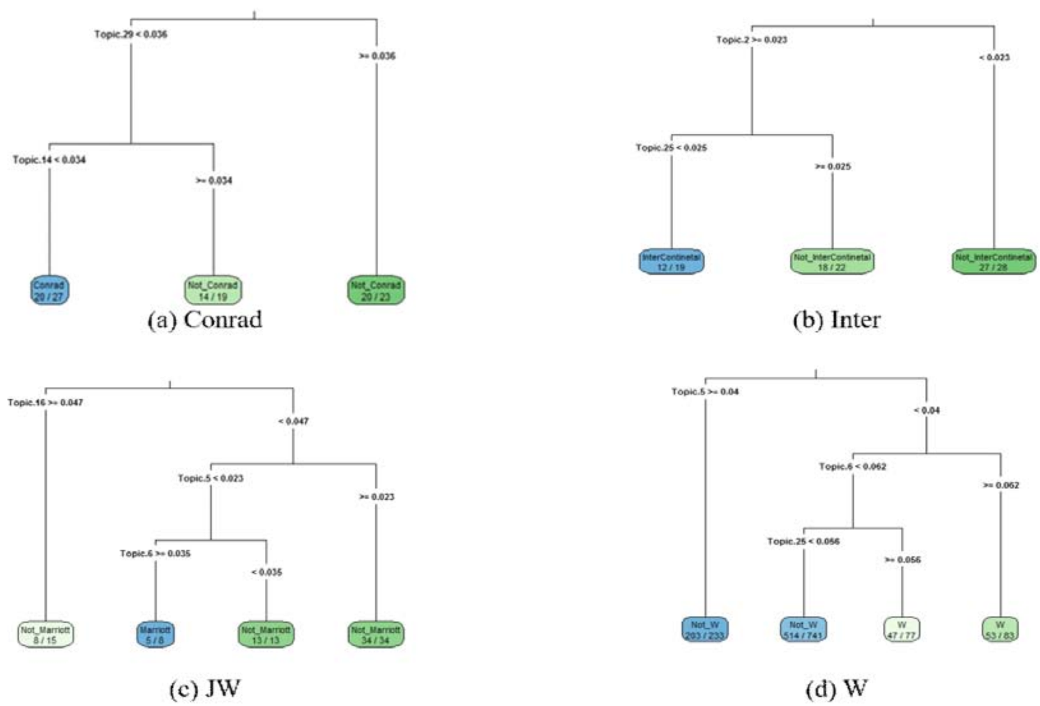
3.3.2 Results of Negative Reviews

In addition to what services should maintain, it is also important to discover what hotels should modify in their services to improve service quality. For this purpose, we then analysed negative reviews for hotels. The performance results for the models of negative reviews are shown in Table 7.

In Table 7, the overall performance is lower than the performance for positive reviews because the number of reviews is too small to construct the CART model. However, as they are shown above the 0.5 value, we can find some insights using these models. The CART results for negative reviews are shown in Figure 5.

〈Table 7〉 The Performance of Models for Negative Reviews

Hotel	# of Reviews	Accuracy	AUC
Conrad	117	0.586	0.571
Inter	57	0.759	**0.700**
JW	64	**0.786**	0.609
W	50	0.759	0.538



〈Figure 5〉 CART Results for Negative Reviews

〈Table 8〉 Main Rules of Negative Reviews for Each Hotel

Hotel	Rule	
Conrad	(Topic 29 < 0.036 & Topic 14 <0.034)	⇒ Conrad
Inter	(Topic 2 ≥ 0.023 & Topic 25 < 0.025)	⇒ Inter
JW	(Topic 16 < 0.047 & Topic 5 < 0.023 & Topic 6 ≥ 0.035)	⇒ JW
W	(Topic 5 < 0.04 & Topic 6 ≥ 0.062) (Topic 5 < 0.04 & Topic 6 < 0.062 & Topic 25 ≥ 0.056)	⇒ W ⇒ W

To find what these hotels can improve, we extracted the main rules of CART results in Figure 5. Table 8 shows these results.

As shown in Table 8, we could not find many rules compared with positive reviews due to the number of negative reviews. These results also mean that it is difficult to find what services influence the negative reviews. Despite these limitations, we could find some insights from negative reviews; they are as follows. First, negative reviews for the Conrad Hong Kong Hotel were not related to topic 14 (Staff) and topic 29 (Facilities). These results correspond to the results for the positive reviews of Conrad because their customers are satisfied with the restaurant, staff, and membership, as mentioned above. However, we could not find specific rules related to what Conrad could not provide in terms of services. Second, in the case of the InterContinental Hong Kong Hotel, their negative reviews were written for topic 2 (View) and not topic 25 (Sound Proof). Specifically, the reviews were not written about sound proofing. That is, we can conclude that their customers are satisfied with their room conditions related to sound proofing. However, it is curious that topic 2 was also found in positive reviews. These results stem from the characteristic of reviews. As a review contains only positive or negative opinions, their negative reviews could also contain positive service factors. Therefore, we can conclude that their customers complain about services in the same review with positive services. Next, the customers of the JW Marriott Hotel Hong Kong wrote their negative reviews about topic 6 (Event). Topic 6 in the LDA analysis means the serendipity of hotel services, which refers to finding valuable things not sought for. That is, their customers could not find any additional enjoyment out of services. Therefore, as their customers were satisfied with the basic facilities such as the restaurant or room conditions, they should pay attention to providing serendipity in customers' experiences. Lastly, the customers of W Hong Kong Hotel also need serendipity in services. Moreover, customers also wrote reviews complaining about room conditions (topic 25: Sound Proof). Although they are satisfied with the restaurant (topic 5), the W Hong Kong Hotel should improve room conditions according to their negative reviews.

3.4 Discussion

Through our experiments to find some insights from online reviews, we found what hotels could improve with regard to for their services. Table 9 summarizes the findings.

〈Table 9〉 The Strengths and Weaknesses of Four Hotels

Hotel	Strength	Weakness
Conrad	Restaurant Staff Membership	-
Inter	View	Location
JW	Room Facilities Location	Event
W	Location Restaurant	Event Room Sound Proof

In Table 9, the left side indicates the strengths of each hotel, which should be maintained for service quality because customers are satisfied with these amenities. To be specific, the customers of the InterContinental Hong Kong Hotel were satisfied with the view from the hotel. Online reviews for the JW Marriott Hotel Hong Kong and the W Hong Kong Hotel revealed positive opinions about their locations and related facilities such as rooms or restaurants.

Next, the right side indicates the weaknesses of each hotel, which should be modified because negative reviews highlighted these issues. For example, as the location of InterContinental Hong Kong Hotel is hard to access, they should provide proper transportation to improve their service quality. In the case of the JW and the W, their customers require some serendipitous services. Therefore, they need to develop some events to maintain customer satisfaction. Moreover, customers of the W also pointed out the sound proofing of rooms in reviews. Based on these results, they should improve their facilities. However, we could not find any specific weakness factors for the Conrad Hong Kong Hotel. These results do not mean that they do not need to modify services; however, it does indicate that we could not find insights because their negative review data set is not sufficient to construct a model.

4. Conclusion

In the tourism and hospitality industry, social media platforms are important data sources used to find insights to improve service quality. Hotels especially want to improve their services based on customers' sincere evaluation. In the past, it was difficult to collect these reviews because of the limitations related to time and place. However, thanks the social media platforms such as TripAdvisor and so on, hotels can collect the truthful reviews of customers. Although many studies indicate that these review data are crucial for the hotel industry, it is hard to analyze reviews because they are written in text, which is hard for computers to understand.

In this study, we suggest an analytical approach to improve service quality of hotels based on online reviews. For our purpose, we propose a 2-phase analytical approach. First, to obtain what customers said in online reviews, we conducted topic mining based on the LDA analysis. Through topic mining, we were able to find some words (topics) that represent reviews, and these words were used to explain each review. Next, we construct a CART model to explain the relationship between ratings and topics. Through our approach, hotels can directly find what they can improve in terms of services from their reliable customers' reviews. In the past, studies for improving service quality relied on surveys or interviews of customers. However, these methods are often costly and time consuming and may be biased because of limitations, such as biased sampling or untrustworthy answers. As a method to overcome these limitations, our proposed approach directly obtains honest feedback from customers' online reviews and draws some insights through a type of big data analytics.

To validate our approach, we experiment with four hotels in the city of Hong Kong. As a result, we find some insights from positive and negative reviews. Especially from positive reviews, we find what these hotels should maintain for service quality. In contrast, we also find what they should modify in their services from negative reviews. These results mean that our approach is useful in finding some insights for the service quality of hotels. Moreover, we also expect that our approach could be used to obtain new knowledge for other hotels or services in the tourism industry because it needs only online reviews and ratings as input data.

However, although this study suggests an analytical approach for an online review data set, it has some limitations. First, in experiments for negative reviews, the size of the negative reviews is too small to construct models. In general, text mining which finds novel information from large textual databases needs to have big data for constructing models. Therefore, we could not find insights for some hotels because negative reviews for those hotels were not large enough to build models. In future studies, we will collect and analyse samples with a big enough data size to find additional insights for improving the service quality of hotels. Next, we selected four hotels in the city of Hong Kong for our experiments. However, the original characteristics of Hong Kong may influence customers' hotel reviews. Therefore, we will expand our experiments to other countries and compare the results to obtain other insights for hotel services. Finally, the explicit rating scores may be affected by the tendency of the reviewer. For example, positive reviewers have rated higher scores for hotels than other reviewers. These tendencies could be modified through proper pre-processing techniques such as a correction with the average per reviewers. Therefore, we will propose a proper correction method for tendencies of reviewers in future works.

References

Archak, N., A. Ghose, and P.G. Ipeirotis, "Deriving the pricing power of product features by mining consumer reviews," *Management science*, Vol.57, No.8(2011), 1485~1509.

Askitas, N., and K.F. Zimmermann, "Google econometrics and unemployment forecasting," *Applied Economics Quarterly*, Vol.55, No.2(2009), 107~120.

Baars, H., and H.G. Kemper, "Management support with structured and unstructured data—an integrated business intelligence framework," *Information Systems Management*, Vol.25, No.2(2008), 132~148.

Blei, D. M., and J. Lafferty, *Text mining: Classification, clustering, and applications*, Chapman & Hall/CRC, 2009.

Blei, D. M., A.Y. Ng, and M.I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, Vol.3, No.1(2003), 993~1022.

Breiman, L, *Classification and regression trees*, Routledge, New York, 1984.

Cambria, E., B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intelligent Systems*, Vol.28, No.2(2013), 15~21.

Chae, S.H., J.I. Lim, and J. Kang, "A Comparative Analysis of Social Commerce and Open Market Using User Reviews in Korean Mobile Commerce," *Journal of Intelligent Information Systems,* Vol.21, No.4(2015), 53~77.

Chen, Y., and J. Xie, "Online consumer review: Word-of-mouth as a new element of marketing communication mix," *Management science*, Vol.54, No.3(2008), 477~491.

Chevalier, J. A., and D. Mayzlin, "The effect of word of mouth on sales: Online book reviews," *Journal of marketing research*, Vol.43, No.3(2006), 345-354.

Choi, H., and H. Varian, "Predicting the present with Google Trends," *Economic Record*, Vol.88(2012), 2-9.

Das, S. R., and M.Y. Chen, "Yahoo! for Amazon: Sentiment extraction from small talk on the web," *Management science*, Vol.53, No.9 (2007), 1375~1388.

Dellarocas, C., "Strategic manipulation of internet opinion forums: Implications for consumers and firms," *Management science*, Vol.52, No.10(2006), 1577~1593.

Dellarocas, C., X. Zhang, and N.F. Awad, "Exploring the value of online product reviews in forecasting sales: The case of motion pictures," *Journal of Interactive marketing*, Vol.21, No.4(2007), 23~45.

Duan, W., B. Gu, and A.B. Whinston, "Do online reviews matter?—An empirical investigation of panel data," *Decision support systems*, Vol.45, No.4(2008), 1007~1016.

Fang, B., Q. Ye, D. Kucukusta, and R. Law, "Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics," *Tourism Management*, Vol.52(2016), 498~506.

Forman, C., A. Ghose, and B. Wiesenfeld, "Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets," *Information Systems Research*, Vol.19, No.3(2008), 291~313.

Gantz, J., and D. Reinsel, "Extracting value from chaos," *IDC review*, Vol.1142(2011), 1~12.

Goel, S., J.M. Hofman, S. Lahaie, D.M. Pennock, and D.J. Watts, "Predicting consumer behavior with Web search," *Proceedings of the National Academy of Sciences(PNAS)* (2010).

Grewal, R., T.W. Cline, and A. Davies, "Early-entrant advantage, word-of-mouth communication, brand similarity, and the consumer decision-making process," *Journal of Consumer Psychology*, Vol.13, No.3 (2003), 187~197.

Griffiths, T. L., M. Steyvers, D.M. Blei, and J.B. Tenenbaum, "Integrating topics and syntax," *In Advances in neural information processing systems* (2005), 537~544.

Hagenau, M., M. Liebmann, and D. Neumann, "Automated news reading: Stock price prediction based on financial news using context-capturing features," *Decision Support Systems*, Vol.55, No.3(2013), 685~697.

Herr, P. M., F.R. Kardes, and J. Kim, "Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnosticity perspective," *Journal of consumer research*, Vol.17, No.4(1991), 454~462.

Huang, J., and C.X. Ling, "Using AUC and accuracy in evaluating learning algorithms," *IEEE Transactions on knowledge and Data Engineering*, Vol.17, No.3(2005), 299~310.

Kusumasondjaja, S., T. Shanka, and C. Marchegiani, "Credibility of online reviews and initial trust: The roles of reviewer's identity and review valence," *Journal of Vacation Marketing*, Vol.18, No.3(2012), 185~195.

Lee, J., "How eWOM reduces uncertainties in decision-making process: using the concept of entropy in information theory," *The Journal of Society for e-Business Studies*, Vol.16, No.4(2011), 241~256.

Lee, M. and H.J. Kim, "Construction of Event Networks from Large News Data Using Text Mining Techniques," *Journal of Intelligent Information Systems,* Vol.24, No.1(2018), 183~203.

Liu, Y., "Word of mouth for movies: Its dynamics and impact on box office revenue," *Journal of marketing*, Vol.70, No.3(2006), 74~89.

Liu, Z., and S. Park, "What makes a useful online review? Implication for travel product websites," *Tourism Management*, Vol.47(2015), 140~151.

Loh, W. Y., "Classification and regression trees," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol.1, No.1(2011), 14~23.

Mudambi, S. M., and D. Schuff, "Research note: What makes a helpful online review? A study of customer reviews on Amazon. com.," *MIS quarterly*, Vol.34, No.1(2010), 185~200.

Öğüt, H., and B.K. Onur Taş, "The influence of internet customer reviews on the online sales and prices in hotel industry," *The Service Industries Journal*, Vol.32, No.2(2012), 197~214.

Pan, B., T. MacLaurin, and J.C. Crotts, "Travel blogs and the implications for destination marketing," *Journal of Travel Research*, Vol.46, No.1(2007), 35~45.

Park, S.M., and B.W. On, "Latent topics-based product reputation mining," *Journal of Intelligent Information Systems,* Vol.23, No.2

(2017), 39~70.

Salehan, M., and D.J. Kim, “Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics,” *Decision Support Systems*, Vol.81(2016), 30~40.

Sparks, B. A., and V. Browning, “The impact of online reviews on hotel booking intentions and perception of trust,” *Tourism management*, Vol.32, No.6(2011), 1310~1323.

Sparks, B. A., H.E. Perkins, and R. Buckley, “Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior,” *Tourism Management*, Vol.39 (2013), 1~9.

Ur-Rahman, N., and J.A. Harding, “Textual data mining for industrial knowledge management and text classification: A business oriented approach,” *Expert Systems with Applications*, Vol.39, No.5(2012), 4729~4739.

Vermeulen, I. E. and D. Seegers, “Tried and tested: The impact of online hotel reviews on consumer consideration,” *Tourism management*, Vol.30, No.1(2009), 123~127.

Werthner, H., and S. Klein, *Information technology and tourism: a challenging relationship*. Springer-Verlag Wien, 1999.

Xiang, Z., and U. Gretzel, “Role of social media in online travel information search,” *Tourism management*, Vol.31, No.2(2010), 179~188.

Xie, K. L., Z. Zhang, and Z. Zhang, “The business value of online consumer reviews and management response to hotel performance,” *International Journal of Hospitality Management*, Vol.43(2014), 1~12.

Ye, Q., R. Law, G. Gu, and W. Chen, “The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings,” *Computers in Human behavior*, Vol.27, No.2(2011), 634~639.

Zhang, X., and C. Dellarocas, “The lord of the ratings: is a movie's fate is influenced by reviews?,” *International Conference on Information Systems 2006 proceedings*, (2006), 117.

Zhang, Z., Q. Ye, R. Law, and Y. Li, “The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews,” *International Journal of Hospitality Management*, Vol.29, No.4(2010), 694~700.

Zhao, F., Y. Zhu, H. Jin, and L.T. Yang, “A personalized hashtag recommendation approach using LDA-based topic model in microblog environment,” *Future Generation Computer Systems*, Vol.65((2016), 196~206.

Zikopoulos, P., and C. Eaton, *Understanding big data: Analytics for enterprise class hadoop and streaming data*, McGraw-Hill Osborne Media, 2011.

국문요약

호텔 산업의 서비스 품질 향상을 위한 토픽 마이닝 기반 분석 방법

문현실* · 성다윗* · 김재경**

정보 기술의 발전으로 온라인에서 활용 가능한 데이터의 양이 급속히 증대되고 있다. 이러한 빅데이터 시대에 많은 연구들이 통찰력을 발견하고 데이터의 효과를 입증하기 위해 노력하고 있다. 특히 관광 산업의 경우 정보에 민감한 사업으로 소셜 미디어의 영향력이 높고 소셜 미디어의 상품 후기에 소비자들이 영향을 많이 받아 많은 기업과 연구자들이 소셜 미디어를 분석하여 새로운 서비스 및 통찰력을 얻고자 시도하였다. 하지만 소셜 미디어의 후기는 텍스트로 이루어진 대표적인 비정형 데이터로 적절한 처리를 하지 않으면 분석에 활용할 수 없다. 또한 후기 데이터의 양이 방대함에 따라 사람이 직접 분석하기도 어려운 실정이다. 따라서, 본 연구에서는 이러한 소셜미디어 상의 온라인 후기로부터 직접 호텔의 서비스 품질 향상을 위한 통찰력을 추출할 수 있는 분석 방법을 제시하고자 한다. 이를 위해 본 연구에서는 먼저 후기 데이터에 포함되어 있는 주제어를 추출하는 토픽 마이닝 기법을 적용하였다. 토픽 마이닝은 대용량의 문서 집합으로부터 문서를 대표하는 단어 집합을 추출하는 기법을 의미하며 본 연구에서는 다양한 연구에서 활용되고 있는 LDA모형을 사용하여 토픽 마이닝을 수행하였다. 하지만, 토픽 마이닝 자체만으로는 주제어와 평점 사이의 관계를 도출할 수 없어 서비스 품질 향상을 위한 통찰력을 발견하기 어렵다. 그에 따라 본 연구에서는 토픽 마이닝의 결과값을 기반으로 의사결정나무 모형을 사용하여 주제어와 평점 사이의 관계를 도출하였다. 이러한 방법론의 유용성을 평가하기 위해 홍콩에 있는 4개 호텔의 온라인 후기를 수집하고 제안한 방법론의 분석 결과를 해석하는 실험을 진행하였다. 실험 결과 긍정 후기를 통해 각 호텔이 유지해야할 서비스 영역을 발견할 수 있었으며 부정 후기를 통해 개선해야할 서비스 영역을 도출할 수 있었다. 따라서, 본 연구에서 제안한 방법론을 사용하여 방대한 양의 후기 데이터로부터 서비스 개선 및 유지 영역을 발견할 수 있으리라 기대된다.

주제어 : 서비스 품질, 토픽 마이닝, 의사결정나무, 빅데이터 분석, 온라인 후기 분석

논문접수일 : 2019년 1월 8일　논문수정일 : 2019년 2월 22일　게재확정일 : 2019년 3월 11일
원고유형 : 학술대회(급행)　교신저자 : 김재경

* 경희대학교 경영학부 & AI경영연구센터
** 교신저자 : 김재경
경희대학교 경영학부 & AI경영연구센터
26, Kyungheedae-ro, Dongdaemun-gu, Seoul 02447, Korea
Tel: +82-2-961-9355, Fax: +82-2-961-0515, E-mail: jaek@khu.ac.kr

저 자 소 개

Hyun Sil Moon

Hyun Sil Moon obtained his M.S. and Ph.D. in Management Information Science (MIS), and his B.S. in Business Administration from KyungHee University. His current research interests focus on big data analysis, recommender systems, text mining, and social network analysis. He has published numerous papers which have appeared in International Journal of Information Management, Asia Pacific Journal of Information System, Journal of Intelligence and Information Systems, and so on.

David Sung

David Sung received the B.S degree from Konkuk University of Mathematics/ Economics and M.S degree from School of Management, KyungHee University. His research interests include big data analysis, text mining, and Machine learning in field of Medical and clinical research.

Jae Kyeong Kim

Jae Kyeong Kim is a professor at School of Management, Kyung Hee University. He obtained his M.S. and Ph.D. in Management Information Systems (MIS) from KAIST (Korea Advanced Institute of Science and Technology), and his B.S. in Industrial Engineering from Seoul National University. His current research interests focus on business analytics, recommender systems, big data analysis, and deep learning. He has published numerous papers which have appeared in IEEE transactions on Systems, Men and Cybernetics, IEEE transactions on Services Computing, Information & Management, International Journal of Information Management, Applied Artificial Intelligence, Artificial Intelligence Review, International Journal of Human-Computer Studies, Expert Systems with Applications, and Group Decision and Negotiations.