[1]Jun Xie

[2,*]Fagang Hu

[3]Mengjie Zhang

Grounded Theory Analysis of Online Shopping Reviews Based on Text Clustering

JES

Journal of Electrical Systems

Abstract: - Grounded Theory Analysis of online shopping reviews can improve the quality and accuracy of review data analysis, providing valuable information and suggestions for relevant businesses and consumers. This article focuses on analyzing mobile phone reviews on JD.com as coding objects, effectively employing machine learning algorithms such as Latent Semantic Indexing (LSI), Latent Dirichlet Allocation (LDA), and text clustering in natural language processing to extract thematic keywords from online shopping review texts and perform axial coding. Subsequently, DBSCAN clustering algorithm is applied to categorize the main categories of design, service, and production under the perspective of businesses and consumers as consumption attraction, while categorizing price and brand under consumption intention, forming selective coding. By applying Grounded Theory to the analysis of online shopping reviews, this study accurately identifies consumer needs and feedback, with the aim of helping businesses improve product and service quality, enhance customer satisfaction and loyalty, and engage in personalized marketing and customized services.

Keywords: Grounded Theory (GT), DBSCAN Clustering, Online Shopping Review, Consumption Attraction, Consumption Intention.

I. INTRODUCTION

With the development of internet technology and social media, the volume of online shopping review data has been continuously increasing. Extracting accurate information and knowledge from a large amount of review data has become a challenging task. The words and expressions frequently used by consumers in online shopping reviews can often reflect their shopping experiences. By reading other consumers' online shopping reviews, individuals can gain a more comprehensive and objective understanding of the strengths and weaknesses of products and services, thus making better shopping decisions.

In recent years, the continuous advancement of artificial intelligence and natural language processing technologies has played a guiding role in various applications of online shopping review analysis, such as brand marketing, product promotion, and social sentiment analysis. The application of grounded theory in online shopping review analysis further enhances the quality and accuracy of review data analysis, providing consumers with useful information and recommendations.

Although the application of grounded theory in analyzing online shopping reviews has received increasing attention, previous literature surveys have shown that few studies have focused on how to promote the sustained development of online consumer behavior, lacking specific recommendations or effective methods to explore new paths for e-commerce marketing. This study applies grounded theory to analyze product reviews on internet e-commerce platforms and reveals consumers' values and knowledge structures regarding product performance, quality, price, and after-sales service. The analysis uncovers consumer needs and concerns regarding product performance, quality, and after-sales service, offering valuable insights for companies to gain a precise understanding of consumer needs and psychology. This information enables them to enhance the quality of their products and services accordingly.

To further enhance the analysis, incorporating elements of text clustering and data analysis methodologies could provide deeper insights into patterns and trends within online shopping reviews [1]. By leveraging computational techniques, such as text clustering algorithms, researchers can systematically categorize and analyze large volumes of text data, identifying common themes, sentiments, and opinions among consumers [2]. Integrating these computational approaches with grounded theory analysis can enrich the understanding of consumer behavior and preferences, enabling more targeted and effective strategies for e-commerce marketing.

II. LITERATURE REVIEW

• Research on Online Shopping Review Analysis: The literature on online shopping review analysis focuses on understanding user needs, the impact of reviews on consumer behavior, and the effectiveness of online

[1]Suzhou University, Suzhou, Anhui, China
[2]Suzhou University, Suzhou, Anhui, China
[3]Suzhou University, Suzhou, Anhui, China
* Corresponding author: Fagang Hu

recommendations. Wei et al. explored public culture cloud activities through online comments, emphasizing user engagement and content optimization [3]. Qiu et al. examined host-guest intimacy in Airbnb reviews, revealing its influence on user experience [4]. Studies by Park et al. and Alzate et al. investigated the role of review helpfulness and visibility in shaping purchasing decisions, highlighting the significance of review characteristics [5,6]. Racherla and Friske, along with Stöckli and Khobzi, delved into the perceived usefulness of reviews and the interplay between recommendation systems and review convergence, underscoring the importance of online reviews in decision-making processes [7,8]. This body of work collectively demonstrates the critical impact of online reviews in guiding consumer preferences and behaviors in the digital marketplace. Li et al.'s study underscores the significant role of user-generated images in enhancing customer engagement [9]. Cheng and Jin focus on the specific concerns of Airbnb users, offering insights into service expectations [10]. Song et al. present a method to quantify service failure risks from customer complaints, providing a unique perspective on addressing consumer concerns [11]. Kostyra et al. dissect the complex effects of online reviews on brand, price, and product attributes, highlighting the multifaceted influence of consumer feedback [12]. Together, these studies contribute to a broader understanding of the dynamics at play in online review systems and their impact on consumer behavior and business strategy.

- Text Clustering Techniques in Online Shopping Review Analysis: Previous research has employed text clustering algorithms to analyze online shopping reviews, revealing both the advantages and challenges associated with using text clustering in review analysis [13, 14]. Zhu et al. utilized a stream of consumers' text comments and overall ratings to obtain a corresponding stream of weights indexed by the comment time. This approach allowed them to capture the temporal dynamics of consumer feedback effectively [15]. In another study, Wang et al. applied the word2vec algorithm, grounded theory, and co-occurrence cluster method to analyze online shopping word-of-mouth texts in relation to consumer behavior theory [16]. Their research aimed to uncover meaningful patterns and associations in the text data, facilitating a deeper understanding of consumer behaviors.

- Integration of Grounded Theory and Text Clustering: Previous research has demonstrated the integration of Grounded Theory and text mining for comprehensive analysis, yielding valuable results and insights. Papathanassis and Knolle revealed that while reviews play a significant role, they are secondary and complement the decision-making process, which also heavily relies on conventional content [17]. Matteucci et al. explored the core principles, epistemological and ontological foundations of Grounded Theory, and its contribution to theory construction in qualitative tourism research [18]. Cheng et al. utilized the Grounded Theory process, including open coding, axial coding, and selective coding, to construct generation models of users' commentary behavior in multi-contextual public opinion [19]. In another study, Hou et al. proposed a method that combines Grounded Theory and multilayer perceptron neural networks to identify the usefulness of online reviews on e-commerce platforms. They suggested that incorporating feature extraction methods for consumer purchase decisions can enhance the effectiveness of classification, providing guidance and recommendations for businesses in managing online stores [20].

III. GROUNDED THEORY AND ITS RESEARCH STEPS

Grounded Theory (GT) is a qualitative research approach that aims to develop theory based on empirical data. Typically, there are no preconceived assumptions before the research begins. Instead, it starts with direct observation of the actual phenomena and induction of experiences from raw data, leading to the development of systematic theory. GT is a bottom-up approach to constructing substantive theory. It involves systematically collecting data and identifying core concepts that reflect the essence of phenomena, and then constructing relevant theoretical frameworks based on the relationships between these concepts. GT relies on empirical support, but its main feature lies not in empiricism, but in the abstraction of new concepts and ideas from empirical facts. GT method is philosophically based on a post-positivist paradigm and emphasizes the falsification of constructed theories. The systematic coding of data is a crucial step in Grounded Theory analysis, which includes three levels of coding: open coding, axial coding, and selective coding. The research step of this article is: data collection→ data preprocessing →data analysis (open coding→axial coding→selective coding)→conclusion presentation and application.

A. Open Coding

Open coding is a method used to study social phenomena, also known as "open-ended research", "open-ended questionnaire" or "free-response method". It differs from closed coding methods (such as multiple-choice questions or scales) in that it does not predefine the research subjects or content. Instead, it allows participants to freely

express themselves and provide more authentic and comprehensive information and data. The fundamental principle of open coding is to minimize restrictions on response formats and content, encouraging participants to express their views and feelings in their own language and thought processes. Grounded Theory researchers typically provide one or several open-ended questions or topics, record participants' responses, and then categorize, summarize, and analyze them. In the context of analyzing online shopping reviews, open coding involves reading and summarizing the text line by line or paragraph by paragraph, extracting common and distinctive language, concepts, and themes, and forming a conceptual system or theoretical framework to fully understand and explain the research subject.

B. Axial Coding

Axial coding is a quantitative research method that belongs to principal component analysis or factor analysis. Its basic idea is to reduce a large number of variables (such as questionnaire items) to a few "axes" (also known as principal components or factors) through mathematical techniques. This reduction helps eliminate redundancy and repetition among variables while preserving as much information and variation as possible. Axial coding can be used to explore the underlying structure and relationships of complex issues, facilitating a deeper understanding and explanation of research phenomena.

C. Selective Coding

Selective coding involves systematically analyzing the results of axial coding to select representative core categories and establish systematic connections between these core categories and other categories. In the process of selective coding, open coding is initially used to describe the text, and then relevant codes related to the research topic are selected for further summarization and abstraction to derive more generalized conclusions.

D. Implementation Steps

The specific implementation steps of GT research are as follows: Preparation → Data Collection → Data Preprocess-ing → Data Analysis → Theme Identification → Synthesis Evaluation → Conclusion Presentation.

Preparation: Clarify the research objectives and research subjects, determine the required data and materials for the study, collect relevant literature and information, and prepare the necessary tools and equipment for the research.

Data Collection: Collect relevant data and materials, such as interviews, questionnaires, observations, experiments, etc., for analysis and research. In this article, data collection was conducted in compliance with relevant laws, regulations, and website usage agreements, avoiding infringement of others' rights. Considering the difficulty and cost of data collection, mature collection methods and web scraping techniques were chosen. Consumer review data was obtained through API interfaces of online shopping platforms, Python-based web scraping programs, and other means. Product review data was collected, and some typical products were selected as research objects.

Data Preprocessing: Process the collected data by cleaning, filtering, and segmenting, to facilitate subsequent analysis.

Data Analysis: Based on the basic principles of grounded theory, analyze the collected data to extract key information and associated factors, and identify the underlying reasons and patterns behind the data. In this article, the review data was analyzed to identify themes, categories, and keywords, and analyze consumers' perceptions, attitudes, and behaviors related to these themes.

Theme Identification: Classify and organize the analyzed information and factors, extract themes and topics, and identify the main factors and associated factors behind the data.

Synthesis Evaluation: Based on in-depth analysis and grounded analysis, comprehensively evaluate the quality and impact of events or phenomena, and propose improvement suggestions and measures.

Conclusion Presentation: Present and explain the research results and conclusions, demonstrate the contribution and value of the study, and identify the limitations and prospects for future research.

IV.   RESEARCH CONTENT AND INNOVATIONS OF THIS ARTICLE

A. Research Content

The main focus of this article is to analyze online shopping reviews using grounded theory and provide corresponding improvement suggestions based on the analysis results. The suggestions include aspects such as product quality improvement, price adjustment, and service optimization, aiming to meet consumer demands and

enhance consumer satisfaction. In the analysis of online shopping reviews, grounded theory is applied to explore consumers' needs and feedback regarding product quality, price, service, and appearance. This article analyzes consumers' cognition, attitudes, and behaviors, synthesizing the main consumer needs and feedback.

B. Innovations of This Article

- Application of grounded theory in online shopping review analysis: Grounded theory is an effective research approach that facilitates in-depth analysis of consumer needs and psychology. This article applies grounded theory to analyze online shopping reviews, enabling a more accurate exploration of consumers' primary needs and feedback. It provides targeted improvement suggestions for businesses, enhancing their competitiveness and market share.

- Improved efficiency in grounded theory analysis: By utilizing natural language processing (NLP), text clustering, and other machine learning algorithms, this article effectively analyzes online shopping reviews, elevating grounded theory research to an automated level and providing more comprehensive research conclusions.

V. CONSTRUCTING GROUNDED THEORY ANALYSIS MODEL FROM ONLINE SHOPPING REVIEWS

This article focuses on conducting grounded analysis of textual reviews collected from the JD platform, specifically mobile phone reviews, and constructing a grounded theory analysis model. The specific steps for constructing the model are as follows:

- Organize the collected textual data of mobile phone reviews from the JD.com and select relevant and valid reviews related to mobile phone consumption.

- Following the principles of grounded theory, categorize the evaluative content based on its importance and relevance. Utilize methods such as thematic analysis and correlation analysis to determine the relationships between each review content.

- Summarize and condense the review content, extracting the main themes of consumer reviews on mobile phones.

- Apply machine learning algorithms such as NLP and text clustering for data fitting and parameter estimation. This process yields the specific form and structure of the grounded theory analysis model. The model should possess interpretability and predictive capability to ensure the accuracy and reliability of the results. Additionally, the model should be regularly updated and adjusted in response to real-world conditions and market changes to meet evolving consumer and market demands.

- Based on the model results and real-world circumstances, make improvements and optimizations in the design, sales, and services of mobile phone products. This will enhance consumer satisfaction and market competitiveness.

A. Requirements for Data Source

When it comes to data sources, there are specific requirements that need to be considered. These requirements ensure the quality, reliability, and relevance of the data used for analysis and research purposes. The following are key requirements for a data source:

- Data Accuracy: The data should be accurate and free from errors or inconsistencies. It should reflect the true values and characteristics of the phenomenon being studied.

- Data Completeness: The data source should provide a comprehensive and complete representation of the relevant information. It should include all necessary variables and attributes required for the analysis.

- Data Reliability: The data should come from reliable and trustworthy sources. It should be collected and recorded using standardized and validated methods, minimizing potential biases or inaccuracies.

- Data Relevance: The data should be directly relevant to the research objectives and questions. It should align with the specific topic or domain under investigation, ensuring its suitability for the intended analysis.

- Data Timeliness: The data should be up-to-date and reflect the most recent information available. Outdated or obsolete data may not accurately reflect the current state of the phenomenon and can lead to inaccurate conclusions.

- Data Ethical Considerations: The data should be obtained in an ethical manner, adhering to legal and privacy regulations. It is important to ensure that the data collection process respects the rights and confidentiality of individuals involved.

- Data Documentation: The data source should provide sufficient documentation, including metadata, to understand the data structure, variables, and any transformations or preprocessing steps that have been applied.

By considering these requirements, researchers can ensure the quality and integrity of the data used in their studies, enabling robust analysis and reliable research findings.

B.   Collection of mobile phone Review Data

In this article, Python language was used to develop a web scraping program to collect mobile phone review data from the JD.com. The program gathered textual reviews from the "All Reviews" section of the product's review area, ensuring a comprehensive collection. A total of 271,605 reviews were obtained. The collected data was then cleaned, organized, and deduplicated in preparation for subsequent analysis and modeling.

C.   Open Coding

Open coding in the analysis of online shopping reviews allows for a deeper exploration and understanding of the meaning and subjects of the statements. It helps uncover the true needs of consumers and yields insightful and comprehensive research results. Additionally, open coding facilitates consumer participation and identification, thereby increasing the credibility and reliability of the research findings. Compared to quantitative research methods, open coding is more flexible, free, and enlightening, enabling the discovery of new ideas and unusual characteristics of the research subject.

Therefore, open coding provides a detailed and profound understanding and knowledge that meets the research requirements. The comments for mobile phones were organized and preprocessed, resulting in the open coding results shown in Table 1.

The above review text data was segmented and word frequency was calculated.

Table 1-Results of Open Coding

Code	Review Sentence	Initial Concept
1	The agreed price guarantee is in place, but it only took a few days to reduce the price by 400 yuan. Negative review	Price guarantee
2	The new battery is not durable at all	Battery quality
3	The appearance of the body is beautiful, and it feels good	Exterior design
4	It doesn't even have any headphones	Free gifts
5	It can't capture the moon clearly at night with 100 million pixels, and often doesn't have autofocus	Phone pixels
…	……	…

D.   Axial Coding

LSI (Latent Semantic Indexing) algorithm and LDA (Latent Dirichlet Allocation) algorithm, which are NLP techniques, were employed to calculate the topic keywords of the mobile phone review text data for axial coding.

- LSI algorithm and LDA algorithm are both probabilistic models used for topic modeling, but they have differences in their assumptions and methods:
- LSI algorithm assumes that the topics of a document are determined by the co-occurrence patterns of words, while LDA algorithm assumes that the topics of a document are determined by a mixture model of probability distributions;
- LSI algorithm uses singular value decomposition for topic modeling, while LDA algorithm uses variational inference algorithm based on Gibbs sampling;
- LSI algorithm mainly relies on the co-occurrence information of words and represents documents as vectors, while LDA algorithm represents documents as probability vectors of topic distributions and represents topics as probability vectors of word distributions;
- The output of LSI algorithm is a low-dimensional topic space that can be used for tasks like text classification and retrieval, while LDA algorithm provides richer topic information by giving the topic distribution of each document and the word distribution of each topic;
- LSI algorithm has lower computational complexity, but the quality of topic modeling may not be as good as LDA algorithm; LDA algorithm has higher computational complexity but usually provides more accurate topic representations.

Both LSI algorithm and LDA algorithm are important algorithms in the field of topic modeling, with different assumptions, advantages and disadvantages. The common results of axial coding obtained by using these two algorithms are such as: Speed, phone, appearance, screen effects, Honor, appearance, sound effects, feel, cost-performance, standby time, JD, Huawei, looking, logistics, quality, color, pixels, curves, issues, price, a bit, shipping, very beautiful, system, ordering, something, elderly, special, lagging, mom, review, baby, buy phone,

battery, play games, overall, customer service, wife, shopping, authentic, performance, video, friends, brand, features, overall, experience, product,…

Table 2-Results of Axial Coding

Code	Initial Concept	Axial coding
1	Exterior design	Appearance, Looking, Beauty, Color, Curved surface
2	Price guarantee	Price difference, Price protection
	Price level	Price, Cost-performance
3	Logistics speed	Logistics, Shipping
	After-sales service	Attitude, Customer service
4	Brand effect	Honor, Domestic
5	Screen quality	Touch panel, Curved surface
	System quality	Speed, Harmony, Stuck, Function
	Battery quality	Battery, Heating
	Pixel level	Pixel, Photography, Photo, Effect

E. Selective Coding

Selective encoding for online shopping review analysis utilizes the DBSCAN clustering algorithm to cluster 271,605 mobile phone reviews according to the five main categories listed in Table 2. This approach aims to interpret consumer needs and psychology, providing reference for business decision-making. Compared to traditional qualitative analysis methods rooted in theory, the selective encoding in this study offers higher efficiency and reliability, accurately reflecting consumer needs and psychology.

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a density-based clustering algorithm used to discover clusters composed of high-density data points and identify noise points. The DBSCAN algorithm categorizes data points into core points, border points, and noise points, forming clusters by connecting data points with sufficient density.

The calculation formula of the DBSCAN algorithm involves neighborhood query and core point determination, using commonly used formulas for neighborhood query and core point determination.

Neighborhood Query: Given a data point p and a neighborhood radius ε, neighborhoods of p can be computed using the Euclidean distance.

Distance formula:

$$\text{dist}(p,q) = \sqrt{(X_p - X_q)^2 + (Y_p - Y_q)^2} \tag{1}$$

where p and q are data points.

Neighborhood query formula:

$$N_{(p)} = \{q | dist(p,q) \leq \varepsilon\} \tag{2}$$

where $N_{(p)}$ represents the ε-neighborhood of data point p, and q is a data point within the neighborhood.

Core Point Determination:

In the DBSCAN algorithm, a core point is a data point with sufficient density, i.e., its neighborhood contains at least MinPts data points. Core point determination formula:

$$core_{(p)} = |N_{(p)}| \geq MinPts\} \tag{3}$$

where $core_{(p)}$ is a Boolean value indicating whether data point p is a core point, and $|N_{(p)}|$ represents the number of data points within p's ε-neighborhood.

In the DBSCAN algorithm, these calculation formulas are used to determine the neighborhood range, core points, and border points of data points. By performing neighborhood queries and core point determinations, the algorithm can identify clusters composed of high-density data points and separate noise points. These calculation formulas are critical steps in the DBSCAN algorithm and are applied to cluster mobile phone review texts according to the main categories, as shown in Table 3.

Table 3-Selective coding

Code	Main Category	Selective coding
1	Design	Consumption attraction

Code	Main Category	Selective coding
	Service	
	Quality	
2	Price	Consumption intention

From the perspective of both businesses and consumers, this article categorizes the main categories of design, service, and production/manufacturing as consumer attraction and the main categories of price and brand as consumer intention, forming selective encoding (Table 3).

- Consumption attraction, viewed from the perspective of businesses, mainly involves design, service, and production/manufacturing. In addition to ensuring quality, businesses should also explore new products to ensure their uniqueness and fashion ability in design. Furthermore, service provided by e-commerce companies is crucial to consumers, primarily referring to logistics service and after-sales service. Currently, people are more inclined to accept good attitudes and feel respected, as opposed to poor service attitudes and the inability to resolve issues, reflecting the constant evolution of consumer demands. Product quality is the fundamental factor that attracts customers to make purchasing decisions. As consumer demands evolve, customers pay more attention to quality assurance and diversity in products.

- Consumption intention, approached from the perspective of consumers, mainly involves brand and price. Nowadays, consumer demands go beyond basic functional needs and shift towards psychological needs. The higher the brand awareness, reputation, and attention, the more it can stimulate and fulfill customers' desires to consume. Price influences supply and demand, and cost-effectiveness is a key factor that consumers prioritize. Consumers tend to compare prices on e-commerce platforms, and there is a strong correlation between price and online shopping consumption.

VI. CONCLUSION

Through the analysis of consumer comments on mobile phones from the JD platform, selective encoding was conducted in terms of consumer attraction and consumer intention. By observing consumer online shopping behavior from the perspectives of both businesses and consumers, the main axes of encoding cover various rich themes. The main axes of consumer attraction encoding include themes such as color, texture, appearance, and craftsmanship. The service category includes logistics service and customer service attitude. The theme of phone quality encompasses aspects such as processor, memory, processing speed, battery life, pixel, and screen performance, which are basic requirements of consumers for products. If there are defects in product quality, it will directly reduce consumer purchase rates.

Price, as a key focus of consumer intention, is an important consideration for consumers when purchasing mobile phones. Consumers generally consider the cost-effectiveness of products. Cost-effectiveness has a positive impact on consumer intention, and the higher the cost-effectiveness, the higher the consumer's intention to purchase. The brand category includes themes such as Honor and domestic brands. When purchasing mobile phones, the brand's image, reputation, and popularity play an important role in consumers' purchasing decisions. Consumers tend to prefer brands that can meet their psychological needs.

Based on the above themes, businesses can carry out market positioning, product design, manufacturing, and marketing to meet the personalized needs of different consumers, and even provide personalized customization services.

VII. IMPROVEMENT SUGGESTIONS FOR RELEVANT BUSINESSES

- Strengthen product quality management and control. Companies need to enhance product quality by establishing sound quality management systems, improving supplier selection criteria, strengthening production process control, and establishing quality monitoring systems.

- Enhance after-sales service. Companies should establish comprehensive after-sales service systems, implement rapid response mechanisms, provide clear and explicit after-sales policies, train professional after-sales service teams, and offer extended warranty periods or value-added services. These measures can improve the level of after-sales service, promptly address consumer issues, and increase consumer trust and loyalty.

- Enhance design capabilities. Companies can improve their design capabilities by establishing dedicated design teams, enhancing the overall competence of design teams, focusing on principles of ergonomics, emphasizing innovation and uniqueness, and paying attention to details and aesthetics.

- Strengthen brand building. Branding is a core competitive advantage for long-term business development. Therefore, companies should strengthen brand building, increase brand awareness and reputation to attract more consumers and enhance consumer loyalty. This can be achieved through improving product and service quality and implementing effective brand promotion strategies.

- Enhance cost-effectiveness. Companies can continuously optimize product performance, functionality, and user experience through cost control, technological innovation, and managing brand premiums. By doing so, they can improve the cost-effectiveness of their products, attracting consumers to make purchases.

- Pay attention to consumer feedback and conduct analysis. Consumer feedback and evaluations are valuable resources for companies. Utilizing technological means to analyze consumer feedback enables companies to understand consumer needs and preferences, take targeted measures for improvement, and effectively enhance consumer satisfaction, thereby gaining a competitive advantage.

References

[1] Jay Kumar, Junming Shao, Rajesh Kumar, Salah Ud Din, Cobbinah B. Mawuli, Qinli Yang, "A context-enhanced Dirichlet model for online clustering in short text streams," Expert Syst. Appl. vol. 228, Oct. 2023, doi:10.1016/j.eswa.2023.120262.

[2] Laith Abualigah, Khaled H. Almotairi, Mohammed A.A. Al-qaness, Ahmed A. Ewees, Dalia Yousri, Mohamed Abd Elaziz, Mohammad H. Nadimi-Shahraki, "Efficient text document clustering approach using multi-search Arithmetic Optimization Algorithm," Knowledge-Based Systems, vol. 248, 2022, doi:10.1016/j.knosys.2022.108833.

[3] Jingzhu Wei, Huizi Cao, and Lele Zhang, "Research on Users' Needs of Public Culture Cloud Activities Based on Online Comments," Library and Information Service, vol. 66, no. 9, pp. 66-73, 2022.

[4] Wei-Wei Qiu, Hai-Tao Yu, Chin-Hsun (Ken) Tsai, Di (Judy) Zhu, Ming-Hsiang Chen, Hyun Jeong Kim, "Understanding the value of host-guest intimacy behind online reviews of Airbnb," International Journal of Hospitality Management, vol. 115, 2023, doi:10.1016/j.ijhm.2023.103599.

[5] Park, J., Li, X., Li, Q., Kim, J. "Impact on recommendation performance of online review helpfulness and consistency." Data technologies and applications, vol. 57, pp. 199-221, 2023, doi:10.1108/dta-04-2022-0172.

[6] Miriam Alzate, Marta Arce-Urriza, Javier Cebollada, "Is review visibility fostering helpful votes? The role of review rank and review characteristics in the adoption of information," Computers in Human Behavior, vol. 153, 2024, doi:10.1016/j.chb.2023.108088.

[7] Pradeep Racherla and Wesley Friske, "Perceived 'usefulness' of online consumer reviews: An exploratory investigation across three services categories." Electronic Commerce Research and Applications, vol. 11, Issue 6, pp. 548-559, 2012, doi:10.1016/j.elerap.2012.06.003.

[8] David Robert Stöckli and Hamid Khobzi, "Recommendation systems and convergence of online reviews: The type of product network matters!" Decision Support Systems, vol. 142, 2021, doi:10.1016/j.dss.2020.113475.

[9] Hengyun Li, Hongbo Liu, Hyejo Hailey Shin, and Haipeng Ji, "Impacts of user-generated images in online reviews on customer engagement: A panel data analysis." Tourism Management, vol. 101, 2024, doi:10.1016/j.tourman.2023.104855.

[10] Mingming Cheng, Xin Jin, "What do Airbnb users care about? An analysis of online review comments." International Journal of Hospitality Management, vol. 76(A), pp. 58-70, 2019, doi:10.1016/j.ijhm.2018.04.004.

[11] Wenyan Song, Wan Rong, and Yuqi Tang, "Quantifying risk of service failure in customer complaints: A textual analysis-based approach." Advanced Engineering Informatics, Vol. 60, 2024,, doi:10.1016/j.aei.2024.102377.

[12] Kostyra, D., Reiner, J., Natter, M., & Klapper, D. "Decomposing the effects of online customer reviews on brand, price, and product attributes." International Journal of Research in Marketing, vol. 33(1), pp. 11–26, 2016.

[13] Babak Amiri, Ramin Karimianghadim, "A novel text clustering model based on topic modelling and social network analysis," Chaos, Solitons & Fractals, vol. 181, 2024, doi:10.1016/j.chaos.2024.114633.

[14] Laith Abualigah, Khaled H. Almotairi, Mohammed A.A. Al-qaness, Ahmed A. Ewees, Dalia Yousri, Mohamed Abd Elaziz, Mohammad H. Nadimi-Shahraki, "Efficient text document clustering approach using multi-search Arithmetic Optimization Algorithm," Knowledge-Based Systems, vol. 248, 2022, doi:10.1016/j.knosys.2022.108833.

[15] Lin Wang, Huaxia Gao, and Yang Zhao, "Research on online shopping contextual cues: refining classification from text mining," Asia Pacific Journal of Marketing and Logistics, vol. ahead-of-print, 2022, doi:10.1108/APJML-10-2022-0847.

[16] Bin Zhu, Dingfei Guo, and LongRen, "Consumer Preference Analysis based on Text Comments and Ratings: A Multi-attribute Decision-making Perspective." Information and Management, vol. 59, issue 3, April 2022.

[17] Alexis Papathanassis, Friederike Knolle, "Exploring the adoption and processing of online holiday reviews: A grounded theory approach," Tourism Management, vol. 32, Issue 2, pp. 215-224, 2011, doi:10.1016/j.tourman.2009.12.005.

[18] Xavier Matteucci and Juergen Gnoth, "Elaborating on grounded theory in tourism research," Annals of Tourism Research, vol. 65, pp. 49-59, 2017.

[19] Junhui Cheng, Mei Li, "Is There Any Change in Public Comments as the Event Develops? -- The Generative Mechanism of Weibo Users' Commentary Behavior in Multi-Contextual Public Opinion Based on Grounded Theory", Journal of Intelligence, vol. 40, no. 10, pp. 45-51+65, 2021.

[20] Hou J., Zhu A., "Identification of Usefulness for Online Reviews Based on Grounded Theory and Multilayer Perceptron Neural Network." Applied Sciences, vol. 13, no. 9 pp. 5321, 2023.