



Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Analyzing the discriminative attributes of products using text mining focused on cosmetic reviews



Kim Sung Guen^a, Kang Juyoung^{b,*}

^a Department of MIS, Business School of Ajou University, San 5, Woncheon-dong, Yeongtong-gu, Suwon 443-749, South Korea

^b Department of e-Business, Business School of Ajou University, San 5, Woncheon-dong, Yeongtong-gu, Suwon 443-749, South Korea

ARTICLE INFO

Keywords:

Text mining
Latent semantic analysis
LSA
Labeled latent Dirichlet allocation
L-LDA
Discriminative attributes of products

ABSTRACT

Consumers evaluate products through online reviews, in addition to sharing their product experiences. Online reviews affect product marketing, and companies use online reviews to investigate consumer attitudes and perceptions of their products. However, when analyzing a review, it is often the case that specific contexts are not taken into consideration and meaningful information is not obtained from the analysis results. This study suggests a methodology for analyzing reviews in the context of comparing two competing products. In addition, by analyzing the discriminative attributes of competing products, we were able to derive more specific information than an overall product analysis. Analyzing the discriminative attributes in the context of comparing competing products provides clarity on analyzing the strengths and weaknesses of competitive products and provides realistic information that can help the company's management activities. Considering this purpose, this study collected a review of the BB Cream product line in the cosmetics field. The analysis was sequentially carried out in three stages. First, we extracted words that represent discriminative attributes by analyzing the percentage difference of words. Second, different attribute words were classified according to the meaning used in the review by using latent semantic analysis. Finally, the polarity of discriminative attribute words was analyzed using Labeled-LDA. This analysis method can be used as a market research method as it can extract more information than a traditional survey or interview method, and can save cost and time through the automation of the program.

1. Introduction

As shown by Zero Moment of Truth (ZMOT), a consumer decision-making model introduced by Google (2011), the moments of reading product information and user reviews online before a consumer directly considers a product are crucial to the consumers' decision making. In fact, according to a Pew Research Center survey, 82% of Americans refer to online reviews when they first purchase a product, and 40% *always* refer to online reviews when purchasing goods. Of Americans respondents surveyed, 46% said online reviews were a great help in convincing them to buy, and 45% attempt to find online reviews using their mobile phones even when they are in physical stores (Smith & Anderson, 2016).

In the past, the moment consumers faced their products was decisive in making their decisions to buy. That has now changed however, as, the purchase decision process is increasingly influenced by online reviews. Many studies have been conducted on the value and usefulness of reviews, and online reviews have been reported to promote product sales (Chen & Xie, 2005; Chevalier & Mayzlin, 2006; Dellarocas, Zhang, & Awad, 2007; Duan, Gu, & Whinston, 2008; Godes & Mayzlin, 2004; Liu, 2006; Resnick &

* Corresponding author.

E-mail address: jygang@ajou.ac.kr (J. Kang).

Zeckhauser, 2002; Zhang & Dellarocas, 2006). As research has shown that product reviews improve sales, many companies are encouraging buyers to write reviews. Therefore, it is no longer significant to study the usefulness of online reviews or their impact on corporate performance. Also, various studies have been conducted to analyze consumers’ emotions and attitudes in their online reviews (Amolik, Jivane, Bhandari, & Venkatesan, 2015; Fang & Zhan, 2015; Narendra et al., 2016; Schouten & Frasincar, 2014; Wahyudi & Kristiyanti, 2016), but this emotional analysis can only identify consumers’ attitudes and emotions toward the product as a whole—it cannot tell the specific causes of the emotions or attitudes. In other words, it is difficult for companies to establish concrete action strategies based on emotional analysis results.

Therefore, in this study, we propose a methodology to extract and analyze the discriminative attributes of competing products by using online reviews to overcome the limitations of existing review studies. The reasons for paying attention to the differentiated attributes of products are as follows. Product differentiation (Chamberlin, 1933) is one of the management strategies of a company. It is common sense that product differentiation determines the competitiveness of products. Product differentiation is achieved by making a difference in the level of the product’s attributes, unless an entirely new product is released through innovation. Therefore, examining consumers’ perceptions of discriminative attributes of a product is similar to obtaining information about the competitive advantage of a product. In addition, if customers’ attitudes are identified at the attribute level of the product, it can be utilized for product improvement or development of new products.

1.1. Research objective

The goal of this study is to collect reviews of two competing products or product groups, extract the discriminative attributes of each product, and analyze consumers’ attitudes toward those attributes. For our purposes, we designed the following three-step analysis procedure. In the first step, we selected two competing products to collect reviews of the products, and we extracted the discriminative attributes of each product in the review. Specifically, we calculated the percentages of the words in the overall reviews for each product, and we extracted the words that showed large differences in the ratios and defined them as “discriminative attribute words,” that is, words that indicate a discriminative attribute; this analysis can identify what consumers perceive to be the distinguishing attributes of competing products. In the second step, the discriminative attribute words were classified based on the degree of similarity, and we classified the attribute words that had high similarity into specific attribute categories; using this analysis, we could identify the context in which words appeared. The third step was to classify the polarities of the discriminative attribute words. We estimated the probability that the discriminative attribute word appeared in the product’s rating (1–5 points). And we classified words with a high probability of appearing in a negative context at 1–2 points, and words with a high probability of appearing in a positive context at 4–5 points, this enabled us to understand consumers’ attitudes and emotions about the products’ attributes. Fig. 1 shows the analysis procedure of this study.

For this study, we combined a variety of machine learning techniques in sequence to design an analytical procedure that met the specific purpose of discriminative attribute analysis. The procedure we propose in this study provides more detailed information than general emotional analysis, and it can be used for product positioning by analyzing the discriminative attributes of each product in a competitive market.

As consumers become more influential, it becomes increasingly important for companies to communicate and interact with them. Customer surveys and interviews require great cost and effort, and it is not possible to identify customer needs in real time. In

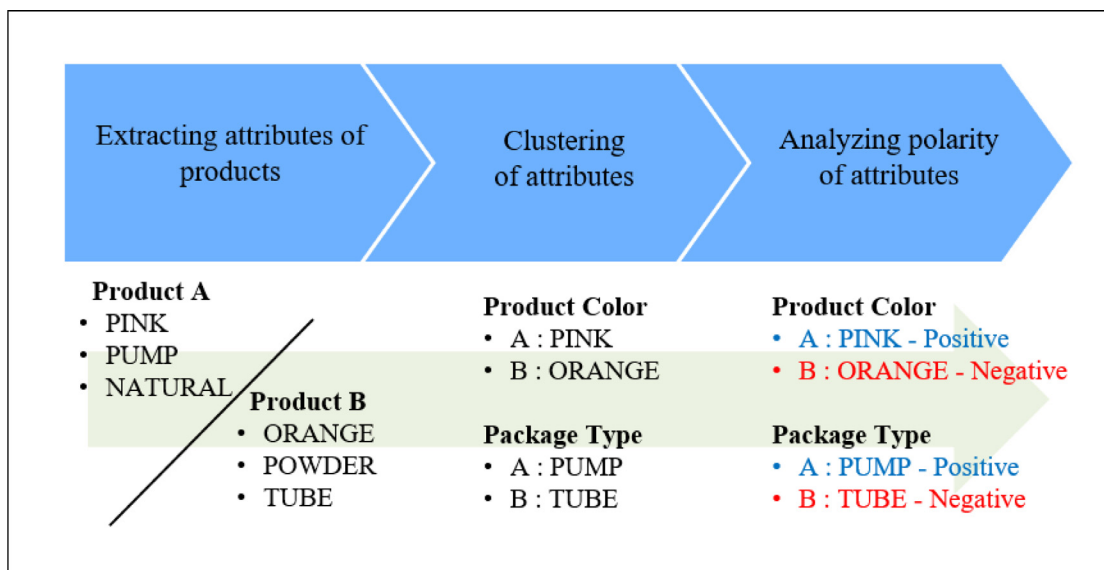


Fig. 1. Research procedure.

Download English Version:

<https://daneshyari.com/en/article/6925916>

Download Persian Version:

<https://daneshyari.com/article/6925916>

[Daneshyari.com](https://daneshyari.com)