

Received April 22, 2020, accepted May 22, 2020, date of publication May 25, 2020, date of current version June 5, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2997522

Framework for Sentiment-Driven Evaluation of Customer Satisfaction With Cosmetics Brands

JAEHUN PARK^{ID}

Major in Industrial and Quality Engineering, Daegu Haany University, Gyeongsan 38610, South Korea

e-mail: pjh3479@dhu.ac.kr

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018R1C1B5033711).

ABSTRACT Cosmetics brand managers' efforts in monitoring customer satisfaction and service quality have suffered due to the lack of effective analysis methods. In order to derive more comprehensive and objective insights into customer opinions on product quality and preferences for cosmetics brands, this study derived an online-review-based process for evaluation of customer satisfaction. The present study developed a systematic approach to the evaluation of relative customer satisfaction with cosmetics brands via sentiment analysis and statistical data analysis, and interpreted the determinants of positive and negative opinions via Term Frequency-Inverse Document Frequency (TF-IDF) analysis. To illustrate the efficacy and applicability of the proposed approach, an empirical case study applying it to the global top 26 cosmetics brands was conducted, which evaluated relative customer satisfaction with brands and examined the main causes of positive and negative opinions. The proposed approach is expected to be employed by cosmetics companies to realize or improve satisfaction with the brands that customers evaluate. Furthermore, we hope that it can be used as a source of fundamental data that could be applied to efforts to improve both brand competitiveness and provision of systematic services.

INDEX TERMS Sentiment analysis, text-mining, customer satisfaction, online reviews, cosmetics brand.

I. INTRODUCTION

Achieving high product and service quality together with high customer satisfaction is a fundamental duty of a business manager in continuously developing business opportunities and running day-to-day operations in a competitive environment. Customer satisfaction affects a company's profit margin, and what determines customers' satisfaction is the product quality offered to them. In other words, customer satisfaction affecting a company's profits is a result of service quality management. Thus, evaluating levels of customer satisfaction is a crucial factor for companies.

Advances in online and mobile technologies have created new consumption patterns for consumers. These technologies have expanded product distribution (which had been limited to the offline realm) to the online world of internet commerce (e-commerce). With the changes to consumers' purchasing patterns, their decision-making processes have come to be affected by opinions influenced not only by cultural leaders but also ordinary people. According to a 2016 consumer

report by DMC Media in South Korea [1], about 51.5% of consumers have shared and posted their consumption experiences online. In other words, consumers share their experiences as to the positive and negative aspects of a given product or service online, and these online reviews are used as important information for new customers in making product-purchase decisions as well as for business managers seeking to understand current customer perceptions of their products. In addition, Archak *et al.* [2] and Ghose and Iperiotis [3] have shown strong evidence that online reviews affect retail sales and consumer purchasing decisions as well as being an important means of analyzing customers' requirements for, and evaluating their satisfaction with, product or service quality.

Online reviews analysis is applicable to various kinds of products, but its value may be greater for the high-sensitivity product category. According to [4], cosmetics are very sensitive to the customers' opinions given the many different types, in addition to the fact that even the same cosmetics may have different effects on the skin depending on the skin characteristics of the individual. For these reasons, cosmetics belong to the high-risk product category, sales of which are

The associate editor coordinating the review of this manuscript and approving it for publication was Santhosh Kumar Gopalan.

highly influenced by customers' product opinions. In general, when a customer wants to buy a cosmetics product on an online website, he or she will typically start by searching for reviews by other consumers of the various offerings so as to complement their own risk perceptions of the product. Due to cosmetics' characteristics such as those just noted, they are highly dependent on the purchasing experiences of customers compared with general products such as home appliances or food [5], [6].

For cosmetics companies that need to continuously improve their product quality and develop new products to gain selling advantage, online reviews are key information from which customer requirements and quality satisfaction can be identified [7]. CJ Mall, an online shopping mall in South Korea, once announced that the sales of products of which customers have positive opinions averaged 2.5 to 5 times those of other products, and emphasized, in that regard, the importance of online review analysis and applications [8]. According to a survey conducted by Timon, a social commerce company in South Korea, online reviews searching ranked second, at 18%, among the factors that consumer considers when purchasing cosmetics. On the other hand, brands (ranked sixth at 5%), product awareness (ranked seventh at 3%), TV commercials and celebrity marketing (0.6%), which are traditionally considered important in the beauty industry, were found to have only low impacts as purchase-decision factors [9].

Given the increased value of online reviews analysis, several studies have presented various methods for extracting knowledge from customer reviews and demonstrating it on product maps [7], [10]. Nevertheless, most of the relevant studies still focus on the lodging and consumer electronic industries [11]–[14]. And, even though some studies have dealt with the application of text-mining techniques in the development of cosmetics products, the majority analyzed only data collected through traditional marketing research methods such as market surveys or questionnaires surveys [7], [15], [16]. However, the traditional marketing research method has two major drawbacks, in that it is limited to collecting various opinions by data sampling, and cannot quickly identify the real-time requirements of consumers (since it takes much time to gather and analyze massive data) [17], [18]. On the other hand, utilizing online reviews offers the advantage of being able to collect and analyze multifaceted information on products from various customers in a short time and at relatively low cost. Above all, online reviews are less distorted respecting product information in that they are written voluntarily by customers. According to [19] and [20], online reviews have been widely utilized as key material, both for evaluating product and service quality satisfaction and for establishing marketing strategies that reflect the real opinions of consumers.

Cosmetics brand managers' efforts to monitor customer satisfaction with product or service quality have suffered due to the lack of any specific methodology for effective analysis. Also, from customers' point of view, it is difficult

to see all of the extensive online reviews, and so they often refer to only company-promoted advertising and a limited number of promotional reviews in making a cosmetics purchasing decision. To derive more comprehensive and objective insights into customer satisfaction and preferences for product and service quality as pertain to cosmetics brands, the present study attempted to obtain customers' opinions and priorities by mining online reviews. Above all, this study focuses on the online review-based relative evaluation of customer satisfaction through customer sentiment analysis with cosmetics brands. The present study developed a systematic approach to the evaluation of relative customer satisfaction via sentiment analysis and statistical data analysis, and interpreted the determinants of positive and negative opinions via term frequency analysis. Dwayne *et al.* [21], Peterson and Meria [22], and Henning *et al.* [23] insisted that customers' sentiments on products affects their quality satisfaction and evaluation, and that the more positive the sentiment in the reviews, the higher the products' quality satisfaction. Thus, the customer satisfaction of this study is evaluated by positive feedback from consumer-generated content, particularly online reviews. To illustrate the efficacy and applicability of the proposed approach, an empirical case study applying it to the global top 26 cosmetics brands was conducted, in which the relative customer satisfactions with brands were evaluated and the main causes of positive and negative opinions were examined.

The main contributions of this paper include: (1) a first-time demonstration of how online reviews are utilized for relative customer satisfaction evaluation of cosmetics brands; (2) an analysis of the main causes of polarity opinions (i.e. positive and negative), with provision of recommendations for improving quality satisfaction for specific cosmetics brands.

Although several researchers have utilized sentiment analysis of online customer reviews in various areas, most of them have focused on sentiment classification and summarization [16], improvement of text-mining techniques to automatically extract knowledge from text [10], [24], [25], customer subjectivity and behavior analysis [26]–[29], and service quality evaluation [30]. They did not cover the evaluation of relative customer satisfaction among competitors or, more importantly, provide any interpretation of the determinants of customer opinions on products and services. In short, this study has a contribution and differs from the previous study in that it applies a statistical data analysis to the well-known sentiment analysis to estimate level of customer satisfaction among cosmetics brands from a relative evaluation perspective. To our best knowledge, our study is the first online sentiment-driven relative evaluation of customer satisfaction with cosmetics brands.

The rest of this study is as follows: Section 2 introduces the basic concepts along with definitions of online customer reviews and sentiment analysis. Section 3 proposes the overall evaluation method and explains in detail the process steps therein. Section 4 presents a case study in which the

proposed approach was applied, and explains the results. Finally, Section 5 concludes the paper and anticipates future research.

II. BACKGROUND RESEARCH

A. ONLINE CUSTOMER REVIEWS

The act of transferring information between individuals in the forms of opinions and sentiments is called Word-of-Mouth (WOM), and online information of this type is called online WOM. A typical form of online WOM is online customer reviews, which are peer-generated product or service evaluations including personal opinions or sentiments that are posted on a company's or third-party website [31]. Since online reviews can share information about various products with many customers beyond space and time, they can have a stronger influence on purchasing decisions and product images than offline WOM [32]. Chatterjee [33] argued that consumer opinions and sentiments on product usage as expressed at online review sites are more objective and credible than the product information unilaterally promoted by companies. According to [12] and [16], online reviews are generally considered to be more honest, unbiased, and comprehensive than information released by sellers. As a source of information, online reviews are seen to be useful for both customers and product providers. On the customers' side, online reviews can be used to support purchasing decisions, and on the product providers' side, they can be utilized to understand customers' current preferences and be applied for product improvement, marketing, and customer relationship management [34].

Online reviews are one of the most useful means, for companies, of obtaining customer requirements and perceptions of products. For this reason, with the development of e-commerce, research on online reviews has become increasingly important. The impact of online reviews on consumer purchasing decisions can vary depending on customers' expressed opinions. Positive reviews, such as compliments or explanations of useful points, have a positive effect on purchases, while negative reviews, such as dissatisfaction with the product, have a negative effect on purchases. In particular, according to [22], [35], and [36], online reviews with negative opinions have more influence on the purchasing intentions of consumers than do reviews with positive opinions, and according to [23], subjective and emotional reviews have a greater influence on purchases than objective and realistic ones. Although these results will vary depending on the nature and types of products, online reviews include not only objective information about the product, but also consumption experience and personal sentiments [21].

Regarding online reviews, several researchers have focused on characteristics based on subjectivity classification, sentiment classification, and opinion summarization [16]. In the literature, most of the studies focus on the improvement of text-mining techniques to automatically extract knowledge from text in more effective and efficient

ways [10], [24], [25]. A few studies have been done on subjectivity analysis, which aims at determining whether a sentence is subjective or objective [26], as well as sentiment analysis of online customer reviews [27]. Also, some researchers have emphasized the relationships among product review content, conceptual cues, and review helpfulness [37], [38]. In addition, several researchers have utilized sentiment analysis for new product development in the cosmetics industry [4], to find out customer preference by analyzing subjective expressions [39] and also for evaluation of quality satisfaction with mobile services [30]. In short, although an abundance of studies have proposed numerous automatic processes and algorithms for the extraction of online reviews, most of them have focused on text data analysis, not evaluation of relative customer satisfaction and analysis of the causes of customer satisfaction.

B. SENTIMENT ANALYSIS

Sentiment analysis is defined as the task of finding the opinions, such as the feelings and emotions, of individuals with respect to specific entities, and their automatic extraction [40]. Amid the constant increase of information in terms of opinions, emotions and feelings, sentiment analysis is attracting more and more interest in the scientific community [41]. Sentiment analysis can be used interchangeably with opinion mining in the sense of deriving consumers' feelings, emotions and opinions from reviews. Studies such as [42]–[46] did not distinguish the definitions of opinion mining and sentiment analysis in analyzing sentiments expressed in text. On the contrary, Karamibekr and Ghorbani [47] argued that opinion mining and sentiment analysis can be defined as an interdisciplinary area situated among the fields of Natural Language Processing (NLP). In addition, studies such as [44], [48] distinguished the concepts of opinion mining and sentiment analysis.

Opinion mining and sentiment analysis have been defined by several researchers from various perspectives. First, the definitions of opinion mining are as follows. According to [44], opinion mining represents a computational study of the opinions, attributes, and evaluations regarding an entity and its aspects, "entity" referring to a product, service or organization, and "aspects" to the attributes of the entity. Guellil and Boukhalfa [49] characterized opinion mining as three main tasks: modeling of opinion, extraction of opinion, and analysis of subjectivity. The modeling of opinion focuses on how an opinion is formalized; the extraction of opinion concerns either a general subject or several subjects or an expression including an opinion or the holder of the opinion; and an opinion is considered to be objective if it contains facts, and subjective if it represents an opinion. Next, the definitions of sentiment analysis are as follows. According to [50], sentiment analysis is an NLP task to identify subjective content, which contains feelings and sentiments, and to classify it as positive, negative or neutral. Liu [44] presented sentimental analysis as a mechanical analysis of the sentiments and opinions of consumers inherent in text

data. Breck and Cardie [51] defined sentiment analysis as a technique for extracting words from a text and analyzing opinions, sentiments, and subjectivity expressed therein in order to help companies understand their customers and to enable them to provide better service thereby.

From these definitions of opinion mining and sentiment analysis, we can see that whereas opinion mining and sentiment analysis are quite similar in the sense of the derivation of consumers' emotions and opinions, opinion mining deals with the broad opinions of consumers including attributes, entities, and aspects, and sentiment analysis focuses on sentiments in their positivity or negativity within consumers' opinions. We note that for this reason, sentiment analysis is more suitable for our proposed method than is opinion mining.

Sentiment analysis can be classified, according to the learning approach, into the following three categories: the machine-learning approach, the lexicon-based approach, and the hybrid approach that combines the first two approaches [52]. The machine-learning approach mainly consists of supervised learning and unsupervised learning. Supervised learning is a major part of machine-learning techniques used for the purpose of sentiment analysis. The most often employed methods for supervised learning are based on vector machines SVM [53], [54], Bayesian network [41], and maximum entropy [55]. The unsupervised learning approaches are based on sentiment lexicons such as dictionaries or corpuses derived from syntactic models [56]. The most representative unsupervised approach is the classification of subjective contents such as sentiments into groups of synonyms and comparing them with a sentiment dictionary, which is composed of a corpus consisting of a positive corpus and a negative corpus. The lexicon-based approach mainly focuses on determining the sentiment lexicon, which is a collection of words in which every word contains a sentiment score that points to the positive, negative or neutral nature of the texts or sentences to be analyzed. The lexicon-based approach is mainly divided into two parts: the dictionary-based approach and the corpus-based approach. The dictionary-based approach proceeds by determining the opinion seeds from reviews and searching the dictionary for antonyms and synonyms, which are added to the list of opinion words taken from reviews. The corpus-based approach constructs a list of seed opinion words and expands it using a large corpus of texts from a single domain. The hybrid approach combines the machine-learning and lexicon-based approaches so as to increase overall performance [57].

Sentiment analysis can be carried out on three different levels according to the analysis depth and type: the document level, sentence level, and entity level [44]. The document level classifies the feelings expressed by reviewers in all text documents as positive, negative or neutral [58]. The sentence level first determines the subjectivity of the sentence, and then determines the polarity (positive, negative or neutral) of subjective sentences [47]. Unlike the document level or sentence level, the entity level can determine the entities or

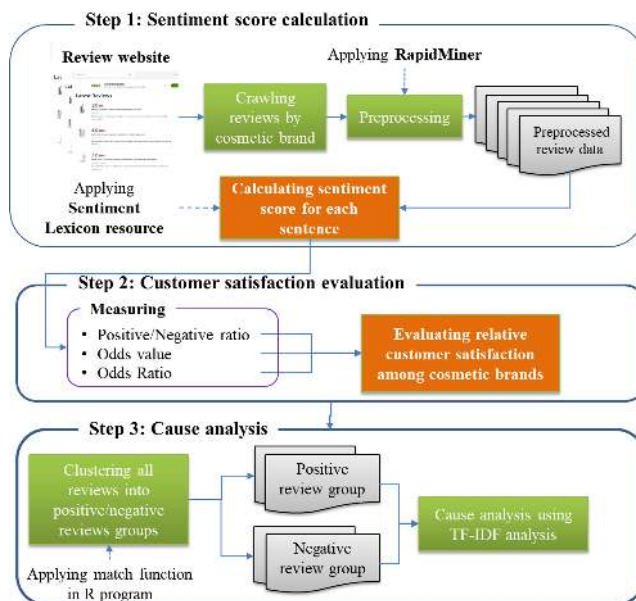


FIGURE 1. Proposed framework.

objects of the emotions and opinions in the text. Even though the overall sentiment of an author's opinion in the document or sentence level is positive, the sentiment for a particular entity or object in the document or sentence can be negative. The most significant feature of the entity level is that it can analyze the particular feeling of an entity or object in the text separately. However, the entity level does not have any clear standard for separating the entity or object in the text, and if the entity or object can be separated, it may still be difficult to grasp the overall sentiment in the document and sentence due to the detailed analysis of entity feeling [58].

III. PROPOSED METHOD: CUSTOMER SATISFACTION EVALUATION

Fig. 1 shows the framework of the proposed method. The framework consists of three steps: sentiment score calculation, customer satisfaction evaluation, and cause analysis. As mentioned in Section I, aim of this study is to introduce a way of evaluating relative customer satisfaction among cosmetics brands through online-reviews analysis applying the well-known text-mining technique and statistical data analysis. Thus, the sentiment score calculation step is the preliminary part, and the customer satisfaction evaluation and cause analysis steps form the backbone of the framework.

Step 1 crawls and collects a corpus of a document (this paper refers to the corpus of the document as reviews) from a cosmetics website, and then calculates the sentiment score to judge the reviews as positive, negative, or neutral through sentiment analysis (Section III.A). Step 2 calculates the satisfaction measures (positive/negative ratio, odds value, and odds ratio), and evaluates relative customer satisfaction for cosmetics brands (Section III.B). Step 3 identifies objects or intentions that cause positive and negative opinions through

integration of cluster analysis and Term Frequency-Inverse Document Frequency (TF-IDF) analysis (Section III.C).

A. SENTIMENT SCORE CALCULATION

Bhuta *et al.* [55] presented a general process of sentiment analysis that consists of four steps: data extraction, pre-processing, data analysis, and identification of knowledge. By extending the general process of sentiment analysis, this step starts with crawling consumers' reviews for each cosmetics brand from an online website (hereafter, "cosmetics brand" and "brand" are used interchangeably). The proposed method was developed by the R programming language, and especially the 'rvest' package provided by the R programming language is utilized in crawling and building a database of reviews. The R programming language is open-source software developed by the R Development Core Team of the R Foundation for statistical calculations (see [59] for more details on R programming).

Since the reviews crawled from online are mostly made up of textual data, NLP needs to be performed for refinement purposes. Xiang *et al.* [60] noted that in order to secure the validity of the text data analysis results, it is necessary to refine the text data by removing unnecessary words and synchronizing similar-meaning words through NLP, a branch of artificial intelligence that deals with the interaction between computers and humans using natural language [61]. The ultimate objective of NLP is to read, decipher, understand, and make sense of human languages in a manner that is valuable. Most NLP techniques rely on machine learning to derive meaning from human languages. The proposed method utilizes RapidMiner software, which is a computer software used in data science, for NLP of review data. RapidMiner software provides features to take advantage of data pre-processing, machine-learning, deep-learning, text-mining, and predictive analysis [62]. In general, online reviews include various letters and symbols that are case-insensitive, including special symbols, or abbreviations. So, we remove numbers, special characters, and symbols, and change uppercase letters to lowercase in the reviews to reduce errors and improve the accuracy of the analysis. Ramadan *et al.* [63] argued that words with the same grammatical meaning and different words with the same meaning must be unified into one word to increase the accuracy of the analysis. So, we unify words that have the same meaning but different expressions into one word (e.g. {*Checkout, Check-Out, Check out*} → *Check-out*).

Each brand generally has multiple reviews, and a single review may contain multiple sentences on the same subject, and a single sentence also can include multiple opinions in the same entities. Thus, for a more fine-grained view of the different opinions as well as for derivation of the various feelings from reviews, this step sets the level of sentiment analysis at the sentence level. Accordingly, the sentiment score is calculated for each sentence, and at least one sentiment score has to be assigned to one sentence. For calculating the sentiment score at the sentence level, the extracted reviews

have to be rearranged in the sentence unit. If a sentence is made up of homogeneous feelings or emotions, it may be ideal to calculate the sentiment score, because it only has to be decided whether the sentence is positive or negative. But if a sentence is made up of heterogeneous emotions such as in a sentimental inversion, it is broken up into clauses where each clause contains a homogeneous opinion. To aid understanding, consider two sentences: '*Staffs are friendly*', '*Staffs are friendly, but facilities are poor.*' The first sentence has a single and homogeneous emotion, and as we know that '*friendly*' is a positive word, we can probably regard this sentence as positive. The second sentence consists of two opposing clauses starting from the reversal connector '*but*', and when we know that '*poor*' is a negative word, this sentence has a heterogeneous emotion; the former can probably be regarded as positive, but the latter can probably be regarded as negative. And although the second sentence is likely to be neutral when we apply the most common sentiment score measurement, it has to be divided into two-sentences "*Staffs are friendly*" and "*facilities are poor*" to further refine the sentence score calculation. Note that a set of linguistic connectors (*and, or, neither-nor, either-or, etc*) and reversal connectors (*but, however, nevertheless, etc*) are used to split a sentence when it includes two opposing clauses.

This step utilizes a common sentiment analysis algorithm [64] that determines the polarity of each sentiment expression based on a sentiment lexicon resource. As mentioned in [65], the sentiment lexicon resource-based sentiment algorithm is the most commonly applied to the various domains because of its relatively simple and easy implementation, and the sentiment lexicon resource is the most crucial for the sentiment analysis algorithm. There are three options of sentiment lexicon resource: the manual approach, by which the lexicon is coded by human hand, the dictionary-based approach, in which a set of seed words is expanded by utilizing resources such as WordNet [66], and the corpus-based approach, in which a set of seed words is expanded by using a large corpus of documents from a single domain. The manual approach is, in general, not feasible, as each domain requires its own lexicon. Since the present research did not limit the subject to a specific domain, the manual approach could not be considered, but the dictionary-based approach was considered for the sentiment lexicon resource. Although the dictionary-based approach has the disadvantage of domain-independence and hence does not capture the specific peculiarities of any specific domain, several studies have reported more advanced sentiment lexicon resources [67], [68]. Thus, the current research utilized these advanced sentiment lexicon resources (hereafter the advanced sentiment lexicon resource is referred to as the sentiment dictionary) in measuring sentiment score.

The sentiment dictionary-based sentiment algorithm is very simple. It detects the frequency of positive and negative words in a sentence by comparing them with the positive and negative lexicons in the sentiment dictionary, and then calculates the sentiment score for each sentence. As a result,

TABLE 1. Review sample for cosmetics brand ‘A’.

Brand	Reviews
‘A’ cosmetic	“My friend recommended me to use this cosmetic. The package is quite good. The price seems reasonable, but the quality is a bit unsatisfactory.”

each sentence is assigned a sentiment score, and the opinion of each sentence is determined as positive, negative or neutral. Let Pe and Ne be a set of positive and negative lexicons in the sentiment dictionary, respectively, and S_i be the i -th sentence in reviews of a certain cosmetics brand. The sentiment score of S_i (SS_i) is calculated as $SS_i = w_i^p - w_i^n$, where w_i^p is the number of positive words found in S_i relative to Pe , and w_i^n is the number of negative words found in S_i relative to Ne . If $SS_i > 1$ entails, it is regarded as positive, if the opposite, it is regarded as negative, whereas otherwise, S_i is neutral. To aid understanding, consider the following sample review from a certain cosmetics brand as listed in Table 1. Although this review consists of three completed sentences, we can identify it as four sentences, namely {‘My friend recommended me to use this cosmetic’, ‘The package is quite good’, ‘The price seems reasonable’, ‘but the quality is a bit unsatisfactory’}, because the last sentence contains two heterogeneous feelings, which are connected by the reversal connector ‘but’. Assume that the sets {‘good’ and ‘reasonable’} and {‘unsatisfactory’} are counted as positive and negative lexicons in the sentiment dictionary, respectively. The sentiment score of the first sentence is regarded as neutral with a sentiment score of 0 (there are no positive or negative words), the second and third sentences are regarded as positive with a sentiment score of 1, and the fourth sentence is regarded as negative with a sentiment score of -1 . As a result, this review contains of two positive sentences, one negative sentence, and one neutral sentence.

B. CUSTOMER SATISFACTION EVALUATION

This step calculates sentiment measures for each cosmetics brand. We note that while the sentiment score is calculated for each sentence, the sentiment measures are calculated for each cosmetics brand. Note that the odds value and odds ratio are calculated for the sentiment measures. In order to calculate the odds value, positive, negative and neutral ratios are measured. Let Ps_r , Ns_r , and Es_r be the aggregation of sentences regarded as positive, negative and neutral in the r -th cosmetics brand, respectively. The positive ratio (PP_r), negative ratio (NP_r), and neutral ratio (EP_r) of the r -th cosmetics brand are measured by Eq. 1-3.

$$PP_r = \frac{Ps_r}{(Ps_r + Ns_r + Es_r)} \times 100 \tag{1}$$

$$NP_r = \frac{Ns_r}{(Ps_r + Ns_r + Es_r)} \times 100 \tag{2}$$

$$EP_r = \frac{Es_r}{(Ps_r + Ns_r + Es_r)} \times 100 \tag{3}$$

Here, the condition of $PP_r + NP_r + EP_r = 100\%$ is satisfied for all cosmetics brands. That is, PP_r means the probability

of having positive opinions in all sentences (the sum of sentences regarded as positive, negative, and neutral) within the r -th brand. However, more precisely, PP_r is not a relative value of positive opinions compared with the others (the sum of negative and neutral) in the r -th brand.

In order to measure the relative degree of positive opinions compared with the other opinions on a brand, the odds value [69] is utilized. We note that the odds value represents the level of customer satisfaction for positive opinions defined in Eq.4.

$$\Omega_r^p = \frac{PP_r}{(100 - PP_r)} \tag{4}$$

where Ω_r^p indicates the relative proportion of having positive opinions relative to the remaining opinions within the r -th brand. A higher Ω_r^p indicates a higher proportion of positive opinions among all, and thus represents higher customer satisfaction with the r -th brand. In particular, when $\Omega_r^p > 1$, the positive opinions on the r -th brand are superior to the sum of the others (negative and neutral), and represent that customer satisfaction with the r -th brand is relatively higher than dissatisfaction or other opinions. An odds value corresponds exactly to a probability value of having positive opinions in the brand, according to its formula defined in Eq. 5.

$$PP_r = \frac{\Omega_r^p}{(100 + \Omega_r^p)} \tag{5}$$

This one-to-one correspondence helps cosmetics managers interpret customer satisfaction based on the number of positive opinions or vice versa.

For better understanding, consider two cosmetics brands A and B. Suppose that the aggregation of positive, negative, and neutral sentences on brands A and B are calculated as (800, 700, 800) and (1450, 600, 500), respectively, when sentiment analysis applied. The positive, negative, and neutral ratios for brands A and B are measured as (34.78, 30.44, 34.78) and (56.86, 23.53, 19.61), respectively, and the sum of all ratios in each brand is 100. Subsequently, the odds values of positivity for brands A and B are measured as 0.53 ($=34.78/(100-34.78)$) and 1.32($=56.86/(100-56.86)$), respectively. In more detail regarding brand A, it can be deduced that the probability of having positive opinions is 0.53 times that for the sum of the other opinions. In addition, the customer satisfaction with brand B is superior to that with brand A, because the odds value of brand B is higher than that of brand A.

Regarding the relative association between cosmetics brands, the odds ratio, which is a pairwise comparison of odds values for brands, would provide useful implications. The odds ratio determines the difference in customer satisfactions between several brands. The relative association between the j -th and j^* -th brands is represented as the odds ratio in Eq. 6.

$$\theta_{(j,j^*)} = \frac{\Omega_j^p}{\Omega_{j^*}^p} = \frac{\frac{PP_j}{(100-PP_j)}}{\frac{PP_{j^*}}{(100-PP_{j^*})}} \tag{6}$$

This means that the likelihood of having better customer satisfaction with the j -th brand is $\theta_{(j,j^*)}$ times higher than

TABLE 2. Contingency table of odds ratios for two sample brands.

j^* \ j	A	B
A	$\theta_{(1,1)} = 1.00$	$\theta_{(2,1)} = 2.49$
B	$\theta_{(1,2)} = 0.40$	$\theta_{(2,2)} = 1.00$
AVG	$(\theta_{(1,1)} + \theta_{(1,2)})/2 = 0.70$	$(\theta_{(2,1)} + \theta_{(2,2)})/2 = 1.75$

that for the j^* -th brand. When the value of $\theta_{(j,j^*)}$ is further from 1, the greater the degree of association with customer satisfaction that is found for the j -th brand. In other words, when $\theta_{(j,j^*)} > 1$, the customer satisfaction with the j -th brand has a higher value than that with the j^* -th brand. Considering sample cosmetics brands A and B as discussed above, we can draw a contingency table as shown in Table 2, through pairwise comparison between the two brands. Here, AVG refers to the average odds ratio of the j -th brand obtained through pairwise comparison with all of the other odds ratios of brands, and this value is also used as the customer satisfaction of a brand. If the AVG value of brand A is greater than that of brand B, it may be determined that the relative customer satisfaction with brand A is higher than that with brand B. In other words, we can say that the likelihood of having better customer satisfaction with brand B is 2.49 times higher than that with brand A, and that consequentially, the relative customer satisfaction with brand B is higher than that with brand A.

C. CAUSE ANALYSIS

This step analyzes, through the importance analysis of words, the objects or intentions that cause positive and negative opinions on each brand. To this end, all of the reviews for each brand are categorized into two groups (see ‘‘Procedure’’ below), where the first group is classified as sentences containing positive lexicons and the second group is classified as sentences containing negative lexicons. In categorizing the reviews, this study used the match function of R programming. As a preliminary step for categorization, the sentiment dictionary is divided into two dictionaries, the first containing only positive lexicons and the second containing only negative lexicons (hereafter, the former will be referred to as the positive sentiment dictionary and the latter as the negative sentiment dictionary). Then, by comparing the reviews in each brand with the positive sentiment dictionary (or negative sentiment dictionary), the sentences containing only positive (or negative) lexicons are identified and classified into separated groups (hereafter, it will be referred to as the positive review group [or negative review group]). This procedure is repeated for all brands, and consequently, the reviews are divided into a positive review group and a negative review group for each brand. Consider the sample review in Table 1 again. When we know that the sets of {‘good’ and ‘reasonable’} and {‘unsatisfactory’} exist in the positive sentiment dictionary and the negative sentiment dictionary, respectively, the second and third sentences will be classified into the positive review group, whereas the fourth sentence

Procedure Categorizing of Reviews Into Positive Review Group and Negative Review Group by Each Brand

Indexes, Set and Parameters:

- i : index of sentences in a review ($i = 1, 2, \dots, I$)
- j : index of brand ($j = 1, 2, \dots, J$)
- Pe : a set of positive lexicons in a positive sentiment dictionary
- Ne : a set of negative lexicons in a negative sentiment dictionary
- S_i : i -th sentence in a review
- w_i^p : number of positive words in S_i
- w_i^n : number of negative words in S_i

Method:

```

[preliminary]Divide sentiment dictionary into  $Pe$  and  $Ne$ 
While  $j \leq J$  do {
  While  $i \leq I$  do {
    Apply the match function to determine if  $Pe$ (or  $Ne$ )
    exists in  $S_i$ 
    Assign the number of positive(or negative) words in  $S_i$ 
    to  $w_i^p$  (or  $w_i^n$ )
    Classify  $S_i$  into the positive (or negative) review group
    if the sum of  $w_i^p$  (or  $w_i^n$ ) is greater than 0
  }
}
    
```

will be classified into the negative review group. Here, the first sentence will not be classified into either group, because no positive or negative lexicons exist in it. If a sentence includes both positive and negative words, it can be classified simultaneously into both the positive and negative review groups.

The TF-IDF method [70] is applied to interpret the determinants of the positive and negative opinions for each brand. TF-IDF is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus, and is used as weighting factor in searches for information retrieval, text mining, and user modeling [71]. In TF-IDF, if a word has a high numerical score in a specific review, the word may be recognized as important in that review.

Though TF-IDF analysis, the words having high numerical scores are calculated for each brand and sorted in descending order. However, the words identified as important can include sentiment lexicons existing in either the positive or negative sentiment dictionary; thus, if the sentiment lexicons are included in the identified words, those lexicons have to be removed. Finally, the remaining words are regarded as most important and frequently mentioned in the positive or negative review group as well as the determinants of positive or negative opinions. Even though various text-mining techniques, such as correlation analysis, frequency analysis and implicit feature extraction [72], can be used for more specific cause analysis, the aim of this study was not to suggest the better or higher-accuracy cause analysis but rather to propose a framework for evaluating customer satisfaction

that includes cause analysis to help brand managers with relatively lower customer satisfaction to improve their results by finding the relevant problems and causes. Thus, this study was limited to the use of TF-IDF analysis for the cause analysis.

IV. EXPERIMENT AND DISCUSSION

To illustrate the efficacy of the proposed approach, an empirical case study applying it to the global top cosmetics brands was conducted. The customer reviews of the brands were crawled from the website of ‘<http://www.Makeupalley.com>’. The structure of the website was analyzed by using the *rvest* package in R programming, while information on the title and content of each of the review comments was crawled and stored into the database temporarily. We crawled the reviews of the top 50 global cosmetics brands published in 2018 by the Brand Finance consulting firm of the UK, and ultimately 26 brands having more than 300 review comments were selected as the final evaluation targets to ensure the reliability of the analysis results.

As shown in the 3rd column of Table 3, a total of 35,617 reviews were crawled for the 26 brands. NLP was performed on the crawled reviews using RapidMiner software. The crawled data was cleaned by removing special characters (e.g., %, !, @, *, #, etc.) and unclarified numbers, by making consistent same-but-different expressions with errata (e.g., “checked-out”, “chekc out”, and “chacking-out”) into one (e.g., “check-out”), and by identifying sentences via characters and tabs (e.g., !, ?, ~, etc.). After breaking up the sentences having heterogeneous opinions (e.g., as in sentimental inversion) into clauses and making those clauses individual sentences, finally, a total of 110,091 sentences were extracted as shown in the 4th column of Table 3. Opinion Lexicon [44] for the sentiment dictionary was used to calculate the sentiment scores of the sentences. Table 4 shows the statistics on the numbers of positive, negative, and neutral sentences derived by the sentiment analysis for the 26 brands. On average, a total of 1,494 sentences were found to be positive, and 414 sentences were found to be negative, indicating that there are more positive opinions than negative. We could see that there were many neutral sentences, because the reviews included much how-to-purchase information as well as explanations of the cosmetics products, all of which is irrelevant to opinion.

Fig. 2 shows the positive, negative, and neutral ratios for each brand. The averages of those ratios for all of the brands were calculated as 35, 10, and 55%, respectively, and most of the brands showed higher positive ratios than negative ones. The brand with the highest positive ratio was Mac*, with a value of 38%, and the lowest positive ratio brands were Neu*, Pan* and Sch*, with a value of 33%. On the other hand, the highest negative ratio was Gil* with a value of 12%, and the lowest negative ratio brand was Sul* with a value of 7%. The brands with the highest neutral ratio were Est* and Pan* with a value of 57%, and the lowest neutral ratio brand was Gil* with a value of 51%. Brands within the top 20%

TABLE 3. Extraction and preprocessing results of reviews for 26 global cosmetics brands.

N0	Brands	Reviews	Sentences
1	Ave*	1,356	4,406
2	Ben*	1,429	4,448
3	Cha*	1,413	4,626
4	Cla*	1,664	4,609
5	Cli*	1,337	4,574
6	Dio*	1,579	4,692
7	Dov*	1,455	4,382
8	Est*	1,536	4,552
9	Gar*	1,368	4,303
10	Gil*	1,265	4,231
11	Gue*	1,539	4,776
12	Hea*	762	2,408
13	Inn*	1,245	3,207
14	Joh*	1,475	4,370
15	Lan*	1,489	4,591
16	L'Oc*	1,474	4,518
17	L'Or*	1,309	4,517
18	Mac*	1,369	4,528
19	May*	1,334	4,397
20	Neu*	1,394	4,591
21	Niv*	1,517	4,351
22	Ola*	1,434	4,536
23	Pan*	1,373	4,461
24	Sch*	1,476	4,485
25	Shi*	1,651	4,600
26	Sul*	374	932
Total		35,617	110,091

TABLE 4. Statistical results on positive/negative/neutral sentences.

	# of positive sentences	# of negative sentences	# of neutral sentences
Max	1,783	505	2,612
Min	268	53	411
Avg	1,494	414	2,319
Std	311.2	95.1	476.3

of the positive and negative ratios were {Mac*, Sul*, Gue*, Gil*, Dio*, Ben*} and {Gil*, Gar*, Hea*, Neu*, Shi*, L'Oc*}, respectively. Especially, Gil* was found to be within the top 20% for both the positive and negative ratios.

The results of the analysis so far are commonly performed in the sentimental analysis of online reviews. Hereafter, we show how to evaluate the customer satisfaction of cosmetics brands from a probabilistic point of view using the results of the sentiment analysis, and provide any interpretation of the determinants of customer opinions on products and services that have not been performed in the previous studies.

The positive odds values for the evaluation of customer satisfaction by each brand are summarized in Fig.3. The brand with the highest positive odds value was Mac*, with a value of 0.61, which indicates that the positive opinions in those reviews were 0.61 times more positive than the other opinions. The brands with positive odds values within the top 20% were Mac*, Ben*, Dio*, Gil*, Gue*, and Sul*, and these brands could be interpreted as having positive feedback from the consumers and high customer satisfaction. On the other hand, brands with positive odds values

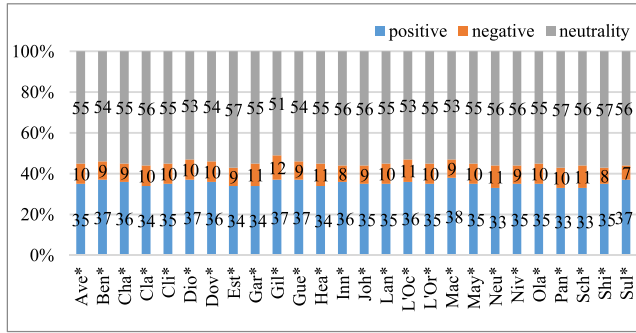


FIGURE 2. Results of positive/negative/neutral ratios for each cosmetics brand.

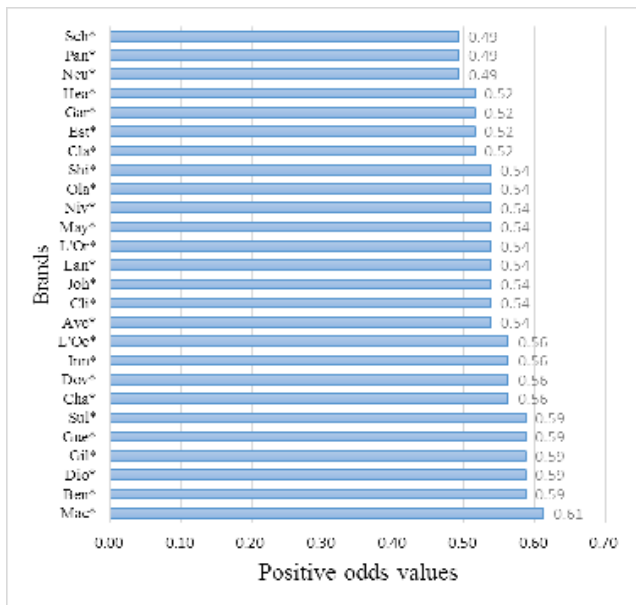


FIGURE 3. Positive odds values from 26 cosmetics brands.

within the bottom 20% were Neu*, Pan*, Sch*, Cla*, Est*, and Gar*, and those brands could be interpreted as inspiring somewhat lower customer satisfaction. Specifically, Mac*'s customer satisfaction was evaluated as the best, and it was regarded as frontier for the relative evaluation with other brands. Ben*, Dio*, Gil*, and Sul* were classified as the second highest customer satisfaction group, and Cha*, Dov*, Inn*, and L'Oc* were classified as the third highest customer satisfaction group, respectively. Consequentially, the Neu*, Pan*, and Sch* brands were evaluated as having the lowest customer satisfaction.

To evaluate the relative associations between brands, the odds ratios were measured by pairwise comparison of the top 20% brands with the bottom 20% brands, and the results are shown in Table 5. Regarding the relative association between Ben* and Cla*, we could say that the likelihood of having better customer satisfaction for Ben* was evaluated to be 1.14 times higher than that for Cla*. Conversely, in terms of Cla*, the likelihood of having better customer satisfaction for Cla* was 0.88 (= 1/1.14) times higher than that for Mac*. That

TABLE 5. Relative associations between top 20% brands and bottom 20% brands.

	Ben*	Dio*	Gil*	Gue*	Mac*	Sul*
Cla*	1.14	1.14	1.14	1.14	1.19	1.14
Est*	1.14	1.14	1.14	1.14	1.19	1.14
Gar*	1.14	1.14	1.14	1.14	1.19	1.14
Neu*	1.19	1.19	1.19	1.19	1.24	1.19
Pan*	1.19	1.19	1.19	1.19	1.24	1.19
Sch*	1.19	1.19	1.19	1.19	1.24	1.19

is, the level of customer satisfaction of Cla* was 88% compared to that for Ben*. From the relative association between Neu* and Ben*, the likelihood of having better customer satisfaction for Ben* was evaluated to be 1.19 times higher than that for Cla*, and the level of customer satisfaction of Neu* was 84% (=1/1.19) compared to that for Cla*. Through the comparison of brands having relatively low customer satisfaction with the frontier brand Mac* (which had the best customer satisfaction, and was regarded as frontier), the levels of customer satisfaction of brands were evaluated as shown in Fig.4. Regarding the relative association between brands Mac* and Ben*, Dio*, Gil*, Gue*, and Sul* (which had the second highest customer satisfaction), we could identify that the level of customer satisfaction of Ben*, Dio*, Gil*, Gue*, and Sul* was 96% of compared to that for Mac*. In the same manner, the level of customer satisfaction of Neu*, Pan*, and Sch* (which had the worst customer satisfaction), was 80% of compared to that for Mac*.

For brands with relatively low customer satisfaction, it is important to analyze what caused such poor results. Cause analysis as a means of enhancing a brand's customer satisfaction has become an essential tool for customer-friendly service provision and service quality improvement. Opinion Lexicon, which was applied for the sentiment dictionary, was divided into a positive sentiment dictionary and a negative sentiment dictionary. As a result, 2,006 positive lexicons were included in the positive sentiment dictionary, and 4,783 negative lexicons were included in the negative sentiment dictionary. In categorizing the reviews of each brand into a positive review group and a negative review group, the average numbers of sentences in the groups were 1,532 and 477, respectively. For a more detail, Fig. 5 shows the result of number of sentences in the positive and negative groups. The averages of the sentences in the positive and negative review groups were found to be somewhat higher than those of the positive and negative sentences in Table 4. This was due to the fact that several sentences containing both positive and negative words could exist simultaneously in the positive and negative review groups.

Cause analysis was performed for the brands having positive odds values within the top 20% and bottom 20%, specifically brands Mac*, Dio* and Sul* among the top 20%, and brands Neu*, Pan* and Ach* among the bottom 20%. Through TF-IDF analysis, six words having high numerical scores were extracted from the positive review group and negative review group by each brand. Table 6 shows the results of the TF-IDF analysis for the brands Mac*, Dio* and Sul*.

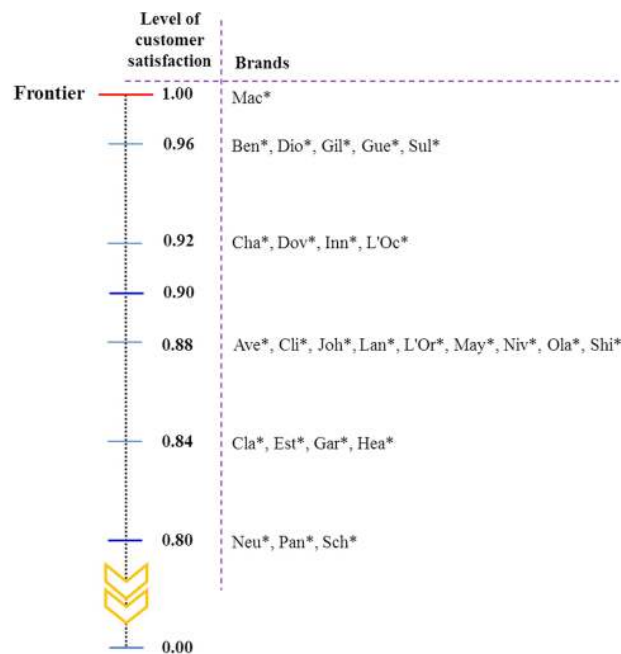


FIGURE 4. Relative level of customer satisfaction.

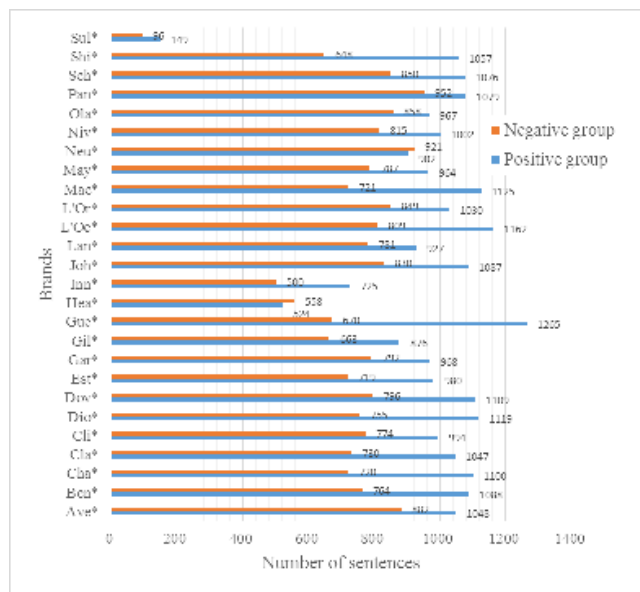


FIGURE 5. Number of sentences in the positive and negative groups.

In the positive review group, we could see that customers frequently mentioned 'skin' as important, and accordingly, 'skin' could be considered to be the main determinant of the positive opinions of those brands. Concretely, although the rank of 'skin' was the same in Mac* and Sul*, the numerical score of 'skin' was higher in Mac* than in Sul*. This could be attributed, simply, to the greater number of reviews for Mac* than for Sul*. Interestingly, in the negative review group, costumers also frequently mentioned 'skin' as important for Mac*, Dio* and Sul*. In other words, 'skin' was identified as the most important word in both the positive and negative

TABLE 6. Results of TF-IDF analysis for top 20% brands in positive and negative review groups.

Positive review group						
No	Mac*		Dio*		Sul*	
	words	score	words	score	words	score
1	skin	47.34	foundation	46.43	skin	9.31
2	color	44.45	skin	43.41	cream	8.14
3	shade	36.93	scent	35.42	smell	6.31
4	lipstick	36.13	packaging	32.87	mask	5.05
5	foundation	27.48	color	27.96	packaging	4.49
6	formula	23.86	glow	26.36	lines	4.27

Negative review group						
No	Mac*		Dio*		Sul*	
	words	score	words	score	words	score
1	matte	40.17	matte	41.97	smell	9.20
2	skin	22.52	smell	35.87	skin	5.87
3	hard	21.70	foundation	30.05	pricey	5.36
4	color	21.65	skin	28.81	herbal	4.95
5	shimmer	21.18	mascara	17.74	dry	3.93
6	dry	20.43	pricey	16.86	worth	3.59

review groups. However, in the numerical scores, there was a slight difference between the positive and negative review groups in that the overall numerical score for 'skin' was higher in the positive review group than in the negative review group. This could be as attributed to the greater number of positive opinions on 'skin' mentioned by customers for Mac*, Dio* and Sul*, even though customers also expressed negative opinions. Overall, customers mentioned to the words 'foundation', 'skin' and 'color' to convey their positive opinions on both Mac* and Dio*; thus, those words could be considered to be the main aspects causing positive opinions of those brands. Specifically, 'skin' ranked higher for Mac* than for Dio*, while 'foundation' ranked higher for Dio* than for Mac*.

Table 7 shows the results of the TF-IDF analysis for brands Neu*, Pan* and Sch* within the bottom 20%. When we looked at the negative review group first, customers mentioned 'smell' and 'greasy' as important words that cause negative opinions for all brands. In Neu*, customers were found to express heavily negative opinions in the forms of 'skin', 'smell', 'face', 'greasy', 'matt', and 'dry', and these words can be interpreted as key determinants causing customer dissatisfaction. Although 'skin' is important in the positive review group, the numerical score is higher in the negative review group than in the positive review group. In other words, we could say that although customers expressed both positive and negative opinions as to 'skin' in Neu*, it causes more negative opinions than positive opinions. Alternatively, 'skin' can be interpreted as a factor that clearly divides likes and dislikes according to customers. At a more strategic level, brand Neu* needs a more in-depth look at the key drivers of negative feedback on skin products along with a strategy for improved customer satisfaction.

We could determine that 'hair', 'shampoo', and 'conditioner' in Pan* were mentioned as important in both the positive and negative review groups. In Pan*, the positive opinions on 'hair', 'shampoo', and 'conditioner' were mentioned with somewhat more importance than were the negative opinions,

TABLE 7. Results of TF-IDF analysis for bottom 20% brands in positive and negative review groups.

Positive review group						
No	Neu*		Pan*		Sch*	
	words	score	words	score	words	score
1	skin	50.55	hair	70.55	skin	61.03
2	makeup	30.93	shampoo	50.79	foundation	34.82
3	clean	27.84	smells	43.59	packaging	26.74
4	cleanser	25.97	conditioner	39.09	eye	24.60
5	face	25.84	clean	39.02	face	23.89
6	gentle	24.87	scent	34.68	right	22.99
Negative review group						
No	Neu*		Pan*		Sch*	
	words	score	words	score	words	score
1	skin	54.83	hair	52.04	smell	65.45
2	smell	52.67	smell	49.13	hair	55.28
3	face	34.13	shampoo	34.03	damaged	30.32
4	greasy	33.88	greasy	31.85	greasy	30.07
5	matte	23.86	dry	31.34	shampoo	20.58
6	dry	23.49	conditioner	26.71	stiff	18.58

but ‘smell’ was mentioned with more importance in the negative review group. In other words, Pan* needs a strategy to resolve negative opinions about smells in the shampoo and the conditioner products in order to improve customer satisfaction. In Sch*, we could see that customers expressed the most negative opinions of ‘smell’ and ‘hair’, and that on the contrary, they expressed the most positive opinions of ‘skin’, ‘foundation’, and ‘packaging’. Since the negative opinions were frequently expressed in terms of ‘smell’, ‘hair’, and ‘greasy’, strategies to improve such opinions may be needed for customer satisfaction improvement.

V. CONCLUSION

This study derived a new, online-review-based process for evaluation of relative customer satisfaction with cosmetics brands as well as interpretation of the determinants of reviewers’ positive and negative sentiments. To illustrate the efficacy of the proposed approach, an empirical case study applying it to the global top 26 cosmetics brands was conducted, and the results of which indicated how it can utilize online reviews for relative customer satisfaction evaluation. In addition, we showed the applicability of the proposed approach by examining the main causes or determinants of positive and negative opinions for specific brands.

According to the sentiment analysis on the global 26 cosmetics brands, the average positive, negative and neutral opinions were 35, 10, and 55%, respectively, and we found that the positive opinion ratio was higher than the negative opinion ratio for most brands. Brand Mac* was recognized as having the highest positive opinions with a ratio of 33%, while Neu*, Pan*, and Sch* were recognized as having the lowest positive opinions. On the contrary, Gil* was recognized as having the highest negative opinion with a ratio of 12%, while Sul* was recognized as having the lowest negative opinions. The brand with the highest customer satisfaction was Mac*, and its positive opinions were relatively 0.61 times higher than the aggregation of negative and neutral opinions. Regarding the relative association between highest

and lowest brands in customer satisfaction, the likelihood of having better customer satisfaction with the highest brands was evaluated as 1.24 times higher than that with the lowest brands. In the brands within the top 20% in customer satisfaction, ‘skin’ was identified as the important word common to both the positive and negative review groups. However, the overall numerical score for ‘skin’ was higher in the positive review group than in the negative review group, and we could interpret that as having stemmed from the greater number of positive opinions than negative opinions on ‘skin’.

The proposed approach is expected to be a new way to replace the questionnaire survey method that was widely applied for analyzing customer satisfaction, and used by cosmetics companies to realize or improve satisfaction with their brands that customers evaluate. Furthermore, we hope that it can be used as a source of fundamental data that could be applied to efforts to improve both brand competitiveness and provision of systematic services. Despite its valuable contributions, this study also has several drawbacks. First, sentiment score is very sensitive depending on the accuracy of the sentiment lexicon resource applied in this approach. Although this study applied the general sentiment lexicon resource widely utilized in other studies, a more highly accurate sentiment dictionary will be needed for more sophisticated sentiment analysis. Second, the case of sentiment activation/deactivation (e.g. ‘not good’, etc) was not considered, which implies that the accuracy of the sentiment score may be somewhat less than reported herein. Both of these issues will be addressed in future research.

REFERENCES

- [1] D. Media, *Consumer Reports by Industry*. Seoul, South Korea, DMC Report, 2016.
- [2] N. Archak, A. Ghose, and P. G. Ipeirotis, “Deriving the pricing power of product features by mining consumer reviews,” *Manage. Sci.*, vol. 57, no. 8, pp. 1485–1509, Aug. 2011.
- [3] A. Ghose and P. G. Ipeirotis, “Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics,” *IEEE Trans. Knowl. Data Eng.*, vol. 23, no. 10, pp. 1498–1512, Oct. 2011.
- [4] M. Haddara, J. Hsieh, A. Fagerström, N. Eriksson, and V. Sigurosson, “Exploring customer online reviews for new product development: The case of identifying reinforcers in the cosmetic industry,” *Managerial Decis. Econ.*, vol. 41, no. 2, pp. 250–273, Mar. 2020.
- [5] E. J. Park, “Effects of the direction of online reviews on purchase decision: The role of product type and brand reputation,” *J. Korea Advertising Soc.*, vol. 18, no. 5, pp. 7–22, 2007.
- [6] E. J. Lee and W. S. Shim, “A study on the behavior characteristics of point of purchase, post purchase and trust evaluation of Internet shopping after notes,” *J. Global e-Bus. Assoc.*, vol. 8, no. 3, pp. 155–170, 2007.
- [7] S.-H. Liao, C.-L. Hsieh, and S.-P. Huang, “Mining product maps for new product development,” *Expert Syst. Appl.*, vol. 34, no. 1, pp. 50–62, Jan. 2008.
- [8] N. R. Jung and Y. R. Kim, “The influence of customer’s purchasing attitude for cosmetic by characteristic of SNS on related information and purchase satisfaction,” *Kor J Aesthet Cosmetol*, vol. 12, no. 4, pp. 565–573, 2014.
- [9] The JoongAng Ilbo. (Mar. 25, 2017). *Testimonials Drive the Market*. The JoongAng Ilbo. Accessed: Dec. 20, 2019. [Online]. Available: <https://news.joins.com/article/21404190>
- [10] J. Zhan, H. T. Loh, and Y. Liu, “Gather customer concerns from online product reviews—A text summarization approach,” *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2107–2115, Mar. 2009.
- [11] J. K. Bae and J. Kim, “Product development with data mining techniques: A case on design of digital camera,” *Expert Syst. Appl.*, vol. 38, no. 8, pp. 9274–9280, Aug. 2011.

- [12] M. Bhandari and S. Rodgers, "What does the brand say? Effects of brand feedback to negative eWOM on brand trust and purchase intentions," *Int. J. Advertising*, vol. 37, no. 1, pp. 125–141, 2018.
- [13] K.-N. Lau, K.-H. Lee, and Y. Ho, "Text mining for the hotel industry," *Cornell Hotel Restaurant Admin. Quart.*, vol. 46, no. 3, pp. 344–362, Aug. 2005.
- [14] H. Li, Q. Ye, and R. Law, "Determinants of customer satisfaction in the hotel industry: An application of online review analysis," *Asia Pacific J. Tourism Res.*, vol. 18, no. 7, pp. 784–802, Oct. 2013.
- [15] Y. Ma and S. Yang, "An empirical study of female e-shopper's satisfaction with cosmetic products in China," *Int. J. Bus. Manage.*, vol. 13, no. 3, pp. 211–219, 2018.
- [16] Y. Zhan, K. H. Tan, Y. Li, and Y. K. Tse, "Unlocking the power of big data in new product development," *Ann. Oper. Res.*, vol. 270, nos. 1–2, pp. 577–595, Nov. 2018.
- [17] H. An and I. Lee, "An Investigation of a Sensibility Evaluation Method Using Big Data," *J. Korean Soc. Clothing Textiles*, vol. 40, no. 6, pp. 1034–1044, 2016.
- [18] A. Gordon, *Future Savvy: Identifying Trends to Make Better Decisions, Manage Uncertainty, and Profit From Change*. New York, NY, USA: American Management Association, 2008.
- [19] S. W. Oh and S. H. Jin, "A study on analysis of Internet shopping mall customers' reviews by text mining," *J. Korean Data Anal. Soc.*, vol. 14, no. 1, pp. 125–137, 2012.
- [20] Q. Gan, B. H. Ferns, Y. Yu, and L. Jin, "A text mining and multidimensional sentiment analysis of online restaurant reviews," *J. Qual. Assurance Hospitality Tourism*, vol. 18, no. 4, pp. 465–492, Oct. 2017.
- [21] D. G. Dwayne, P. G. Kevin, and W. B. Stephen, "Generating positive word of mouth communication through customer-employee relationships," *Int. J. Service Ind. Manage.*, vol. 12, no. 1, pp. 5–44, 2001.
- [22] R. A. Peterson and M. C. Merino, "Consumer information search behavior and the Internet," *Psychol. Marketing*, vol. 20, no. 2, pp. 99–121, Feb. 2003.
- [23] T. Hennig-Thurau, K. P. Gwinner, G. Walsh, and D. D. Gremler, "Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet?" *J. Interact. Marketing*, vol. 18, no. 1, pp. 38–52, Jan. 2004.
- [24] J. Yi and W. Niblack, "Sentiment mining in WebFountain," in *Proc. 21st Int. Conf. Data Eng. (ICDE)*, Apr. 2005, pp. 1073–1083.
- [25] T. Y. Lee and E. T. Bradlow, "Automated marketing research using online customer reviews," *J. Marketing Res.*, vol. 48, no. 5, pp. 881–894, Oct. 2011.
- [26] S. Arora, M. Joshi, and C. P. Rosé, "Identifying types of claims in online customer reviews," in *Proc. Hum. Lang. Technol., Annu. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Companion Volume, Short Papers NAACL*, 2009, pp. 37–40.
- [27] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: Sentiment classification using machine learning techniques," in *Proc. ACL Conf. Empirical Methods Natural Lang. Process. - Volume*, vol. 10, 2002, pp. 79–86.
- [28] N. Hu, T. Zhang, B. Gao, and I. Bose, "What do hotel customers complain about? Text analysis using," *Tourism Manage.*, vol. 72, pp. 417–426, Jun. 2019.
- [29] P. Padma and J. Ahn, "Guest satisfaction & dissatisfaction in luxury hotels: An application of big data," *Int. J. Hospitality Manage.*, vol. 84, Jan. 2020, Art. no. 102318.
- [30] D. Kang and Y. Park, "Review-based measurement of customer satisfaction in mobile service: Sentiment analysis and VIKOR approach," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1041–1050, Mar. 2014.
- [31] S. M. Mudambi and D. Schuff, "Research note: What makes a helpful online review? A study of customer reviews on Amazon.com," *MIS Quart.*, vol. 34, no. 1, pp. 185–200, 2010.
- [32] Y. I. Yoon and D. H. Ha, "The effects of source credibility to Internet WOM communication on brand equity: Focused on customers of deluxe hotel," *J. Korean Academic Soc. Hospitality Admin.*, vol. 19, no. 4, pp. 81–97, 2010.
- [33] P. Chatterjee, "Online reviews: Do consumers use them," in *Advances in Consumer Research*, vol. 28, Mary C. Gilly and J. Meyers-Levy, eds. Valdosta, GA, USA: The Association for Consumer Research, 2001, pp. 129–133.
- [34] G. Somprasertsri and P. Lalitrojwong, "Mining feature-opinion in online customer reviews for opinion summarization," *J. Universal Comput. Sci.*, vol. 16, no. 6, pp. 938–955, 2010.
- [35] J.-S. Chiou and C. Cheng, "Should a company have message boards on its Web sites?" *J. Interact. Marketing*, vol. 17, no. 3, pp. 50–61, Jan. 2003.
- [36] E. Clemons, G. Gao, and L. Hitt, "When online reviews meet hyper differentiation: A study of the craft beer industry," *J. Manage. Inf. Syst.*, vol. 23, no. 2, pp. 149–171, 2006.
- [37] A. Ghose and P. G. Ipeiritos, "Designing ranking systems for consumer reviews: The impact of review subjectivity on product sales and review quality," in *Proc. 16th Annu. Workshop Inf. Technol. Syst.*, Chicago IL, USA, 2006, pp. 303–310.
- [38] M. S. I. Malik and A. Hussain, "Helpfulness of product reviews as a function of discrete positive and negative emotions," *Comput. Hum. Behav.*, vol. 73, pp. 290–302, Aug. 2017.
- [39] S. Riaz, M. Fatima, M. Kamran, and M. W. Nisar, "Opinion mining on large scale data using sentiment analysis and k-means clustering," *Cluster Comput.*, vol. 22, no. S3, pp. 7149–7164, May 2019.
- [40] P. Gundecha and H. Liu, "Mining social media: A brief introduction," *Tuts. Oper. Res.*, vol. 1, no. 4, pp. 1–17, 2012.
- [41] P.-W. Liang and B.-R. Dai, "Opinion mining on social media data," in *Proc. IEEE 14th Int. Conf. Mobile Data Manage.*, Jun. 2013, pp. 91–96.
- [42] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intell. Syst.*, vol. 28, no. 2, pp. 15–21, Mar. 2013.
- [43] S. R. El-Beltagy and A. Ali, "Open issues in the sentiment analysis of arabic social media: A case study," in *Proc. 9th Int. Conf. Innov. Inf. Technol. (IIT)*, Mar. 2013, pp. 215–220.
- [44] B. Liu, "Sentiment analysis and opinion mining," *Synthesis Lectures Hum. Lang. Technol.*, vol. 5, no. 1, pp. 1–167, 2012.
- [45] F. Tuveri and M. Angioni, "An opinion mining model for generic domains," in *Distributed Systems and Applications of Information Filtering and Retrieval*. Berlin, Germany: Springer, 2014, pp. 51–64.
- [46] M. Karamibekr and A. A. Ghorbani, "Sentence subjectivity analysis in social domains," in *Proc. IEEE/WIC/ACM Int. Joint Conf. Web Intell. (WI Intell. Agent Technol. (IAT))*, Nov. 2013, pp. 268–275.
- [47] M. Karamibekr and A. A. Ghorbani, "A structure for opinion in social domains," in *Proc. Int. Conf. Social Comput.*, Sep. 2013, pp. 264–271.
- [48] H. Chen and D. Zimbra, "AI and opinion mining," *IEEE Intell. Syst.*, vol. 25, no. 3, pp. 74–80, May 2010.
- [49] I. Guellil and K. Boukhalfa, "Social big data mining: A survey focused on opinion mining and sentiments analysis," in *Proc. 12th Int. Symp. Program. Syst. (ISPS)*, Apr. 2015, pp. 1–10.
- [50] V. K. Singh, R. Piryani, A. Uddin, and P. Waila, "Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification," in *Proc. Int. Multi-Conf. Autom., Comput., Commun., Control Compressed Sens. (iMac4s)*, Mar. 2013, pp. 712–717.
- [51] E. Breck and C. Cardie, "Opinion mining and sentiment analysis," in *The Oxford Handbook of Computational Linguistics*, 2nd ed. 2017.
- [52] X. Wan, "A comparative study of cross-lingual sentiment classification," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. Intell. Agent Technol.*, Dec. 2012, pp. 24–31.
- [53] F. Wogenstein, J. Drescher, D. Reinel, S. Rill, and J. Scheidt, "Evaluation of an algorithm for aspect-based opinion mining using a lexicon-based approach," in *Proc. 2nd Int. Workshop Issues Sentiment Discovery Opinion Mining WISDOM*, 2013, pp. 1–8.
- [54] X. Zhou, X. Tao, J. Yong, and Z. Yang, "Sentiment analysis on tweets for social events," in *Proc. IEEE 17th Int. Conf. Comput. Supported Cooperat. Work Design (CSCWD)*, Jun. 2013, pp. 557–562.
- [55] S. Bhuta, A. Doshi, U. Doshi, and M. Narvekar, "A review of techniques for sentiment analysis of Twitter data," in *Proc. Int. Conf. Issues Challenges Intell. Comput. Techn. (ICICT)*, Feb. 2014, pp. 583–591.
- [56] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Eng. J.*, vol. 5, no. 4, pp. 1093–1113, Dec. 2014.
- [57] A. Mukwazvure and K. P. Supreethi, "A hybrid approach to sentiment analysis of news comments," in *Proc. 4th Int. Conf. Rel., Infocom Technol. Optim. (ICRITO) (Trends Future Directions)*, Sep. 2015, pp. 1–6.
- [58] N. D. Valakunde and M. S. Patwardhan, "Multi-aspect and multi-class based document sentiment analysis of educational data catering accreditation process," in *Proc. Int. Conf. Cloud Ubiquitous Comput. Emerg. Technol.*, Nov. 2013, pp. 188–192.
- [59] R Core Team. (2016). *R, A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Accessed: Dec. 20, 2019. [Online]. Available: <http://www.R-project.org/>
- [60] Z. Xiang, Z. Schwartz, J. H. Gerdes, and M. Uysal, "What can big data and text analytics tell us about hotel guest experience and satisfaction?" *Int. J. Hospitality Manage.*, vol. 44, pp. 120–130, Jan. 2015.

- [61] K. Li, R. G. Dewar, and R. J. Pooley, "Requirements capture in natural language problem statements," Heriot-Watt, Edinburgh, Scotland, Tech. Rep. HW-MACS-TR-002, 2004.
- [62] G. Deutsch. (Mar. 18, 2010). *RapidMiner From Rapid-I at CeBIT 2010*. Accessed: Dec. 20, 2019. [Online]. Available: <http://www.data-mining-blog.com/cloud-mining/rapidminer-cebit-2010/>
- [63] S. Ramadan, H. Ibrahim Baqapuri, E. Roecher, and K. Mathiak, "Process mining of logged gaming behavior," in *Proc. Int. Conf. Process Mining (ICPM)*, Jun. 2019, pp. 57–64.
- [64] X. Ding, B. Liu, and P. S. Yu, "A holistic lexicon-based approach to opinion mining," in *Proc. Int. Conf. Web Search Web Data Mining - WSDM*, 2008, pp. 231–240.
- [65] R. Feldman, "Techniques and applications for sentiment analysis," *Commun. ACM*, vol. 56, no. 4, pp. 82–89, Apr. 2013.
- [66] C. Fellbaum, *WordNet: An Electronic Lexical Database*. Cambridge, MA, USA: MIT Press, 1988.
- [67] E. C. Dragut, C. Yu, P. Sistla, and W. Meng, "Construction of a sentimental word dictionary," in *Proc. 19th ACM Int. Conf. Inf. Knowl. Manage. CIKM*, 2010, pp. 1761–1764.
- [68] W. Peng and D. H. Park, "Generate adjective sentiment dictionary for social media sentiment analysis using constrained nonnegative matrix factorization," in *Proc. 5th Int. AAI Conf. Weblogs Social Media*, Barcelona, Spain, Jul. 2011, pp. 273–280.
- [69] A. Agresti, *Categorical Data Analysis*, 2nd ed. Hoboken, NJ, USA: Wiley, 2002.
- [70] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Inf. Process. Manage.*, vol. 24, no. 5, pp. 513–523, Jan. 1988.
- [71] A. Rajaraman and J. D. Ullman, "Data Mining," in *Mining of Massive Datasets*. New York, NY, USA, Cambridge Univ. Press, 2011, pp. 1–17.
- [72] Y. Zhang and W. Zhu, "Extracting implicit features in online customer reviews for opinion mining," in *Proc. 22nd Int. Conf. World Wide Web WWW Companion*, 2013, pp. 103–104.



JAEHUN PARK received the Ph.D. degree in industrial engineering from Pusan National University, South Korea, in 2012. From 2012 to 2014, he was a Postdoctoral Research Fellow with the Worcester Polytechnic Institute, Worcester, MA, USA. From 2014 to 2017, he was a Senior Researcher with the Defense Agency for Technology and Quality, South Korea. He is currently an Assistant Professor of industrial and quality engineering with the Division of Cosmetic Science and Technology, Daegu Haany University. His research interests include performance evaluation, big data analysis, and smart factory.

• • •