Print ISSN: 1738-3110 / Online ISSN 2093-7717
JDS website: http://www,jds.or.kr/
http://dx.doi.org/10.15722/jds.19.4.202104.25

Topics and Sentiment Analysis Based on Reviews of Omni-Channel Retailing*

Soon-Hong KIM [1], Byong-Kook YOO[2]

Received: January 13, 2021. Revised: March 04, 2021. Accepted: April 05, 2021.

Abstract

Purpose: This study aims to analyze the factors affecting customer satisfaction in the customer reviews of omni-channel, posted on Internet blogs, cafes, and YouTube using text mining analysis. **Research, data, and Methodology:** In this study, frequency analysis is performed and the LDA (Latent Dirichlet Allocation) is used to analyze social big data to respond to reviewers' reaction to the recently opened omni-channel shopping reviews by L Shopping Company. Additionally, based on the topic analysis, we conduct a sentiment analysis on purchase reviews and analyze the characteristics of each topic on the positive or negative sentiments of omni-channel app users. **Results:** As a result of a topic analysis, four main topics are derived: delivery and events, economic value, recommendations and convenience, and product quality and brand awareness. The emotional analysis reveals that the reviewers have many positive evaluations for price policy and product promotion, but negative evaluations for app use, delivery, and product quality. **Conclusions:** Retailers can establish customized marketing strategies by identifying the customer's major interests through text mining analysis. Additionally, the analysis of sentiment by subject becomes an important indicator for developing products and services that customers want by identifying areas that satisfy customers and areas that evoke negative reactions.

Key words: Omni-channel, Customer Satisfaction, TOPIC, LDA (Latent Dirichlet Allocation), Emotion Analysis

JEL: L10, L81, L86, M15, M30

1. Introduction

With the rapid development of IT technology, we are living in the era of big data consumption. Due to technological advancements such as AI and IoT, omni-channel distribution channels are also attracting attention. An omni-channel is a new type of distribution system that combines online and offline distribution channels such that information identified through an online shopping mall can be applied in an offline store and vice versa to buy products. As non-face-to-face contact has become a part of daily life due to the recent COVID-19 situation, purchases in offline stores such as department stores and hypermarkets have decreased, and online shopping has been increasing rapidly. According to statistics released by the Ministry of Trade, Industry and Energy in March 2020, the percentage of online retailers among the total sales of the retail industry in February 2020 (10,600 billion won) was 49.0%, an increase of 9.2% compared to the same month last year as of February.

In omni-channel research, analyzing data through questionnaire surveys is the most dominant method. However, in the Internet era, it is difficult to accurately reflect the rapidly changing directions of large-scale customers across various online channels, with the number of samples reaching several hundred. In recent years,

* This study was supported by the Incheon National University research grant in 2017.
This study was revised by and supplemented paper published at summer conference of IPACT 2020.
1 First Author, Professor. Division of International Trade, Incheon National University, Korea, Email: snow8817@inu.ac.kr
2 Corresponding Author. Professor, Division of International Trade, Incheon National University, Korea, Email: bkyoo@inu.ac.kr

research on the analysis of user reviews has been actively conducted in the fields of product reviews, tourism and hotel (Pearce & Wu, 2015; Cenni & Goethals, 2017), but there are not many studies on social big data analysis in the omni-channel field. A review of the extant literature related to the satisfaction of omni-channel shopping use shows that economic value, satisfaction with product quality, brand awareness, delivery and customer relationship management services are suggested as important factors for the purchase satisfaction of customers (Verhoef, 2012; Han, 2017; Ryu, 2019). This study is significant because it analyzed unstructured data through social big data analysis, unlike regression analysis and structural equations that use surveys, and determined the satisfaction level and complaints of omni-channel customers. Analyzing various opinions of customers directly using unstructured data, such as the number of reviews and recommendations from various portals, is a new merit for social big data analysis. Therefore, the purpose of this study is to analyze sentences and words appearing in reviews by using text mining analysis techniques targeting the reviews of omni-channel customers posted on Internet blogs, cafes, and YouTube as well as what factors may affect customer satisfaction or dissatisfaction.

Text mining is the process of finding useful information from unstructured text data or partially structured, from Internet websites, blogs, comments, reviews, and consumers' shopping experiences (Gupta & Lehal, 2009). Beginning with information extraction such as text summary and document retrieval, text mining techniques include various analysis methods such as text classification, document clustering analysis, and sentiment analysis (Khan & Kanth, 2016). Kim, Park and Suh (2018) conducted frequency analysis through unstructured text analysis, and examined civil complaints using design thinking. In this study, the LDA (Latent Dirichlet Allocation) technique, the most widely used social big data technique, is utilized to respond to reviewers' reaction to the recently opened omni-channel shopping reviews by L Shopping Company. Additionally, based on topic analysis, we conduct sentiment analysis on purchase reviews and analyze the characteristics of each topic on the positive/negative sentiment of omni-channel app users.

2. Theoretical Background

2.1. Concepts of Omni-channel and Cases of Domestic and Overseas Establishment of Omni-channel

While a multi-channel simply means that a single retailer operates with multiple distribution channels, an omni-channel is differentiated from a multi-channel in that the omni-channel strategy is not simply to increase the number of channels but also to optimize the customer's overall shopping process itself (Rigby, 2011). An omni-channel means that customers can always buy the same product at the same price and promotion, whether online or offline. The prominent feature of an omni-channel is that the information identified through the online shopping mall can be applied to the offline store and vice versa such that customers can purchase products and obtain information under the same conditions without discrimination when purchasing products.

The following are examples of successful establishment of omni-channels in Korea and abroad. In the case of Wal-Mart in the U.S., sales increased by more than 10% after an omni-channel service called 'Click & Collect' was developed following COVID-19. In Korea, Lotte Shopping has been providing, following the outbreak of COVID-19, an omni-channel (O4O: Online for Offline) distribution system (LOTTE ON) was introduced in April 2020. Shinsegae Distribution is actively promoting online shopping through SSG. The delivery function is being reinforced with the goal of ‘same day delivery’ and ‘on-time delivery.’

2.2. Previous Studies on Satisfaction-related Factors of Omni-channel Shopping

Studies on omni-channel have recently been actively conducted. Avery, Steenburgh, Deighton and Caravella (2012) and Verhoef (2012) have been conducting research on collision and integration with existing channels. In addition, there are studies on the concept of omni-channels and customer satisfaction (Bhalla, 2014; Bell, Gallino, & Moreno, 2014; Park, 2016)

The following is a summary of the academic terms related to the omni-channel satisfaction factors. Customer satisfaction does not simply refer to the satisfaction that consumers feel when purchasing products or services; it refers to the overall satisfaction that consumers expect from their consumption experiences accumulated from past experiences(Oh, & Cheon, 2018;-Hoang, Chi., Thi., & Thi. 2020). Homburg, Vollmayr and Hahn (2014) also defined “customer satisfaction as evaluation after consumption for quality perceived in response to quality expected before consumption.” The antecedent factors for customer satisfaction include perceived quality, product experience, perceived value, service, and brand reputation. Aaker (1991) also stated that perceived quality is consumers’ overall evaluation of the quality of products or services (Wang & Yum, 2015; Yang & Shim, 2018). Evaluating quality involves intrinsic factors, such as product function, convenience, information and design, and extrinsic factors, such as price, brand image, and store image (Won & Kim, 2020).

Perceived value can be defined as an evaluation of the value of the benefit utility of a product attribute, not just a sense of satisfaction with the physical characteristics or objective value of the product itself that consumers feel when purchasing a product (Anderson & Srinivasan, 2003; Niu & Lee, 2018). In other words, when consumers judge that the quality of the product is higher than expected compared to the cost they paid when using the product, they view it as having high value (Kim & Shim, 2017). Among the previous studies on the factors of omni-channel customer satisfaction, Ryu (2019) conducted a study on the effect of fashion innovation, technological innovation, and fashion purchase immersion on consumers' omni-channel shopping intentions. Trong, Truc, Phuong, and Hoang (2020) stated that the perceived convenience and usefulness of omni-channels affect their use as well as customers' satisfaction with them. Additionally, in an analysis of fashion products, Verhoef, Scott, Neslin, and Vroomen (2007) stated that search convenience, service quality, purchase risk and enjoyment have a positive effect on muti-channel shopper (Zhang et al., 2010). Using the Unified Theory of Acceptance and Use of Technology (UTAUT) model, Lee, Becker, and Potluri (2019) analyzed that the convenience of use, usefulness, social impact, and promotion conditions affect omni-channel acceptance intention (Yim & Han, 2016).

Omni-channels have now become an important research topic for distributors. A review of the extant literature related to the satisfaction of omni-channel shopping use shows that economic value, the ease of use of online apps, brand awareness, satisfaction with product quality, delivery, and CRM such as coupons are suggested as important factors for the satisfaction of customers.

2.3. Topic Model and Sentiment Analysis

2.3.1. Topic Model and LDA

Topic modeling is a text mining technique that automatically extracts clusters related to a specific topic based on the frequency of words within the sentence from text data. In this case, one topic can be regarded as a set of words having a similar meaning. In addition, topic model includes the feature that one document can have several topics.

LDA estimates both the distribution of words for each topic and the probability distribution of topics for each document by calculating the probability of each word in a document being included in a specific topic (Blei & Jordan, 2003). The LDA analysis procedure is outlined below. The analyst first sets the number of topics k for the topic to be analyzed. On the second step, when the number of topics k is entered, it is assumed that k topics are distributed throughout documents d. In the third step, all words are assigned to one of k topics. Also, one topic out of k topics is randomly assigned to every word in all documents. On the final step, each document has a topic, and the topic extracts the word distribution. The analysis procedure of LDA can be summarized as shown in Figure 1.

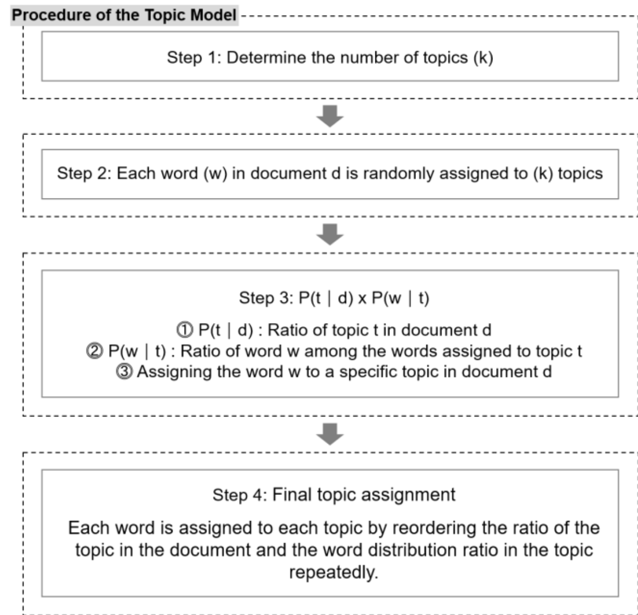

Figure 1: Procedure of LDA Model

And LDA analysis is assumed to follow the Dirichlet probability distribution. Dirichlet distribution is a type of multinomial distribution in which the binomial distribution is expanded to the number of N cases (Blei & Jordan, 2003).

The number of topics must be determined first when using the LDA topic model. In previous studies, several methods have been proposed to determine the number of topics. One method involves randomly selecting several topics until the cluster classification is clear. To extract the number of topics using this method, the theoretical basis should be supported through a review of previous studies on related research topics (Grant, Cordy, & Skillicorn, 2013). Cao, Xia, Li, Zhang, and Tang (2009) proposed a second method involving minimizing the relevance of topics by calculating the distance between topics through visualized data to select an appropriate number of topics. In addition, there are various statistical verification methods to determine the optimal number of topics.

Studies on the topic model are as follows, Pearce and Wu (2018) carried out the topic analysis of the reviews from foreign tourists in China. Cenni, and Goethals (2017) conducted topic and sentiment analyses to determine whether there were differences among different foreign languages in TripAdvisor's negative hotel reviews (Gang & Hong, 2017).

2.3.2. Sentiment Analysis

The sentiment analysis is a method of discriminating subjective features in languages such as blogs, news, product reviews, and SNS (Social Network Service) of Internet portal sites using the text mining technique by quantifying the degree of subjectivity. Recently, research has been conducted in various fields, such as improvement of customer service, by analyzing customer satisfaction at the level of marketing strategy, such as through social issues or product reviews.

The sentiment analysis method is mainly divided into supervised learning and unsupervised learning. The supervised learning technique is a method of classifying the level of the buyers' feelings into synonym groups using a sentiment dictionary when consumers write their purchase reviews, while the unsupervised learning analysis method is a technique to determine whether the sentiment of review sentences or documents is simply positive or negative (Ravi & Ravi, 2015; Duan, Yu, Cao, & Levy, 2016).

Sentiment analysis can also be classified into document-level, sentence-level, and topic-level analysis. Document-level sentiment analysis is a method of classifying documents that indicate positive or negative sentiment according to emotional expression based on specific terms in the document, sentence-level sentiment analysis is a method of analyzing positive or negative emotions of sentences contained in a document, and topic-level sentiment analysis is a method of analyzing positive or negative documents according to the sentences included in various topics included in the document. Based on this trend from previous studies, this study also attempts sentiment analysis based on the topic model. Sentiment analysis using the topic model can derive more detailed information about customer satisfaction or reviews of specific products.

Introducing previous research on sentiment analysis, Duan et al. (2016) conducted a sentiment analysis on complaints about reviewers with low star ratings, targeting reviews including star ratings in a study on hotel ratings. Lee and Lee (2020) analyzed smart tourism awareness through big data analysis. Liu (2015) suggested that sentiment analysis includes sentiment analysis at the document level, at the sentence level, and at the topic level. Gang and Hong (2017) employed sentiment analysis with online destination images based on the topic. Research related to topic-based sentiment analysis is also on the rise (Marrese-Taylor, Velasquez, & Bravo-Marquez, 2014; Nguyen, Shirai, & Velcin, 2015**).** In the marketing sector, the results of the sentiment analysis were also related to satisfaction or preference to improve services.

3. Research Methodology

3.1. Research Questions

Even if customers generally show favor in product purchase reviews, the sentiment for each topic may differ, such as that the price was low, but the service was not good. Therefore, to understand customers' satisfaction with product purchase in depth, basic sentiment analysis on each topic is required in addition to comprehensive evaluation. In this study, the LDA model is used for the topic analysis, and sentiment analysis is performed on the sentence and document units based on the extracted topics. Therefore, we analyze the following research tasks:

Research question 1: What are the main keywords related to the reviews of omni-channel shopping app users that were recently collected and analyzed by NAVER, Daum, and YouTube?

Research question 2: Into what topics can the keywords related to customer satisfaction factors for omni-channel shopping recently collected and analyzed by NAVER, Daum, and YouTube be classified, and what are the characteristics of each group?

Research question 3: What are the characteristics of each topics sentiment analysis for topics recently collected and analyzed by NAVER, Daum, and YouTube?

3.2. Research Method and Procedure

3.2.1. Data Collection and Analysis Target

This study analyzes the factors that customers consider important when using omni-channels with a focus on the review data posted by customers who have used omni-channel apps on NAVER, Daum, and YouTube through keyword analysis related to customer satisfaction. We analyzed the omni-channel usage review data of L Shopping company, which built an on-offline integrated omni-channel app. The collection period was based on user reviews for approximately 6 months from the end of April to October 31, 2020, when L Shopping established an omni-channel app, and the collected data were customer reviews posted on Naver, Daum, and YouTube, the major portal sites in Korea.

To collect texts from major portals, unstructured text mining data were collected for each period using TEXTOM, an online big data service solution. The amount of data collected was 1782, including 483 Naver Blog, 280 Naver Cafe, 465 Daum Blog, 4 Daum Cafe, and 550 YouTube video posts, totaling 692 KB (Table1).

Table1: Amount of Data Collected on Omni-channel Reviews.

Channel	Section	Collection amount (case)	Volume (%)
NAVER	blog	483	27.2
	Cafe	280	15.7
Daum	blog	465	26.1
	Cafe	4	0.2
Google	YouTube	550	30.8
Sum		1,782	692KB

3.2.2. Data analysis procedure

The process of LDA analysis is as follows:

1: Collect and accumulate sentences from L Shopping 's omni-channel app user reviews posted on the blogs and cafes of NAVER, Daum, and YouTube.

2: We perform the frequency analysis to derive important keywords, based on the extracted words, and extract topics on the satisfaction of omni-channel reviews.

3: Based on the derived topic analysis, we measure the reputation of each topic in the omni-channel app reviews through sentiment analysis**.**

The research procedure is shown in Figure 2 below.

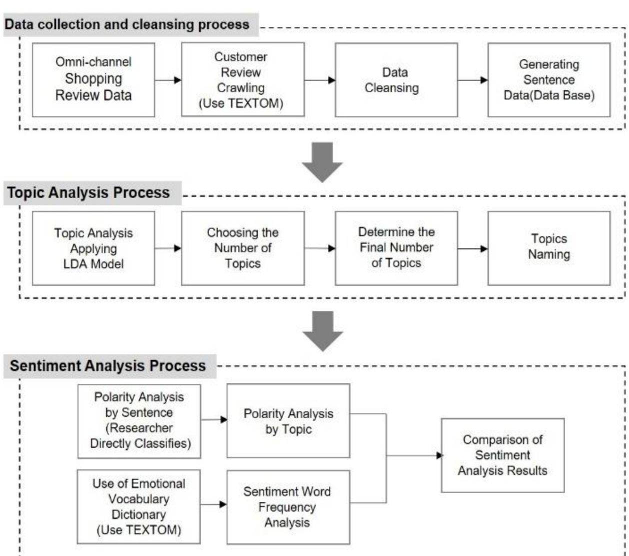

Figure 2: Research Progress Model

4. Results of the Analysis

First, for the data refining process, sentences posted on the Internet were targeted, and the title and posted text were examined for text mining analysis. When explaining the morpheme analysis method, the language used for analysis was Korean, and the subject of analysis included common nouns, proper nouns, dependent nouns, and pronouns.

4.1. Results of TF-IDF Analysis

First, an analysis of the frequency of key words related to omni-channel shopping app reviews was conducted. As an analysis method, we analyzed the reviews of users of the omni-channel shopping app for L Shopping and derived relevant keywords to analyze what factors are important in relation to consumer interest and customer satisfaction in omni-channel apps.

A word with a larger TF value can be considered to have higher importance. However, a word with a large TF value may be a word that frequently appears in all documents, that is, a word having a large 'Document Frequency (DF)' value. To filter out words that frequently appear in all documents, the 'Term Frequency-Inverse Document Frequency (TF-IDF)' technique is used. This is not simple word frequency processing but, rather, reprocessing of the appearance frequency based on the probability of the appearance of the word. Among the nouns derived for this frequency analysis, proper nouns such as the name of a specific company, incomplete nouns such as '~jeok(thing), ~'bun', '~ba' were excluded.

The frequency analysis results are presented in Table 2. The contents commonly frequently mentioned in reviews by consumers who use L Shopping 's omni-channel app include words such as reviews, purchase, recommendation, product, event, delivery, price. Overall, in this TF-IDF analysis, words related to the factors mentioned in the literature study on the satisfaction of omnichannel use were extracted. The analysis result of TF also showed similar results, but the event was ranked 8th in the frequency analysis, while the event was ranked 5th in the TF-IDF analysis. In addition, the brand ranked 18th in the frequency (TF) analysis but 12th in the TF-IDF analysis, showing some differences in ranking. TF-IDF is a value obtained by multiplying TF, which represents the importance of words in a specific document, by IDF, which is the importance of words related to the entire document, and the higher the frequency of words in a specific document and the fewer documents including the word among all documents, the higher the value.

Table 2: Result of TF-IDF Analysis of Keywords from Customer Reviews about L Shopping 's Omni-channel Shopping App

Rank	Words	TF	Rank	Words	TF--IDF
1	Reviews	538	1	Reviews	665.28
2	Purchase	313	2	Purchase	548.13
3	Recommendation	119	3	Recommendation	374.17
4	Products	118	4	Products	339.63
5	Price	115	5	Event	332.59
6	Discount	107	6	Delivery	304.27
7	Buy	106	7	Today	300.91
8	Event	90	8	Price	298.58
9	Delivery	87	9	Discount	295.15
10	Prepare	78	10	Offline	271.41
11	Posting	77	11	Distribution	270.60
12	Goods	72	12	Brand	250.67
13	Distribution	69	13	Video	249.92
14	Today	69	14	Goods	248.14
15	Offline	65	15	Online	248.14
16	Use	63	16	Coupang	242.30
17	Order	61	17	Use	235.49
18	Brand	57	18	Prepare	230.41
19	Size	56	19	Order	229.99
20	Gift	54	20	L Department	229.38

4.2. Topic Model Analysis

In Figure 3, each circle shows one topic in the derived topic figure. The size of the circle indicates the relative importance of the topic in the analysis data, and the number displayed in the circle is the number of the analyzed topics. The location of each topic is expressed by similarity distance according to multidimensional scaling, which is useful for identifying relationships, differences, and similarities among topics. The multidimensionality of several topics is represented on a two-dimensional (PC1, PC2) plane such that the topics belonging to the same quadrant are semantically related.

As a result of topic analysis using the LDA analysis method, 10 topics were first derived, as shown in Figure 3. However, in the results of this topic analysis, factors for the satisfaction of omni-channel use are not clearly presented. Considering the preceding studies on omni-channel satisfaction factors, this study reduced the 10 topics so that word groups for each topic were clearly separated. Table 3 shows the words drawn from 10 topics. As a result of reducing the topics until the classification of each topic group was clearly separated, four models were finally selected in which the topic classification was relatively clearly separated. Figure 4 shows the result of reducing the topics to four. Table 3 shows the results of the first topic analysis of omnichannel user reviews.

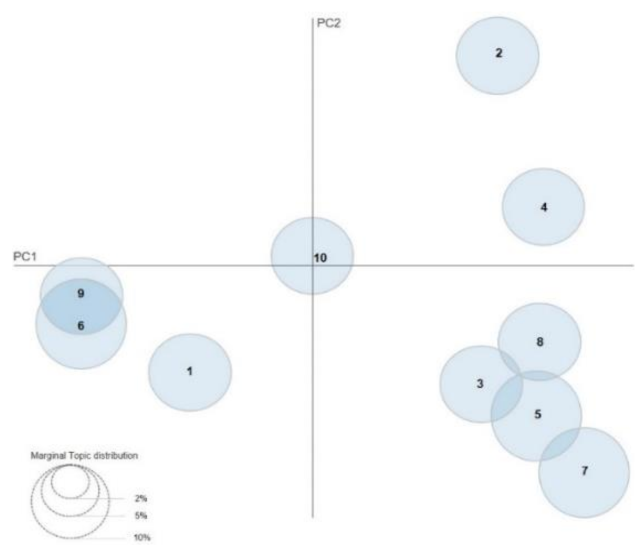

Figure 3: Analysis of the First Topic from Omni-channel User Reviews

Table 3: The Results of Analysis of the First Topic of Omni-channel User Review

Topic number	Topic Category	Key words
1	**Brand, Product and Delivery**	review, recommendations, brand, products, shopping, offline, design, delivery, distribution, event.
2	**Review and Recommendation**	review, purchase, product, summer, payment, online, recommendation, discount, after purchase, what I bought with my money.
3	**Discount and Event**	purchase, review, event, discount, video, after purchase, use, delivery, recommendation, discount coupon.
4	**Review and Use**	review, purchase, videos, recommendation, orders, use, today, products, products, brand.
5	**Purchase and Price**	review, purchase, products, goods, recommendations, online, Coupang, NIKE, price, gifts.
6	**Delivery and Event**	review, offline, event, wallet, delivery, distribution, community, 2 hours, Starbucks, M coupon.
7	**Price and Product**	price, 11th Street, purchase, Auction, products, Halloween, buy, gift, goods, posting.
8	**Products and Brand**	review, beauty, products, video, recommendation, brand, discount, purchase, bag, payment.
9	**Delivery and Event**	review, delivery, community, hint chain, hammock app, event, L pay, Starbucks, 2 hours, distribution.
10	**Purchase and Event**	review, purchase, event, price, recommendations, discounts, masks, what I bought with my money, video, sales.

Table 4 shows the results of the final topic analysis of L Shopping's omnichannel user reviews. As a result of analysis, four major topics were derived. Topic 1 includes words such as review, delivery, offline, event, distribution, purchase, community, wallet, Starbucks coupon, and address. This topic consists of words related to delivery and CRM (Customer Relationship Management) services. This topic corresponds to the first topic analysis groups No.1, No.6 and No.9.

Topic 2 is composed of words such as purchase, review, products, post-purchase, discount, price, online, payment, point, and size. It is mainly composed of words related to economic reputation. This topic corresponds to the first topic analysis groups No.3, No.5 and No.7. Topic 3 contains words such as purchase, review, recommendation, video, use, gift, discount, writing, preparation, and event. It is composed of reviews and topics related to purchasing experience and corresponds to the first topic analysis groups No.2 and No.4. Topic 4 is mainly composed of words related to product quality and brand evaluation such as purchase, quality, recommendation, products, order, brand, use, sold out, L department store, and sales and corresponds to the first topic analysis groups No.8 and No.10. When naming the four topics, we named Topic 1 the delivery and CRM topic, Topic 2 the economic value topic, Topic 3 the review and use experience topic, and Topic 4 the product quality and brand reputation topic. Based on these four topics, we conducted sentiment analysis of each topic and analyzed their features. This topic analysis allows for derivation of more detailed information about customers' responses to product purchases by analyzing the reviews of customers who use purchasing apps.

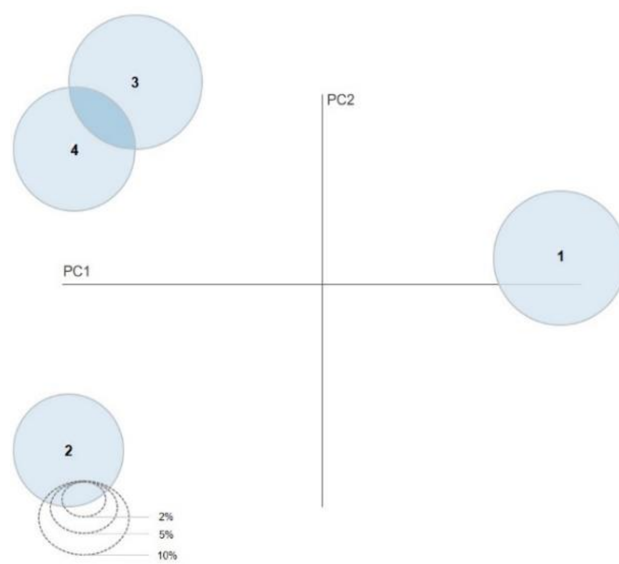

Figure 4: Results of Analysis of the Final Topics from Omni-channel User Reviews

Table 4: Results of the Final Topic Analysis: Four Topics and Words Related to Each Topic

Topic number	Topic name	Keywords
1	**Delivery and CRM**	reviews, delivery, offline, event, distribution, purchase, community, wallet, Starbucks coupon, address.
2	**Economic Value**	purchase, review, products, post-purchase, discount, price, online, payment, point, size
3	**Reviews and Use Experience**	purchase, review, recommendation, video, use, gift, discount, writing, preparation, event
4	**Product Quality and Brand Reputation**	purchase, quality, recommendation, products, order, brand, use, sold out, L department store, sales

4.3. Sentiment Analysis

In this study, polarity sentiment analysis was conducted to identify positive and negative contexts based on the corpus of sentences in the document for the reviews from customers who used omni-channel online transactions of L Shopping. Through the analysis of four topics relating to the satisfaction factor of L Shopping 's omni-channel app user reviews, a sentiment analysis with a focus on positive and negative contexts was conducted for 1782 comment sentences and related site documents.

To identify positive and negative sentences, the sites related to each sentence were checked to derive to which topic the main words in the review document belong and then classified by whether the derived sentence context was positive or negative. The sentiment dictionary provided by TEXTOM was then used to compare and analyze the ratio of positive and negative sentences. The result of the verification showed that there was some difference between the positive and negative results in the sentiment dictionary and the actual site verification results.

The result of polarity sentiment analysis showed that there are 831 positive sentences and 81 negative sentences. Documents in which the same sentence overlapped several times or which had neutral content, general publicity articles, and product ad sentences or app ads used by L Shopping itself were excluded from the sentiment analysis. The derived analysis results can be expressed as the following positive and negative indexes:

$$\text{Positive index} = \frac{\text{NP}}{\text{NP+NN}} \times 100 \quad (1)$$

$$\text{Negative index} = \frac{NN}{NP+NN} \times 100 \quad (2)$$

NP: number of positive cases,
NN: number of negative cases

Table 5: Results of Sentiment Analysis for Each Topic Unit (case, %)

Category	TOPIC 1 (Delivery and CRM)	TOPIC 2 (Economic Value)	TOPIC 3 (Reviews and Use Experience)	TOPIC 4 (Product Quality and Brand Reputation)	Total
Positive	126 (87.5%)	356 (97.27)	216 (86.75)	139(86.88)	837(91.01)
Negative	18 (12.5)	10 (2.73)	33 (13.25)	21(13.12)	82 (8.99)
Total	144	366	249	160	919
Neutral & (N. A.)					863
Total					1782

Table 5 shows the result of the sentiment analysis for each topic unit. The overall analysis result for the positive index was found to be 91.01%, and that for the negative index was 8.99%. In the case of the positive/negative index by topic, it was found that the positive index was 87.5% and negative index was 12.5% for Topic 1, the positive index was 97.27% and negative index was 2.73% for Topic 2, the positive index was 86.75% and negative index was 13.25% for Topic 3, and the positive index was 86.88% and negative index was 13.12% for Topic 4.

The sentiment index results can be summarized as follows. For topic 2 (the economic value topic), the reviews were relatively very satisfactory compared to other topics in terms of economic feasibility, such as discounts and points, while for topic 1 (the delivery and CRM), there were many complaints about delivery accidents. For Topic 3 (the reviews and use experience), there were negative reactions to bad reputation reviews written by others and negative effects from factors such as inconvenience of using the app itself, while in the case of Topic 4 (the product quality and brand awareness topic), there were negative reputation factors due to not being offered the expected brand or product.

Table 6: The Result of the Sentiment Analysis Using the Sentiment Dictionary

Detailed Sentiment Words	Detailed Sentiment Frequency (cases)	Detailed Sentiment Ratio (%)
Crush	989	65.95
Joy	141	8.65
Interest	140	8.27
Sadness	72	5.25
Fear	51	2.6
Rejection	123	6.61
Anger	18	1.06
Surprise	14	1.04
pain	11	0.57
Positive	1270	81.46
Negative	289	18.54
TOTAL	1559	100%

Table 6 shows the results of sentiment analysis using the sentiment dictionary. The result of the sentiment analysis using the sentiment dictionary with TEXTOM showed that the positive index was 81.46% and the negative index was 18.54%, a somewhat different result from the sentence and document analysis of this study, which suggests that there is a limitation that the context of the entire document or sentence cannot be filtered out in the mechanical text mining analysis. This is a case of reflecting only the results of sentiment word analysis in a simple sentence. For example, there are many examples of sentences such as "I was satisfied with the price and product after the actual purchase even though the reviews were not good."

5. Conclusion

5.1. Summary of Findings–

The significance of this study of omni-channel use is that this is a new attempt on omni-channel customer satisfaction research by comparing and analyzing the results of quantitative studies on the satisfaction factors of omni-channel use and qualitative data of omni-channel.

The differences between social big data analysis and existing studies are as follows.

First, through social big data analysis, unlike conventional survey analysis, the study can accurately reflect purchase desires across various channels of consumers as unstructured data has a lot of samples. Second, social big data analysis can quantify qualitative data such as reviews to some extent and reflect it in the marketing strategies of companies. Third, social big data analysis techniques such as sentiment analysis can identify the polarity of the positive or negative aspects of omni-channel app users, so that retailers can immediately see which areas consumers are satisfied and dissatisfied with.

The summary of the research results is as follows. As a result of topic analysis, four main topics are derived: delivery and events, economic value, recommendations and convenience, and product quality and brand awareness. The emotional analysis reveal that the reviewers have many positive evaluations for price policy and product promotion, but negative evaluations for app use, delivery, and product quality.

5.2. Implications

Social big data analysis can accurately identify consumers' reactions to each aspect, such as price, delivery, event, quality, and brand, by analyzing comments and reviews from online commerce. Companies can establish customized marketing strategies by determining the major interests of customers in real time through TF-IDF analysis and TOPIC analysis by topic. In addition, the sentiment analysis for each topic becomes an important index for developing products and services that customers want by more closely identifying which areas customers are satisfied with and toward which areas they exhibit negative reactions.

Policy implications based on the results of this study are as follows:

First, omni-channel distribution companies should strengthen their customer promotion strategies through events such as product promotion videos screened at live shows and coupon issuance to customers, rather than simply reducing prices. Second, when customers purchase products, reviews posted by others become an important factor. Additionally, consumers themselves also participate directly in the reviews and influence others. Therefore, retailers must implement a strategy for maintaining customer relationships to manage their product reviews and to ensure that customers continuously maintain their products. Third, complaints about omni-channel apps use and product shortages were found to be major complaints of customers in sentiment analysis. Therefore, omni-channel developers should develop apps that customers can use more conveniently and quickly and ensure that there is no shortage of stocks for products ordered by customers. In sum, in the era of big data, it is necessary to increase the loyalty of commenters using omni-channels, so that comments recommended by commenters with high loyalty to a product increase the trust of other customers and induce re-purchase. As offline commercial transactions contract in response to the COVID-19 outbreak, each distributor introduces an omni-channel system, which is becoming a means to compensate for these shrinking offline transactions. It is now apparent that Korea is in the early stage of omni-channels. We hope that the introduction of big data technology will be actively utilized in the distribution field to activate products, transaction services, and customer relationship management policies that can further satisfy customers using omni-channels.

5.3. Limitations of Research

As LDA topic analysis is a mechanical method that extracts through the probability of words in a sentence, there is a limitation in that the entire contents of the purchase review document cannot be completely reflected. There are cases where multiple topics are mixed within the app review sentence, and the same word may even be included across multiple topics depending on the context. For example, words such as reviews, discounts, and purchases are duplicated in several topics. In the case of sentiment analysis, there were cases in which the reviews themselves, such as those in blogs and cafes, were converted into an event by the company, and in some cases, the text writing was done for payment, but there was a limitation in which it was difficult to completely filter these out from the collected data. In this study, the results of sentiment analysis using a sentiment dictionary showed somewhat different from the sentence and document analysis of this study. This indicates that there is a limitation in that the entire context or sentence cannot be filtered out in text mining analysis.

As revealed in the limitations of the above analysis, LDA analysis/sentiment analysis is a qualitative and technical analysis that analyzes unstructured data such as purchase reviews on the Internet based on word frequency. It cannot derive a quantitative causal result regarding the satisfaction factors of consumers' choices for app user reviews. Since this study is a qualitative analysis using social big data, it is not possible to directly secure the validity and reliability of the variables of quantitative analysis. To compensate for this, the topic analysis tried to secure the validity of the factors by using the proximity of the distance of the multidimensional scaling method and, also tried to secure the validity by comparing and confirming the customer satisfaction preceding variables of the existing theoretical studies. As for reliability, the positive and negative ratio of the sentiment analysis for each topic is much higher than the number of general questionnaires, therefore we think it can cover the for reliability of the customer's response to the group classification of the customer satisfaction level of each topic. In the future research, we will try to perform a comparative analysis using big data as well as a structured questionnaire analysis.

References

Aaker, D. A. (1991). Managing Brand Equity. *The Free Press*, New York.

Anderson, R. E., & Srinivasan, S. S. (2003). E‐ satisfaction and E‐ loyalty: A Contingency Framework. Psychology & marketing, *20*(2), 123-138.

Avery, J., Steenburgh, J., Deighton, T. J., & Caravella, M. (2012). Adding Bricks to Clicks: Predicting the Patterns of Cross-channel Elasticities Over Time. *Journal of Marketing, 76*(3), 96-111.

Bell, D. R., Gallino, S., & Moreno, A. (2014). How to Win in an Omni channel World. *MIT Sloan Management Review, 56*(1), 45-53.

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

Bhalla, R. (2014). The Omni-channel Customer Experience: Driving Engagement Through Digitisation. *Journal of Digital & Social Media Marketing*, *1*(4), 365-372.

Cao, J., Xia, T., Li, J. X. J., Zhang. Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection, *Neuro computing*, *72*(7-9), 1775-1781.

Duan, W., Yu, Y., Cao, Q., & Levy, S. (2015). Exploring the impact of social media on hotel service performance: A sentimental analysis approach**.** *Cornell Hospitality Quarterly, 57*(3), 282-296.

Cenni, I., & Goethals, P. (2017). Negative hotel reviews on TripAdvisor: A cross-linguistic analysis. *Discourse Context Media, 16*, 22–30.

Gang R., & Hong T. H. (2017). Investigating Online Destination Images Using a Topic-Based Sentiment Analysis Approach, *Sustainability*, 1765 (1-18).

Grant, S., Cordy, J. R., & Skillicorn, D. B. (2013). Using heuristic to estimate an appropriate number of latent topics in source code analysis, *Science of Computer Programming, 78*(9), 1663-1678.

Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence*, *1*(1), 60-76.

Han, S. S. (2017). Effect on Brand Loyalty in Omni-Channel: Focus on Category Knowledge. *Journal of Distribution Science 15*(3), 61-72.

Hoang B. H. L., Chi T. N., Thi T. H. T., & Thi T. P. N. (2020). Factor Affecting Customers' Decision to Use Mobile Banking Service: A Case of Thanh Hoa Province, Vietnam. *Journal of Asian Finance, Economics and Business, 7*(2), 205-212.

Homburg, C., Vollmayr, J., & Hahn. A. (2014). Firm Value Creation Through Major Channel Expansions: Evidence from an Event Study in the United States, Germany and China. *Journal of Marketing, 78* (3), 38–61.

Khan, R. A., & Kanth, S. (2016). Text Mining: Knowledge Discovery from Unstructured data, Artificial Intelligent Systems and Machine Learning, *8*(2), 71-77.

Kim, H. S., & Shim, J. H. (2017). The Effects of Service Qualities on Customer Satisfaction and Behavioral Intention in Coffee Shops. *International Journal of Industrial Distribution & Business, 10*(10), 35-43.

Kim, T. H., Park, B. J., & Suh, E. K. (2018). Irregular Big Data Analysis and Considerations for Civil Complaint Based on Design Thinking. *International Journal of Industrial Distribution & Business, 9*(8), 51-60.

Lee, J. W., Becker, K., & Potluri, R. M. (2019). Antecedents of Corporate Adoption of Social Media and the Role of the Technology Acceptance Model in the Path. *The Journal of Asian Finance, Economics and Business, 3*(2), 67-76.

Lee, S. Y., & Lee H. S. (2020). A Study on the Smart Tourism Awareness through Bigdata Analysis. *Journal of Industrial Distribution & Business*, *11* (5), 45-52.

Liu B. (2015). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. *Cambridge University Press*, Cambridge.

Marrese-Taylor, E., Velasquez, J. D., & Bravo-Marquez, F. (2014). A Novel Deterministic Approach for Aspect-based Opinion Mining in Tourism Products Reviews. *Expert Systems with Applications, 41*(17), 7764-7775.

Niu, L. X., & Lee, J. H. (2018). The Intention of Repurchase on e-Service Quality by Online Travel Agency Site. *International Journal of Industrial Distribution & Business, 9*(7). 61-70.

Oh, H. S., & Cheon, H. S. (2018). The Impact of Nature of Purchase and Purchase Utility on Purchase Intention According to Retail entertainment. *Journal of Distribution Science, 16* (12), 57-68.

Park, S. H. (2016). Effect of Omni-Channel Use and Customer-Brand Relationship. *Journal of Distribution Science, 14*(11), 129-138.

Pearce, PL., & Wu, M. Y. (2018). Entertaining International Tourists: An Empirical Study of an Iconic Site in China. *Journal of Hospitality & Tourism Research, 42*(5), 772-792.

Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems, 89*, 14-46.

Rigby, D. (2011). The Future of Shopping. *Harvard Business Review, 89*(12), 65–76.

Ryu, J. S. (2019). Consumer Characteristics and Shopping for Fashion in the Omni-channel Retail Environment. *Journal of Business, Economics and Environmental Studies, 9*(4), 15-22.

Trong N. P., Truc V. H., Phuong V. Le, & Hoang (2020). Factors Affecting the Behavioral Intention and Behavior of Using E–Wallets of Youth in Vietnam. *Journal of Asian Finance, Economics and Business 7*(10*)*, 295–302.

Verhoef, P. C., Scott A., Neslin, S. A., & Vroomen, B. (2007). Multi-Channel Customer Management: Understanding the Research Shopper Phenomenon. *International Journal of Research in Marketing, 24*(2), 129-148.

Verhoef, P. C. (2012) Multi-Channel Customer Management Strategy in Handbook of Marketing Strategy. Shankar V. and Carpenter G., eds. Cheltenham: Edward Elgar Publishing Limited.

Yang, S. K., Shim, J. H. (2018). The Effects of Service Qualities on Customer Satisfaction, Trust, and Behavioral Intention in Smartphone Shopping Malls. *International Journal of Industrial Distribution & Business. 9*(12). 31-43.

Yim, D. S., & Han, S. S. (2016). Omni-channel's Perceived Ease of Use and Perceived Usefulness Effect on Omni-channel Use and Customer-Brand Relationship. *Journal of Distribution Science, 14*(7), 83-90.

Wang, W., & Yum, C. S. (2015). The Effect of Service Quality on Customer Satisfaction of Smart Phone Banking. *Journal of KIIT, 13*(4), 95-102.

Won, J., & Kim, B. Y. (2020). The Effect of Consumer Motivations on Purchase Intention of Online Fashion - Sharing Platform. *Journal of Asian Finance, Economics and Business, 7*(6), 197-207.

Zhang, J., Farris, P. W., Irvin, J. W., Kushwaha, T. T., Steenburgh, J., & Weitz, B. A. (2010). Crafting Integrated Multi- Channel Retailing Strategies. *Journal of Interactive Marketing, 24*(2), 168-180.