HARNESSING THE PREDICTIVE VALUE OF ONLINE WORD-OF-MOUTH FOR IDENTIFYING MARKET SUCCESS OF NEW AUTOMOBILES: INPUT VERSUS OUTPUT WORD-OF-MOUTH PERSPECTIVES

Jaewon Choi[1], Hong Joo Lee[2], Joon Yeon Choeh[3]

[1] *Soonchunhyang University, Department of Business Administration, Republic of Korea, ORCID: 0000-0001-9795-553X, jaewonchoi@sch.ac.kr;*
[2] *Catholic University of Korea, Department of Business Administration, Republic of Korea, hongjoo@catholic.ac.kr;*
[3] *Sejong University, Department of Software, Republic of Korea, ORCID: 0000-0002-6604-5944, zoon@sejong.edu (corresponding author).*

Abstract: *The automotive industry evaluates various success factors to achieve competitive advantage in selling products. Existing studies have predicted the success of newly launched automobiles based on an economic perspective. However, factors such as dynamic changes in consumer preferences and the emergence of numerous automobile brands pose difficulty in understanding product quality. This study proposes a method of understanding the automotive market using text mining techniques and online user opinions for newly launched cars. By analyzing customer experiences and expectations through their opinions, we can anticipate automobile demand in the market more easily. The proposed method is based on online reviews from an online portal for automobiles. Based on a literature review, this study presents a framework for analyzing input versus output word-of-mouth (WOM). It also integrates the success factors from existing automobile studies and derives functional categories and relevant keywords. The analysis identifies differences in consumer-interest factors that lead to short-term success or normal results in automobile sales. In addition, it confirms that the elements of WOM produces varying results depending on the timing these are employed in relation to the product launch (i.e., before or after a product's launch). It revealed which dimensions of automobile characteristics are important factors in identifying sales volume and market share for specific types and brands of automobile models. The results of this study provide theoretical advantage in predicting market success in the automobile industry. In addition, the study derives practical insights into characteristics of classification information for market forecasts in the automotive industry. The paper provides empirical insights about how input WOM and output WOM which are analyzed differently can have predictive power in forecasting market share and sales volume for automobiles.*

Keywords: *Market success, word-of-mouth, text mining, LDA, purchase decision.*

JEL Classification: *C55, C80, D91, L62.*

APA Style Citation: Choi, J., Lee, H. J., & Choeh, J. Y. (2022). Harnessing the Predictive Value of Online Word-of-mouth for Identifying Market Success of New Automobiles: Input versus Output Word-of-mouth Perspectives. *E&M Economics and Management, 25*(2), 183–201. https://doi.org/10.15240/tul/001/2022-2-012

Introduction

Social interactions such as user discussions, reviews, and replies on websites have become a major communication channel for customers to discuss their opinions on various topics (Appel et al., 2020). Social networking sites provide real-time communication and opportunity to connect to an unlimited number of friends at no extra cost. As a result, this enables consumers to share and discuss various opportunities in the market.

Getting information about products across to potential customers is critical as their purchase decisions rely heavily on other customers' opinions from social media (Park et al., 2007). As Simon (2019) pointed out, information systems have increased information flow between opinion leaders and the general public to the point of information overload. Thus, companies are tracking the perceptions of their products from customer reviews and social media to monitor their brand image (Cruz & Lee, 2014). The insights from the tracking can be reflected in their new products (Lee et al., 2008). Predicting the success of a product launch or demand by using online opinions for new or existing products can be helpful for manufacturers to prepare new product launches or revamp existing products (De Janosi, 1959). Accurate prediction is essential for companies to augment their profits and decrease their costs since a new product launch or facelift requires huge costs (Kahn & Chase, 2018).

Researchers have proposed approaches to rank products (Liu et al., 2017), visualize market structure (Chen et al., 2015) and predict sales (Schneider & Gupta, 2016) based on customer opinions. Most studies on sales prediction using online opinions incorporate text mining with econometric modeling (Hsiao et al., 2017) and machine learning approaches (Sohrabpour et al., 2021). Despite the opportunities to use customer opinions to predict the success of a new product launch, studies still lag in two critical aspects. These two aspects, as discussed below, are the motivations for this study.

First, prior studies that analyzed the market structure and predicted product sales using consumer reviews did not consider the lifecycle of newly launched cars. Studies that distinguished product lifecycle such as new and existing (Schneider & Gupta, 2016) are rare. Also, previous studies did not consider when posts are uploaded according to the product lifecycle. It is beneficial to use customer opinions with consideration of whether posts concern expectations of upcoming products or sharing of experiences on released products. Moreover, changes in the frequency by which a product is mentioned in relation to the product's launch schedule can be an important factor in predicting its success. Kim and Hong (2015) proposed the BMIC model as an alternative to overcoming these limitations and increasing compliance with high initial demand due to atmospheric demand, such as records, new drugs, smartphones, and automobiles. In particular, when new products or new technologies are first introduced, they are adopted by a very few adventurous people or institutions, and then gradually spread around them through word of mouth, resulting in a rapid increase in the number of adopters, reaching a level where the market is saturated, and the number of adopters. It will then be handed over to new products or new technologies. Thus, when the cumulative number of new adopters of a new product is modeled on a diffusion curve, the life cycle of the new technology is divided between the introduction and the growth period (Guo, 2014). The introduction period is the period from the time the product is introduced to the take-off point. During this period, mainly innovative consumers or institutions will adopt new or new technologies regardless of the quality or price of new products or new technologies. They are called innovators groups. The period of growth refers to the period when demand for a product or technology expands, from the time of take-off to the time when demand reaches its peak. In the dual market theory, the market is divided into early and major markets, which is when the main market is formed (Muller & Yogev, 2006). The maturity period refers to the period from the peak of demand for a product or technology to the point where demand is sharply reduced. The period of decline refers to the period from the time when the demand for a product or technology is rapidly reduced to the time when there is little demand. Thus, it is important to consider the timing of customers' opinions in relation to a product's launch schedule.

Second, many studies used measured values such as sentiment (Chong et al., 2016) and helpfulness (Choi & Leon, 2020) based on holistic reviews. However, customers have diverse tastes on different types of features such as design, quality, and brand, among others. Therefore, analyzing customer opinions on product features is important since this affects overall ratings and the success of a new product release. Thus, there have been many studies on how to automatically extract product features for diverse products (Titov & McDonald, 2008). Most of the studies, which were based on machine learning approaches and dependency relation, are product independent since they do not need any a priori product knowledge (Yan et al., 2015). These approaches mainly depend on the appearance probabilities of terms and their co-occurrence. However, a lexicon-based approach extracts product features based on a word list that includes manually predefined product-dependent features (Yan et al., 2015). Although this lexicon-based approach is impossible to apply for every product and manually categorized product attributes and distinctive words (Schneider & Gupta, 2016), the categories of product features can be built based on academic studies and inputs from domain experts. Moreover, identifying latent topics and their related terms can validate the categories and their characteristics.

We have two research questions. The first concerns what product features should be considered and how do these considerations change throughout the life cycle of a new product's launch. The second question focuses on what factors have a significant impact on the evaluation of a newly launched automobile by considering customers' interests in product features. We chose the automobile industry as our target since it has distinct features. Moreover, we are able to track launching schedules as well as collect exact sales data for this industry. We gathered online reviews regarding newly launched automobiles launched in South Korea from a web board in Naver.com and investigated the changes in consumers' interests concerning the schedule of the product launch (including periods before and after the launch) via a lens of text analysis. We identified clusters of product features from existing studies and specified synonymous terms in the clusters via a survey. Moreover, we empirically tested the effect of reviews on identifying the important factors of a new product launch by incorporating a machine learning approach.

This study can be expected to improve the existing studies mentioned above as follows. First, compared with previous studies that have used economic variables for prediction, this study can distinguish purchase-related characteristics in the automobile field. Applying text mining and machine learning to analyze user reviews would be a new attempt to develop marketing strategies of automobile company. Second, unlike the single WOM perspective of existing studies, the distinction between input and output perspective in this study can propose a more detailed and clear market strategy. Finally, as the importance of attribution modeling to understand automobile sales and market share increases, organic changes in purchasing determinants by period can be confirmed.

The remainder of this paper is organized as follows. Section 1 reviews the existing literature on text mining and machine learning for market structure. Section 2 presents conceptual framework and hypotheses about input and output word-of-mouth. Section 3 demonstrates the method used to test the hypotheses, with Section 4 presenting the results. Discussion section discusses the theoretical and practical implications of the findings. Conclusion section summarizes the conclusions of this study.

1. Theoretical Background

Text mining strives to solve the information overload problem by using techniques from data mining, machine learning, natural language processing, information retrieval, and knowledge management (Hassani et al., 2020). There have been many studies based on text mining on how to rank products, visualize the market structure, and predict sales with customer opinions (Ghose et al., 2012) proposed a hotel ranking system based on the average utility gain a consumer obtains from staying in a certain hotel (with hotel reservation data and attribute values attained using techniques such as text mining from various social media sources).

Prior studies have shown that market structures derived from online consumer reviews are useful in understanding consumers' perceptions (Mitra & Jenamani, 2020). This understanding of consumers' perceptions and market structure is beneficial for product development, pricing, promotion/campaign, and brand positioning (Chen et al., 2015). Netzer et al. (2012) proposed a market structure surveillance method by utilizing text-mining and semantic network analysis for sedan automobiles and diabetes drugs. Common terms totaling 1,200 and composed of noun phrases and adjectives were used to calculate similarities among automobiles and brands. The selections of terms require human intervention to distinguish similar product features (Chen et al., 2015). Thus, there are some studies that tried to automate obtaining market structure by applying K-means clustering (Lee & Bradlow, 2011) to identify product attributes and topic modeling (Chen et al., 2015).

Many of the existing approaches to the text mining analysis of consumer reviews rely on the extraction of product features and evaluative phrases. There have been many studies to identify product features automatically from online reviews and Yan et al. (2015) classified these approaches into three categories: lexicon-based, dependency-relation-based, and machine-learning-based approaches. These approaches are briefly explained as follows:

(1) Lexicon-based approach utilizes a predefined product-dependent feature lexicon to extract product features and measure users' feature inclination with a sentiment dictionary (Li et al., 2009).

(2) Prior studies identified that there are relations between feature terms and sentiment words in review sentences (Qiu et al., 2011). The dependency-relation-based approach analyzes these relations and applies some rules and algorithms to extract product features from the identified dependency relations (Ahmad & Doja, 2012; Huang et al., 2012).

(3) Machine-learning based approach is similar to the dependency-relation-based approach, but it applies some machine learning algorithms to extract candidate product features. Association rule mining (Wei et al., 2008), undirected graph model (Wong & Lam, 2008) have been utilized to extract product features, and adjectives adjacent to the features are used to measure the sentiment on the product features.

Most of the studies based on machine-learning approaches and dependency relation are product-independent since they do not need any a priori product knowledge (Yan et al., 2015) and are good for applying to many product categories. Though the lexicon-based approach is not feasible to apply to diverse product categories (since it requires terms of product features), it can be beneficial in a certain product category if we identify product features based on prior academic studies and inputs from domain experts. However, many studies identify product features manually and even without proper explanations. Ghose et al. (2012) manually selected features of hotel services such as indoor swimming pools, high-speed Internet, free breakfasts, hairdryers, and parking facilities without any clear explanation, Netzer et al. (2012) also did not explain the selection procedures of the terms in their study. Thus, we used lexicon-based approach to identify product features based on prior studies of predicting automobile sales and similar word sets of the product features are attained through a survey.

2. Conceptual Framework

In the case of the automobile industry, customers often buy their automobiles after some serious thought before making a decision. In particular, automobiles generate high levels of immersion in which there are various factors to consider such as the price, design, and performance of the automobile. Therefore, consumers go through diverse stages of decision-making based on a variety of information. Specifically, consumers go through the stages of search, evaluation, and the final step of purchase decision as influenced by their desires. The five stages of a purchase decision illustrate that the stages of information search and evaluation are important steps in a consumer purchase decision (Dewey, 2007). This process can be seen as a particular form of cost-benefit analysis in the presence of multiple alternatives.

First, need recognition starts from the individual's instinctive need for daily survival. Need recognition for automobiles can also occur when a friend or other sees a masterfully crafted product in front of a parking lot or restaurant or sees a neighbor next door or friend drive by in a new automobile. Thus, the need can be triggered by internal or external stimuli (Kotler, 2009).

If the consumer is aware of his/her need for something, the second step is to search the information. Nowadays, consumers get more information from online communities for automobile to evaluate specific automobiles. As such, the importance of considering online reviews for product demand forecasts is growing. Especially in a high-content society, where the members of society share information more closely, information gathering among people is common and highly reliable. This represents the potential buyer for the automobile's effort of identifying and observing sources of information related to the specific buying decision. This is the basis of the recent online WOM marketing. However, there is a strategy that is more basic and lasts longer than the sponsored WOM marketing. Especially for products of complex nature such as an automobile, i.e., utilitarian and hedonic, the purchase decision requires a high level of involvement (Karimi et al., 2018). Therefore, the consumer for automobile compares web information from various users to make a purchase decision. The online environment allows consumers to exchange views and share reviews of products. It has become common, especially for online purchases of automobiles, to consult other customers' reviews and discuss them.

Third, the collected information is evaluated based on the knowledge, beliefs, circumstances, and preferences already possessed by the consumers. Moreover, consumers evaluate how they value each of the alternatives as well as the related costs. Consumers then compare each other's alternatives, both online and offline. Based on this evaluated information, consumers choose products that have the highest value and lowest opportunity cost for themselves. Specifically, in recent years, consumers have been acquiring diverse information such as discussions of new automobiles, expert testimonies, and video materials on various online sites. This information also stimulates the purchase demand of consumers.

After the stage of information search, consumers evaluate different automobiles or brands based on varying product attributes and related benefits that they are seeking (Kotler, 2009). If customer involvement is high, then they will evaluate several automobile brands; whereas if it is low, only one or a few brands will be evaluated.

After evaluating products, consumers make decisions to purchase products based on their expectations concerning alternatives. Consumers often hold expectations about an alternative's performance on an attribute based on its performance on a different attribute that they see as correlated (Lv et al., 2018). This means some qualified dimensions are more strongly correlated with consumer purchases than other weak attributes (Karimi et al., 2018). Normally, the price is the most important factor in a consumer's purchase decision. However, consumers consider various attributes in their purchase decision according to their level of involvement. In the case of automobiles, functional attributes such as horsepower, acceleration, and low failure probability affect purchase decisions. However, other emotional aspects such as the color, interior, or exterior of the car are also important factors in the purchase decision. According to the attitude formation viewpoint presented in the multi-attribute theory (Fishbein & Ajzen, 1977), the weight of each product property changes because the values of consumers change according to the purchase context. The impact of the various attributes of products are influenced by changing their weight depending on the context of the consumer. Likewise, the subjective norm of others (Fishbein & Ajzen, 1977) and the opinions of others (Choi et al., 2017) are other key factors affecting consumer purchase decisions. Therefore, purchase decisions are made based on various attributes evaluated by consumers, weights by attributes, and results of combining consumers' situations. From the perspective of the company that sells the product, it is important to understand what attributes of the product are important to consumers to understand the success of the product in the market.

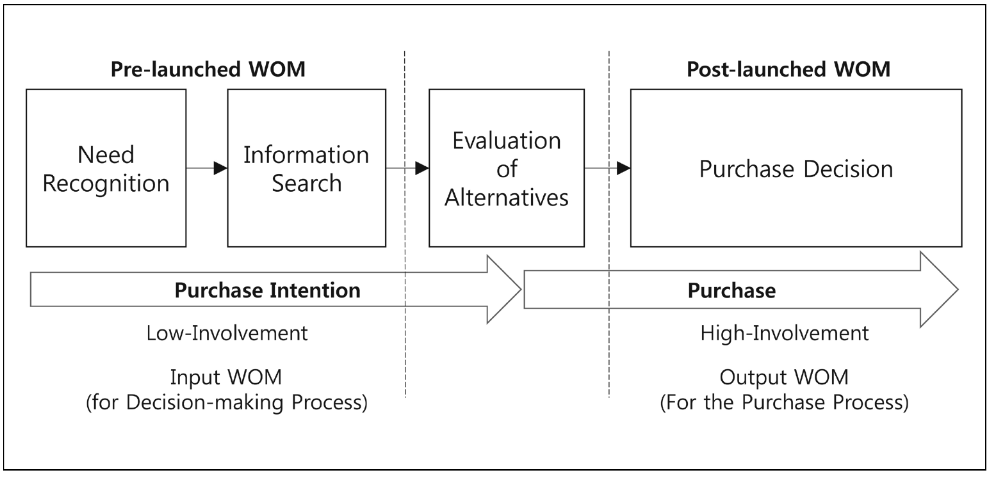
Recent studies suggest opinions and reviews of others as an important factor for consumers' decision to purchase a product of high involvement (Vieira et al., 2018). Consumers value their evaluation, emotion, and preference for low involvement purchases. However, in purchasing a costly, high-involvement product a customer considers the product's benefits from a wide variety of perspectives. Particularly, the final purchase decision can be impeded by two factors: negative feedback from other customers and the level of motivation to comply or accept the feedback (Kotler, 2009). Therefore, this study suggests that it is possible to evaluate the demand for an automobile based on the information that consumers discuss online. Thus, it is important to find out how to use the conventional effect of WOM in automobile-demand evaluation.

According to the social exchange theory, the online environment allows people to develop emotional attachment, form consensus, and experience positive feelings through a mutual exchange (West & Turner, 2007). Moreover, since consumers join online communities according to their will, they more naturally form a consensus among themselves and develop a culture of their own. Emotional cohesion with people on the web can lead to stronger behavioral intentions. Moreover, emotional cohesion can create positive intentions for service and intent to engage in ongoing relationships.

In particular, the WOM effect among online customers serves as a recommender (Bhattacherjee, 2001). The information generated by another customer is more valuable in terms of trust than the information generated by the seller. The review is a major factor in WOM, and customers use product purchase reviews positively to reduce risk and adverse outcomes (Cheung et al., 2007). Online WOM has a great impact on consumers' choice of products. Particularly positive and negative WOM have a strong influence on customers' purchase based on its valence function (Maru File et al., 1994), maximizing the profit from the purchase of products by consumers and minimizing the risk. Therefore, it is important to analyze WOM communication between consumers about the product rather than the product reviews of the experts. The positive and negative information in the online WOM greatly influence consumers in the stage of searching and evaluating information. This implies that from a firm's perspective, the two-sided WOM that does not distinguish between the positive and negative contents of WOM information is more effective in predicting demand. Therefore, WOM information about products posted by users online can be used as proof of comprehensive judgment rather than positive and negative directions of messages. In particular, consideration of negative information can improve information reliability among consumers because information with only positive content is not reliable on a website.

Automobile companies conduct marketing activities in various ways before they introduce a car to the market for sale. Therefore, before a consumer buys a newly launched automobile, the automobile market is already actively engaged in online discussions on various topics by potential consumers, car enthusiasts, and experts. Since the price of an automobile is high, consumers have high involvement in this purchase decision and tend to acquire the maximum amount of information. Generally, consumers evaluate automobiles before they are actually introduced to the market. The automobile has special characteristics that are evaluated again based on user experience when the automobile is introduced to the market.

Fig. 1: Conceptual framework



Source: own

H1: Web user-based categories of automobile characteristics have different effects on user evaluation for the automobile.

Although a review rating reflects the overall experience of a consumer, it is difficult to predict demand based on ratings before consumers experience driving the cars. Likewise, it is unclear what factors should be highlighted before and after launch to ensure accurate demand forecasting and demand retention. Thus, a company should consider both pre and post WOM communications. To this end, we developed a conceptual framework for time-based WOM communication as shown in Fig. 1.

The process of deciding to buy an automobile involves several stages, namely, need recognition, information search, alternative evaluation, and finally, purchase decision. In this process, automakers start advertising before they introduce a car to the market for sales, and potential consumers start getting involved in WOM online. An automobile is a utilitarian good, which gradually increases the consumer's involvement until the purchase decision, and it is also a hedonic product that involves design and emotion as well. Therefore, the stages of information search and alternative evaluation are related to the purchase intention. The consumer then confirms the actual purchase decision through evaluation.

The functions of WOM are valence (Maru File et al., 1994), focus (Christopher et al., 1991), timing (Buttle, 1998), solicitation (Buttle, 1998), and intervention. WOM should especially be expressed timely either before or after purchase (Buttle, 1998). Accordingly, as set in the research framework, we distinguish between pre-launch WOM (Input WOM) and post-launch WOM (Output WOM) based on WOM's timing function. The input WOM is particularly relevant where the product is characterized by search quality rather than easily evaluated by experience quality (Zeithaml et al., 1992). Input WOM provides an important source of pre-purchase information. Consumers ask for recommendations to reduce risk perception in the intangible area (Herr et al., 1991). The reliability of the input WOM mainly influences the customer's decision-making process. Because input WOM provides additional information to customers from trusted sources, giving them greater confidence to make their decision (Empson et al., 2015). Thus, the input WOM is related to pre-purchase WOM solicited from 3rd parties. The output WOM has different characteristics related to the pre-purchase process. Output WOM focuses on WOM after the purchase or consumption experience. Consumers might put their idea or opinion after the purchase or consumption experience.

A negative output WOM can produce a greater effect than a positive output WOM (TARP, 1985). The negative WOM communication is complicated by the emotional context of the conversation (Richins, 1983). Consumers display preferences that convey bad news in negative environments and deliver good news from positive ones (Heath, 1996). However, buyers are still more interested in bad news. East et al. (2007) found that positive WOMs were produced more than negative WOMs. However, for these studies, data were utilized between 2001 and 2003. It also presented aggregated results through questionnaires. The scope of the study, which is more familiar with web activities and based on expectations of new cars, determined that users would be more interested in negative content. Managing negative WOMs by handling complaints has no impact on increasing customer loyalty, but it can impact customer retention (Fornell & Wernerfelt, 1988). This means companies' management of negative WOMs is helpful when consumers decide to buy them. In the long run, however, excessive management of negative WOMs cannot improve customer loyalty to companies, not sales (Jones et al., 2007).

H2: The WOM process has different effects on user evaluation for the automobile in terms of input and output WOM dimensions.

Millions of new consumers, who have secured spending capabilities as emerging markets grow rapidly, are facing a new complex and fast-evolving marketing environment, just like consumers in developed countries. AIDA was presented to explain customer purchasing behavior that customer attention-interest-desire-action is important for purchasing decisions (Rawal, 2013). However, the current point of consumption advantage, not supply, is changing AIDA into the purchasing paradigm of AISAS (Attention-Interest-Search-Action-Share) (Tseng & Wei, 2020).

As in advanced markets, technology has increased the participation of emerging market consumers at each stage of the purchasing decision-making process, but several important changes need to be made to reflect the differences in brand and product characteristics. Thus, it is important to utilize the power of WOM effect, which is judged to play a significant role in the purchasing decision-making process of emerging market consumers. However, in emerging markets compared to advanced markets, the time to search for products is longer and more important in the process of consumers' purchasing decisions (Stokes & Lomax, 2002). So, it is necessary to focus on changing the determinants of customers at the time of purchase.

In other words, for product and customer management through WOM, a company's manager needs to understand the process of multidimensional WOM, i.e., purchasing phase and time flow, among others rather than a single concept (Fornell & Wernerfelt, 1988). Thus, investigation of input and output WOM processes amongst customers and other important influencers is important to an automobile launch-based specific context. The output WOM is stimulated by the degree of engagement. Thus, investigation of input and output WOM processes amongst customers and other important influencers are important to automobile launch-based specific contexts.

H3: The purpose (long-term market share vs. rapid sales performance) of a product and its success in the market changes as influenced by various factors related to the timing of product information exposure.

3. Research Methodology

To carry out the analysis according to the purpose of the research, we used data from an online forum site for the automobile. The online community is the most realistic way to see the reaction of prospective buyers to new automobile models. Therefore, automobile-related reviews of the online community can be regarded as the most sensitive data for the development of a market strategy for automobile sales based on the effect of WOM (Netzer et al., 2012).

We collected data from the automobile section of Naver (http://auto.naver.com/), which is the most popular web portal of South Korea. In the automobile section, there are online bulletin boards that have 87,756 posts by 30,751 users on average per year. The number of collected reviews is 401,322 from Naver online bulletin boards that occurred during the period from July 2010 to August 2016. From the collected data, we normalized a large chunk of review text for reducing noises such as typing mistakes or linguistic errors. After normalization step, we tokenized the review text and conduct POS (Parts of Speech) tagging to identify and extract keywords. The average length of the posts was 78.6 words. We collected user reviews for 18 automobile models and the selections were made to compare nine successful and nine normal models with similar size as shown in Tab. 1. The distinction between the success model and the normal model separated was made based on sales volume data. The sales volume data were divided into monthly sales data for each vehicle model in the Korea Automated Manufacturers Association. All the automobile models are manufactured in South Korea. The successful model is defined as an automobile that has a significant hike in sales in the first month after its launch when compared to the monthly sales of other cars. On the contrary, the normal model refers to a group of automobiles that has no rapid increase in sales in the first month after its launch.

This study developed a word category related to automobile characteristics based on previous studies (Jalilvand & Samiei, 2012; Wang et al., 2009). As shown in Tab. 2, the category characteristics are identified namely, country of origin, price, brand, design, emotional function, technical function, advertising and PR, service quality, environment for company, and industry. To use our text data from numerous online user reviews and to develop a word dictionary for the automotive category, we first identified keywords related to each automobile characteristic. The initial keywords in Tab. 2 were also identified from the previous studies on each characteristic.

However, customers use different expressions for the same concepts and use colloquial expressions and casual words in online communities. To identify the meaning of these words, we need to develop a dictionary for generating thesaurus and categorizing words into one of the automobile characteristics. To develop the dictionary, we performed an experiment to identify the related words and colloquial expressions for each characteristic by hiring sixty participants. Each participant was asked to write related words for the initial keywords in each characteristic. After we collected similar words

Tab. 1: The selected automobile models

Type	Automobile model	Size	Type	Automobile model	Size
Successful model	Chevrolet Spark	Compact	Normal model	Kia Picanto	Compact
	Hyundai Elantra GT	Compact		Chevrolet Cruze	Compact
	Chevrolet Malibu	Mid-size		Kia Optima	Mid-size
	Renault Talisman	Mid-size		GM Buick Lacrosse	Mid-size
	Hyundai Azera	Mid-size		Kia Cadenza	Mid-size
	Hyundai Genesis	Luxury		Kia K900	Luxury
	Hyundai SantaFe	Mid-size SUV		Chevrolet Captiva	Mid-size SUV
	Kia Sorento	Mid-size SUV		Renault Koleos	Mid-size SUV
	Kia Sedona	Minivan		Ssangyong Turismo	Minivan

Source: own

Tab. 2: Automobile characteristics and the extracted keywords – Part 1

Category		Initial keywords	Additional keywords
Country of origin		Location specific, Showroom-oriented variables, Purchase-oriented factors	Domestic, Import, Foreign trade, North America, Export, National, Foreign, Germany, World, Store
Price		Sales volume, Contract, Price, Sales, Discount, etc	Sales, Contract, Price, Sale, Discount, Purchase, Cash, Installment, Advance, Price, Premium, Maintenance, Share, Cost, Initial cost, Release, Base, Total cost, Used
Brand		Brand logo (Differentiation, Uniqueness, Color, Vividness, Preference), Brand personality factors	Each brand name (Referring to Tab. 1)
Design	Inner design	Color, Inner design, Door handle, Lamp, etc	Design, Luxury, Indoors, Cruise, Basic, Color, Seat, Real, Photography, Space, Premium, Art
	Outer design	Variety of color, Bumper hall, Molding, etc	Design, Luxury, Cruise, Medium, Large, Model, Basic, Rear view, Color, Life, Large, Real, Photography, Space, Line, Premium
Emotional function		Performance (Stability, comfortability)	Convenience, Daewoo, Feeling safe, Luxury, Comfort, Seats, Interior, Ideas, Best, Advance, Real, Important, Differences, Standards

Tab. 2: Automobile characteristics and the extracted keywords – Part 2

Category		Initial keywords	Additional keywords
Technical function	Hardware (HW)	HW performance (Durability, Fuel efficiency)	Compare, Options, Engine, Fuel Economy, Defects, Diesel, Spare parts, Big, Problem, Difference
	Software (SW)	Acceleration, Maximum speed, Handling	Travel, Compare, Options, Performance, Class, Issues, Differences, Faults, Ride, Fast, Horsepower, Startup, Output, Speed, Acceleration, Function, Gout
Advertising and PR		Attitude for BBBs site (Preference, Quality of information, Design, Satisfaction for searching)	Internet, Media, Release, Bulletin Board, Share, Photo, First, News, Line, Now
Service quality		Service quality (Fixing service, Emergency call, Customer satisfaction, Repurchase intention, Intent to reputation, Salesperson)	Test, Consumer, Maintenance, Installment, Discount, Recall, Issue, Stock, Repair, Exchange
Environment for company and industry		Creativity, Utility, Uniqueness, Industrial factors (Law, Social, Economic, Technical), Mega trend (Eco-friendly, High-efficiency, etc)	Market, Recall, Release, Company, Premium, Bad, Issue, Mental, Tax, Export, Competition, Lately, Domestic, Enterprise, Share, Person, Used

Source: own

for each category, we compiled the word list and used the developed lexicon for analyzing user reviews. The results of this experiment are the additional keywords in Tab. 2.

In this study, we conducted LDA (Latent Dirichlet Allocation) modeling to identify the cluster arrangement by user-centered keywords for the automobile model. LDA-based topic modeling analysis is performed to confirm the clarity of the classified word groups. LDA modeling has been used to explore and derive topics in the characteristics of automobiles that reflect recent trends (Blei et al., 2003; Chen et al., 2012). Thus, topic modeling can analyze specific words based on vast text data to identify specific topics by word usage and how these words are connected. Moreover, the importance of automobile-related keywords that are derived from actual automobile users can be confirmed as opposed to distinguishing automobile characteristics in the manufacturer's context. As such, considering the purpose of this study, the LDA method is a suitable method for validating the lexicon-based approach related to the keywords of automobile demand as gathered from web discussion. We conducted the LDA through python and gensim libraries (https://radimrehurek.com/gensim/). The data used in the LDA were nouns, verbs, and adjectives. The number of topics was selected through the highest topic coherence value through several repetitive experiments. We calculated the probability values of all the words included in the vocabulary list and then performed topic modeling. Based on the defined lexical dictionary, we calculated co-occurrences of words and core words in the topic using LDA.

Tab. 2 represents the categories and related attributes of automobile characteristics based on interviews and LDA analysis. Specifically, previous studies have identified the following characteristics related to automobile purchase: country of origin, sale price, brand, design, emotional function, technical function, advertising and PR, service quality, and managerial environment for company and industry (Breteau & Weber, 2013; Wang et al., 2009). Since the process of determining the number of clusters according to the characteristics of the automobiles to be classified is a prerequisite, the optimal number of clusters was determined for the classified clusters based on the previous research (Jalilvand & Samiei, 2012; Wang et al., 2009). To test *H1,* we compared the variation rate for keyword order for successful and normal models. And then support vector machine was used for identifying *H2. H3* was also tested with the distribution of keyword categories for the successful and normal model.

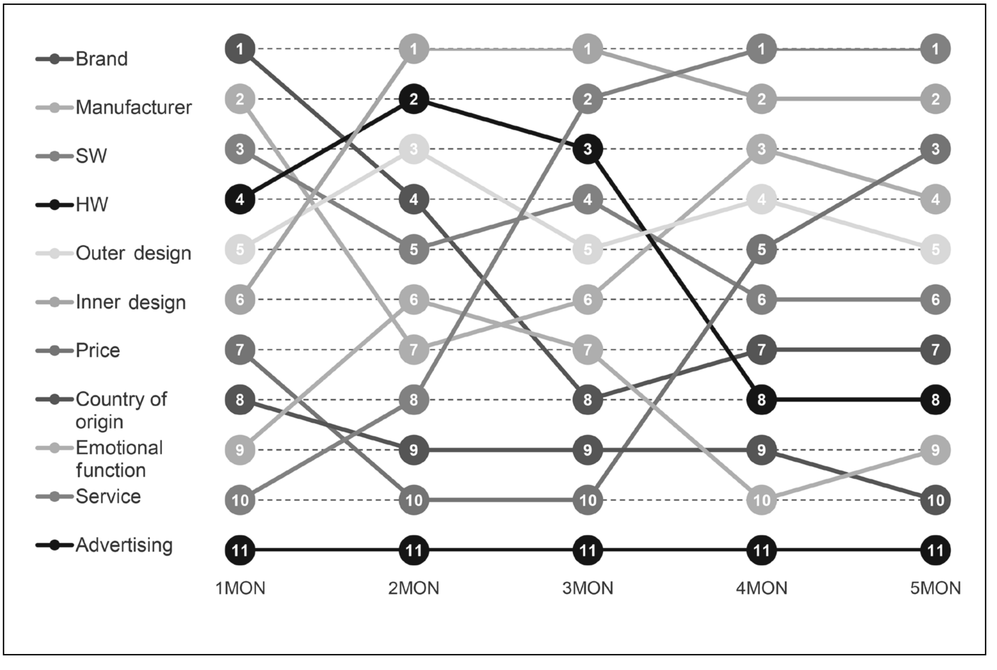
4. Research Results

For the automobile models that are considered successful, Fig. 2 shows the results of analyzing the keywords based on the dictionary we built. This picture confirms the rate of change before and after the launch. In calculating the rate of change, 1MON is a measure of the rate of change in the window size one month before the release and one month after the release. If the distribution ratio of 'Brand' one month before the launch was 10% among all categories and the distribution ration of 'Brand' one month after the launch was 12%, the rate of change is 20%. Likewise, 2MON is a calculation of the rate of change in two months prior to launch and two months after launch. We calculated the rate of change for all the categories from 1MON to 5MON and Fig. 2 and 3 show the rankings of the rate of change for the categories from the successful and normal model, respectively.

The keywords that emerged from the product launch date include brand, exterior design, interior design, and SW function. It can be confirmed that the initial information about the product is mainly evaluated by design and brand. In Fig. 2, 'Brand' appears on top at 1MON. This means that the successful model has a marked increase in brand mentality in the period one month before and after launch. However, practical aspects such as exterior design, country of origin, and HW functions became more prominent two months after an automobile's release. In particular, when we look at the frequency of occurrence of monthly keywords, some of the keyword categories such as 'Brand' and 'Hardware' that emerged right before the automobile launch are generally low by the fifth month. However, until two months before launch, some of these keyword categories emerged in high-level debates.

Tab. 3 shows the distribution of keyword categories before and after the launch of the successful models. As Tab. 3 shows, brand and manufacturer, H/W and S/W are considered

Fig. 2: Variation rate for keyword order (successful model)



Source: own

important factors at the beginning of the launch. However, from two months after the launch, it may be advantageous to adopt a strategy in the market by emphasizing interior design, emotional function, and service-oriented factors referring to Fig. 2. Meanwhile, keywords such as price, country of origin, and advertising are rarely presented as important factors. In other words, the rapid increase in the number of automobiles sold since its launch has been largely dependent on design and brand evaluation.

By contrast, the normal model in Fig. 3 shows a rapid decline in keywords for elements such as country of origin, advertising, brand, and exterior design within a month of its launch. In Fig. 3, 'Price' appears on top at 1MON. This means that the successful model has a marked increase in brand mentality for one month before and after launch. In the rapidly growing themes, there are many practical themes such as service, price, and interior design. Tab. 4 shows the distribution of keyword categories before and after the launch of the normal models. As shown in Tab. 4, the majority of related topics are sharply debated over all the category topics within a month after the automobile model is released, specifically regarding performance, but this is also reduced within one month. Overall, in the case of the normal model, the frequency of WOM on the web is greatly reduced within one month.

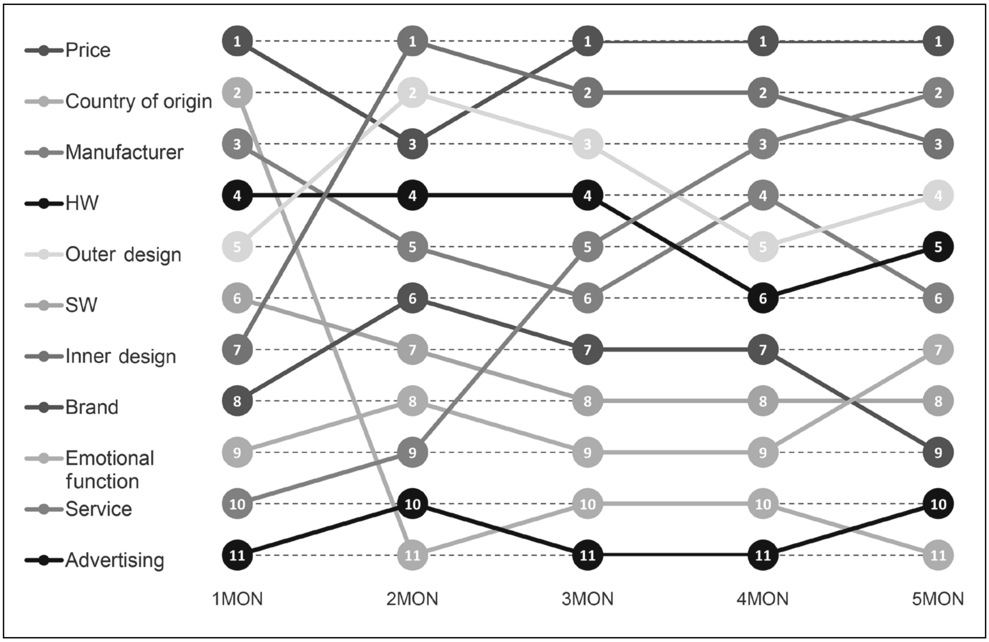
Support vector machine (SVM) was performed based on the word characteristics of the selected automobile models to set up a total of three models to test *H2.* Similar to discriminant analysis, a linear support vector machine technique tries to find a hyperplane that separates the target classes in the space of the input variables. Like most classification techniques, support vector machines can be trained most effectively when the distribution of the target variable is approximately equal in the training set. The first model consisted of "the number of comments before launch", "the number of comments one month after launch", "the number of comment users before launch" and "the number of comment users one month after launch." Model 2 was added to Model 1 with additional variables. The added variables are "number of words per word category before launch" and "number of words after one month." Model 3 included "number of words per word category before launch" and "growth rate of one month before launch" in model 1. For training a SVM model, six models were selected from the nine successful models and six models also were randomly selected from the nine normal models. The rest of the data was used as a test set. This random splitting of dataset for training and testing SVM models was performed 30 repetitions. The analyzed results are shown in Tab. 5 below and the accuracy values are the average accuracy values of the 30 repetitions. In this study, SVM function in e1071 library (Meyer et al., 2021) of R was

Tab. 3: Distribution of keyword categories (successful model)

Category	Before launch	After 30d	After 60d	After 90d	After 120d	After 150d
Country of origin	4%	5%	4%	4%	4%	4%
Price	7%	8%	7%	7%	8%	10%
Brand	14%	13%	14%	16%	17%	13%
Inner design	7%	5%	5%	5%	4%	6%
Outer design	8%	7%	8%	7%	5%	6%
Emotional function	11%	11%	11%	10%	9%	10%
HW	20%	21%	20%	19%	19%	19%
SW	11%	12%	12%	12%	13%	12%
Advertising	2%	1%	2%	2%	2%	1%
Service	4%	5%	5%	5%	4%	7%
Manufacturer	11%	12%	12%	12%	14%	12%

Source: own

Fig. 3: Variation rate for keyword order (normal model)



Source: own

Tab. 4: Distribution of keyword categories (normal model)

Category	Before launch	After 30d	After 60d	After 90d	After 120d	After 150d
Country of origin	4%	3%	5%	3%	4%	4%
Price	8%	9%	9%	8%	9%	8%
Brand	9%	7%	12%	11%	7%	9%
Inner design	10%	10%	8%	9%	8%	10%
Outer design	11%	13%	8%	9%	8%	11%
Emotional function	12%	12%	11%	14%	14%	12%
HW	20%	19%	19%	18%	21%	20%
SW	11%	12%	9%	14%	14%	12%
Advertising	2%	1%	1%	1%	1%	1%
Service	3%	5%	5%	4%	6%	5%
Manufacturer	11%	8%	12%	9%	8%	8%

Source: own

used and the default parameter values were used for the test. The SVM results show that Model 3 has better results than Models 1 and 2. The analysis shows that sales evaluations are more advantageous when more comments are used.

Tab. 5: Result for SVM

Model	Variables	Accuracy
1	- Number of comments before launch - Number of users who posted comments for one month after launch - Nnumber of users who posted comments before launch - Number of comments after one month of launch	0.872
2	- Model 1 variables - Number of words per word category before launch - Number of words for one month after launch	0.878
3	- Model 1 variables - Number of words per word category before launch - Growth rate for one month after launch	0.933

Source: own

Discussion

This study explored if the market demand for products can be predicted based on user feedback found in online communities. In relation to purchase decisions, the automobile industry in which various influencing factors exist was used as a target area. Factors affecting the purchase of the product include various psychological factors (i.e., attitudes, beliefs, and behavioral intentions), functional factors, emotional factors, advertising, and media, among others. This study distinguished the important factors affecting purchase demand based on the changing perspectives of target consumers throughout the periods before and after a product's market launch. Based on user opinions expressed in online discussion boards, users' interests were identified and monitored by time flow for changes in the periods before and after a product's launch. Moreover, we compared the effects of the identified success factors on successful and normal cars.

The results of the study provide the following implications. First, it can be seen that the rankings of the subcategories vary based on the particular timing considered (i.e., before or after the automobile's launch). Moreover, the results also differ between successful and normal car models. In the case of the successful models, brand, manufacturer, and price became unimportant factors after product launch while inner design, emotional function, and service became important factors to consider after the product launch. However, the advertisement itself was not a big issue regardless of the stages of product launch. Therefore, in the case of the successful model, brand and manufacturer, H/W and S/W, and outer and inner designs are considered important factors at the beginning of the launch as shown in previous study (Mazurova, 2017). However, from two months after the launch, it may be advantageous to adopt a strategy in the market by emphasizing interior design, emotional function, and service-oriented factors. In particular, H/W and country of origin can be treated as less important factors than price three months after the product launch. This result means that the characteristic elements classified on a user-based basis have different influences over the evaluation of the car over time (Tseng & Wei, 2020). Especially, variables that are more detailed than existing studies can be determined to be used in user evaluations at different frequencies for success models. Thus, *H1* was accepted.

Second, based on the SVM results, predicting the success of an automobile launch can be improved by considering the parameters presented in this study. Companies should consider in detail how their reputation is perceived in online communities to improve their market prediction (Chen et al., 2016). However, as time elapses from the launch date, companies will have to readjust their market-oriented strategies by applying different weights to the customers' decision factors (Schlosser & McNaughton, 2009). Marketing strategies and sub-factors related to consumers' decision-making can have different effects on consumer purchases depending on the product's life cycle. While previous studies have pointed out the general WOM elements and their importance (Frasquet et al., 2017; Reimer & Benkenstein, 2016), we tested that the impact of different influencing factors may vary at different points of time before and after a product's launch. This means that the roles of the input and output WOMs are different and that each WOM process has a different function in the market. Therefore, *H2* was accepted.

Third, the effects of success factors were compared between the successful and normal models. Based on our analysis, the country of origin, brand, and software lose their importance after the product launch. In particular, the country of origin was not considered at all by the consumers from the second month after the release date. However, the price and inner and outer designs of the automobile were considered extremely important factors, and this sentiment continued to strengthen over time. Therefore, focusing on the price and design is more appropriate for a company to increase its long-term market share instead of momentarily achieving rapid sales of the automobile after the launch (Bone, 1995). Generally speaking, it is important to sell several different automobile products. However, each product deserves a distinct strategy for the market. A product being market-oriented means responding to changes in purchase-related factors that affect consumers (Bone, 1995). Before and after the product launch, companies need to consider that the market vision of the product can change. Considering the distribution of keyword categories of successful and normal models, we can see that the references to important variables vary over time. As a result, the importance of the variables used in user evaluations changes with time. Thus, we can say that *H3* was accepted.

Conclusions

This study focused on the predictive relationship between online opinions and sales of automobiles. Although there are several approaches to predict sales of automobiles such as economic modeling (Brownstone & Train, 1998) and user-based survey (Vardakas et al., 2014) among others, we have presented that user opinions analyzed through text mining can have predictive power in forecasting market share and sales volume for automobiles. In this study, we have developed a method of categorizing automobile characteristics and their related variables. When an automobile manufacturer predicts market share and success of their new product, they need to consider pre-WOM and post-WOM before and after a product's launch date.

Academically, this study explored the effectiveness of online text mining in predicting sales and market share of automobiles as follows. First, previous researches have focused on economic variables to predict sales and market share (Chowdhury et al., 2016; O'Brien, 2018; Sallee et al., 2016). This study separated the categories of automobile characteristics more clearly by utilizing text mining and machine learning. Compared with previous studies (Chen et al., 2016; Frasquet et al., 2017), our categories can be used to measure the success probability of automobile sales in the market in a more theoretical manner. Second, this study presented how input WOM and output WOM should be differently considered based on various user reviews in the online communities when latent customers consider purchasing automobiles. Existing WOM studies have investigated the characteristics of online sharing of opinions, its applicability, and the sociality of users (Bone, 1995; Chen et al., 2016; Frasquet et al., 2017; Reimer & Benkenstein, 2016). However, this study empirically divided the characteristics of WOM into input and output WOM. As such, dividing WOM into two types (input and output) and applying both types to the purchase process has theoretical implications.

Third, this study revealed which dimensions of automobile characteristics are important factors in identifying sales volume and market share for specific types and brands of automobile models. The elements of market success in automobiles that have been raised from previous studies are quite diverse. Through data analysis and experiment, this study integrated the separate success factors of various viewpoints based on the characteristics of the automobile. According to sub-characteristics of the automobile classified, significant success-factors as the result of this research can be integrated for understanding market success. Therefore, we conclude that the developed clusters and keywords for automobile success generated in this study provide theoretically important information to understand automobile research.

Practically, this study suggests how text analysis of online reviews can be used to identify sales possibilities and market share before automobiles are even introduced to the public. Through text mining, LDA, and SVM technique, online users' opinions for brand-new automobile models can help brand managers develop effective marketing strategies for increasing automobile sales by considering model-specific and brand-specific factors. Second, companies that anticipate accurate prediction of automobile demand need to recognize the importance of online reviews. There is a strategic limit to companies estimating the success of automobiles in competitive markets simply by calculating the net profit to cost in relation to demand forecasts for newly launched automobiles or by expert reviews alone. As suggested in this study, the opinions of online prospective buyers are of great significance in predicting sales of newly launched automobiles and securing competitive advantage in the market. Third, as presented in this study, companies need to identify the viewpoints of various categories of market forecasts for newly released automobiles. As noted in this study, the major attributes of each category based on automobile characteristics can serve as a useful guide for companies to evaluate the marketability of automobiles before and after market launch. In particular, the market strategy for newly launched automobiles needs to be weighted differently depending on the period before and after launch.

Based on the results of this study, future research topics arise from various combinations of themes. It is necessary to analyze the evaluation of consumers' vehicles from various viewpoints such as comparisons of domestic and imported vehicles, and new and used cars. Besides, when analyzing the characteristics of online WOM, it is possible to identify cultural differences in customer reviews as part of the success factors for building a global automotive market strategy in the future. It is also possible to establish a predictive strategy by matching dynamic topic modeling that reflects time flow, improving on the LDA modeling, and by using specific weekly market sales information per vehicle brand. Nevertheless, this study contributes to the literature by integrating information on automobile success factors with the use of various data mining technology and by developing a data dictionary from user experiments that could prove useful for future researches in this field.

Acknowledgments: *This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT). [NRF-2020R1A2C2099665] and the Soonchunhyang University Research Fund.*

References

Ahmad, T., & Doja, M. N. (2012). Rule Based System for Enhancing Recall for Feature Mining from Short Sentences in Customer Review Documents. *International Journal of Computer Science and Engineering*, *4*(6), 1211–1219.

Appel, G., Grewal, L., Hadi, R., & Stephen, A. T. (2020). The future of social media in marketing. *Journal of the Academy of Marketing Science, 48*(1), 79–95. https://doi.org/10.1007/s11747-019-00695-1

Bhattacherjee, A. (2001). Understanding Information Systems Continuance: An Expectation-Confirmation Model. *MIS Quarterly, 25*(3), 351–370. https://doi.org/10.2307/3250921

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Bone, P. F. (1995). Word-of-mouth effects on short-term and long-term product judgments. *Journal of Business Research, 32*(3), 213–223. https://doi.org/10.1016/0148-2963(94)00047-I

Breteau, V., & Weber, S. (2013). Reconsidering the Choice between Gasoline- and Diesel-Powered Cars: Modeling Demand and Automakers' Reactions. *Transportation Research Record: Journal of the Transportation Research Board, 2375*(1), 18–28. https://doi.org/10.3141/2375-03

Brownstone, D., & Train, K. (1998). Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics, 89*(1–2), 109–129. https://doi.org/10.1016/S0304-4076(98)00057-8

Buttle, F. A. (1998). Word of mouth: understanding and managing referral marketing. *Journal of Strategic Marketing, 6*(3), 241–254. https://doi.org/10.1080/096525498346658

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly, 36*(4), 1165–1188. https://doi.org/10.2307/41703503

Chen, K., Kou, G., Shang, J., & Chen, Y. (2015). Visualizing market structure through online product reviews: Integrate topic modeling, TOPSIS, and multi-dimensional scaling approaches. *Electronic Commerce Research and Applications, 14*(1), 58–74. https://doi.org/10.1016/j.elerap.2014.11.004

Chen, Y., Yang, S., & Wang, Z. (2016). Service cooperation and marketing strategies of infomediary and online retailer with eWOM effect. *Information Technology and Management, 17*(2), 109–118. https://doi.org/10.1007/s10799-015-0237-1

Cheung, M.-S., Anitsal, M. M., & Anitsal, I. (2007). Revisiting Word-of-Mouth Communications: A Cross-National Exploration. *Journal of Marketing Theory and Practice, 15*(3), 235–249. https://doi.org/10.2753/MTP1069-6679150304

Choi, H. S., & Leon, S. (2020). An empirical investigation of online review helpfulness: A big data perspective. *Decision Support Systems, 139,* 113403. https://doi.org/10.1016/j.dss.2020.113403

Choi, J., Lee, H. J., & Kim, H.-W. (2017). Examining the effects of personalized App recommender systems on purchase intention: A self and social-interaction perspective. *Journal of Electronic Commerce Research, 18*(1), 73–102.

Chong, A. Y. L., Li, B., Ngai, E. W. T., Ch'ng, E., & Lee, F. (2016). Predicting online product sales via online reviews, sentiments, and promotion strategies. *International Journal of Operations & Production Management, 36*(4), 358–383. https://doi.org/10.1108/IJOPM-03-2015-0151

Chowdhury, M., Salam, K., & Tay, R. (2016). Consumer preferences and policy implications for the green car market. *Marketing Intelligence & Planning, 34*(6), 810–827. https://doi.org/10.1108/MIP-08-2015-0167

Christopher, M., Payne, A., & Ballantyne, D. (1991). *Relationship Marketing: Bringing quality customer service and marketing together*. Oxford: Butterworth-Heinemann.

Cruz, R. A. B., & Lee, H. J. (2014). The Brand Personality Effect: Communicating Brand Personality on Twitter and its Influence on Online Community Engagement. *Journal of Intelligence and Information Systems, 20*(1), 67–101. https://doi.org/10.13088/jiis.2014.20.1.067

De Janosi, P. E. (1959). Factors Influencing the Demand for New Automobiles. *Journal of Marketing, 23*(4), 412–418. https://doi.org/10.1177/002224295902300408

Dewey, J. (2007). *How We Think & Education And Experience*. London: Frederick Ellis.

East, R., Hammond, K., & Wright, M. (2007). The relative incidence of positive and negative word of mouth: A multi-category study. *International Journal of Research in Marketing, 24*(2), 175–184. https://doi.org/10.1016/j.ijresmar.2006.12.004

Empson, L., Muzio, D., Broschak, J. P., & Hinings, C. R. (2015). *The Oxford Handbook of Professional Service Firms*. Oxford University Press.

Fishbein, M., & Ajzen, I. (1977). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Boston, MA: Addison-Wesley.

Fornell, C., & Wernerfelt, B. (1988). A Model for Customer Complaint Management. *Marketing Science, 7*(3), 287–298. https://doi.org/10.1287/mksc.7.3.287

Frasquet, M., Miquel, M. J., & Mollá, A. (2017). Complaining at the Store or Through Social Media: The Influence of the Purchase Channel, Satisfaction, and Commitment. In F. Martínez-López, J. Gázquez-Abad, K. Ailawadi, & M. Yagüe-Guillén (Eds.), *Advances in National Brand and Private Label Marketing* (pp. 87–94). Cham: Springer.

Ghose, A., Ipeirotis, P. G., & Li, B. (2012). Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content. *Marketing Science, 31*(3), 493–520. https://doi.org/10.1287/mksc.1110.0700

Guo, X. (2014). A novel Bass-type model for product life cycle quantification using aggregate market data. *International Journal of Production Economics, 158,* 208–216. https://doi.org/10.1016/j.ijpe.2014.07.018

Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text Mining in Big Data Analytics. *Big Data and Cognitive Computing, 4*(1), 1. https://doi.org/10.3390/bdcc4010001

Heath, C. (1996). Do People Prefer to Pass Along Good or Bad News? Valence and Relevance of News as Predictors of Transmission Propensity. *Organizational Behavior and Human Decision Processes, 68*(2), 79–94. https://doi.org/10.1006/obhd.1996.0091

Herr, P. M., Kardes, F. R., & Kim, J. (1991). Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective. *Journal of Consumer Research, 17*(4), 454–462. https://doi.org/10.1086/208570

Hsiao, Y.-H., Chen, M.-C., & Liao, W.-C. (2017). Logistics service design for cross-border E-commerce using Kansei engineering with text-mining-based online content analysis. *Telematics and Informatics, 34*(4), 284–302. https://doi.org/10.1016/j.tele.2016.08.002

Huang, S., Liu, X., Peng, X., & Niu, Z. (2012). Fine-grained product features extraction and categorization in reviews opinion mining. In *Proceedings of the 2012 IEEE 12th International Conference on Data Mining Workshops*. Washington, DC: IEEE Computer Society.

Jalilvand, M. R., & Samiei, N. (2012). The effect of electronic word of mouth on brand image and purchase intention. *Marketing Intelligence & Planning, 30*(4), 460–476. https://doi.org/10.1108/02634501211231946

Jones, M. A., Reynolds, K. E., Mothersbaugh, D. L., & Beatty, S. E. (2007). The Positive and Negative Effects of Switching Costs on Relational Outcomes. *Journal of Service Research, 9*(4), 335–355. https://doi.org/10.1177/1094670507299382

Kahn, K. B., & Chase, C. W. (2018). The State of New-Product Forecasting. *Foresight: The International Journal of Applied Forecasting, 51,* 24–31.

Karimi, S., Holland, C. P., & Papamichail, K. N. (2018). The impact of consumer archetypes on online purchase decision-making processes and outcomes: A behavioural process perspective. *Journal of Business Research, 91,* 71–82. https://doi.org/10.1016/j.jbusres.2018.05.038

Kim, T., & Hong, J. (2015). Bass model with integration constant and its applications on initial demand and left-truncated data. *Technological Forecasting and Social Change, 95,* 120–134. https://doi.org/10.1016/j.techfore.2015.02.009

Kotler, P. (2009). *Marketing Management: A South Asian Perspective*. Pearson Education India.

Lee, J., Park, D.-H., & Han, I. (2008). The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic Commerce Research and Applications, 7*(3), 341–352. https://doi.org/10.1016/j.elerap.2007.05.004

Lee, T. Y., & Bradlow, E. T. (2011). Automated Marketing Research Using Online Customer Reviews. *Journal of Marketing Research, 48*(5), 881–894. https://doi.org/10.1509/jmkr.48.5.881

Li, Z., Zhang, M., Ma, S., Zhou, B., & Sun, Y. (2009). Automatic Extraction for Product Feature Words from Comments on the Web. In *Asia Information Retrieval Symposium, Lecture Notes in Computer Science* (Vol. 5839, pp. 112–123). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-04769-5_10

Liu, Y., Bi, J.-W., & Fan, Z.-P. (2017). Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory. *Information Fusion, 36,* 149–161. https://doi.org/10.1016/j.inffus.2016.11.012

Lv, Z., Jin, Y., & Huang, J. (2018). How do sellers use live chat to influence consumer purchase decision in China? *Electronic Commerce Research and Applications, 28,* 102–113. https://doi.org/10.1016/j.elerap.2018.01.003

Maru File, K., Cermak, D. S., & Alan Prince, R. (1994). Word-of-Mouth Effects in Professional Services Buyer Behaviour. *Service Industries Journal, 14*(3), 301–314. https://doi.org/10.1080/02642069400000035

Mazurova, E. (2017). Exploratory Analysis of the Factors Affecting Consumer Choice in E-Commerce: Conjoint Analysis. *Journal of Information Systems Engineering and Management, 2*(2), 12. https://doi.org/10.20897/jisem.201712

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2021). *Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071)*. Wien: TU. Retrieved from https://CRAN.R-project.org/package=e107

Mitra, S., & Jenamani, M. (2020). OBIM: A computational model to estimate brand image from online consumer review. *Journal of Business Research, 114,* 213–226. https://doi.org/10.1016/j.jbusres.2020.04.003

Muller, E., & Yogev, G. (2006). When does the majority become a majority? Empirical analysis of the time at which main market adopters purchase the bulk of our sales. *Technological Forecasting and Social Change, 73*(9), 1107–1120. https://doi.org/10.1016/j.techfore.2005.12.009

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine Your Own Business: Market-Structure Surveillance Through Text Mining. *Marketing Science, 31*(3), 521–543. https://doi.org/10.1287/mksc.1120.0713

O’Brien, J. (2018). Age, autos, and the value of a statistical life. *Journal of Risk and Uncertainty, 57*(1), 51–79. https://doi.org/10.1007/s11166-018-9285-3

Park, D.-H., Lee, J., & Han, I. (2007). The Effect of On-Line Consumer Reviews on Consumer Purchasing Intention: The Moderating Role of Involvement. *International Journal of Electronic Commerce, 11*(4), 125–148. https://doi.org/10.2753/JEC1086-4415110405

Qiu, G., Liu, B., Bu, J., & Chen, C. (2011). Opinion Word Expansion and Target Extraction through Double Propagation. *Computational Linguistics, 37*(1), 9–27. https://doi.org/10.1162/coli_a_00034

Rawal, P. (2013). AIDA Marketing Communication Model: Stimulating a Purchase Decision in the Minds of the Consumers through a Linear Progression of Steps. *IRC's International Journal of Multidisciplinary Research in Social Management, 1,* 37–44.

Reimer, T., & Benkenstein, M. (2016). When good WOM hurts and bad WOM gains: The effect of untrustworthy online reviews. *Journal of Business Research, 69*(12), 5993–6001. https://doi.org/10.1016/j.jbusres.2016.05.014

Richins, M. L. (1983). Negative Word-of-Mouth by Dissatisfied Consumers: A Pilot Study. *Journal of Marketing, 47*(1), 68–78. https://doi.org/10.1177/002224298304700107

Sallee, J. M., West, S. E., & Fan, W. (2016). Do consumers recognize the value of fuel economy? Evidence from used car prices and gasoline price fluctuations. *Journal of Public Economics, 135,* 61–73. https://doi.org/10.1016/j.jpubeco.2016.01.003

Schlosser, F. K., & McNaughton, R. B. (2009). Using the I-MARKOR scale to identify market-oriented individuals in the financial services sector. *Journal of Services Marketing, 23*(4), 236–248. https://doi.org/10.1108/08876040910965575

Schneider, M. J., & Gupta, S. (2016). Forecasting sales of new and existing products using consumer reviews: A random projections approach. *International Journal of Forecasting, 32*(2), 243–256. https://doi.org/10.1016/j.ijforecast.2015.08.005

Simon, H. A. (2019). *The Sciences of the Artificial*. Cambridge, MA: MIT Press.

Sohrabpour, V., Oghazi, P., Toorajipour, R., & Nazarpour, A. (2021). Export sales forecasting using artificial intelligence. *Technological Forecasting and Social Change, 163,* 120480. https://doi.org/10.1016/j.techfore.2020.120480

Stokes, D., & Lomax, W. (2002). Taking control of word of mouth marketing: the case of an entrepreneurial hotelier. *Journal of Small Business and Enterprise Development, 9*(4), 349–357. https://doi.org/10.1108/14626000210450531

TARP. (1985). *Consumer Complaint Handling in America: An Update Study*. Technical Assistance Research Programs Institute.

Titov, I., & McDonald, R. (2008). Modeling online reviews with multi-grain topic models. In *Proceedings of the 17th International Conference on World Wide Web* (pp. 111–120). Beijing: ACM. https://doi.org/10.1145/1367497.1367513

Tseng, C. H., & Wei, L. F. (2020). The efficiency of mobile media richness across different stages of online consumer behavior. *International Journal of Information Management, 50,* 353–364. https://doi.org/10.1016/j.ijinfomgt.2019.08.010

Vardakas, J. S., Zorba, N., & Verikoukis, C. V. (2014). A Survey on Demand Response Programs in Smart Grids: Pricing Methods and Optimization Algorithms. *IEEE Communications Surveys & Tutorials, 17*(1), 152–178. https://doi.org/COMST.2014.2341586

Vieira, V., Santini, F. O., & Araujo, C. F. (2018). A meta-analytic review of hedonic and utilitarian shopping values. *Journal of Consumer Marketing, 35*(4), 426–437. https://doi.org/10.1108/JCM-08-2016-1914

Wang, X., Yang, Z., & Liu, N. R. (2009). The Impacts of Brand Personality and Congruity on Purchase Intention: Evidence From the Chinese Mainland's Automobile Market. *Journal of Global Marketing, 22*(3), 199–215. https://doi.org/10.1080/08911760902845023

Wei, W., Liu, H., He, J., Yang, H., & Du, X. (2008). Extracting feature and opinion words effectively from chinese product reviews. In *Proceedings of the 5th International Conference on Fuzzy Systems and Knowledge Discovery* (Vol. 4, pp. 170–174). Shandong: IEEE Computer Society.

West, R., & Turner, L. (2007). *Introducing Communication Theory*. New York, NY: McGraw Hill.

Wong, T.-L., & Lam, W. (2008). Learning to extract and summarize hot item features from multiple auction web sites. *Knowledge and Information Systems, 14*(2), 143–160. https://doi.org/10.1007/s10115-007-0078-2

Yan, Z., Xing, M., Zhang, D., & Ma, B. (2015). EXPRS: An extended pagerank method for product feature extraction from online consumer reviews. *Information & Management, 52*(7), 850–858. https://doi.org/10.1016/j.im.2015.02.002

Zeithaml, V. A., Parasuraman, A., & Berry, L. L. (1992). Strategic positioning on the dimensions of service quality. *Advances in Services Marketing and Management, 2,* 207–228.