



Article

Text Mining for Consumers' Sentiment Tendency and Strategies for Promoting Cross-Border E-Commerce Marketing Using Consumers' Online Review Data

Changting Liu ^{1,2}, Tao Chen ^{1,2,*}, Qiang Pu ¹ and Ying Jin ¹

¹ Business School, Ningbo University, Ningbo 315211, China; leyyou@163.com (C.L.)

² Merchants' Guild Economics and Cultural Intelligent Computing Laboratory, Ningbo University, Ningbo 315211, China

* Correspondence: chentao@nbu.edu.cn

Abstract: With the rapid advancement of information technology and the increasing maturity of online shopping platforms, cross-border shopping has experienced rapid growth. Online consumer reviews, as an essential part of the online shopping process, have become a vital way for merchants to obtain user feedback and gain insights into market demands. The research employs Python tools (Jupyter Notebook 7.0.8) to analyze the 14,078 pieces of review text data from the top four best-selling products in a certain product category on a certain cross-border e-commerce platform. By applying social network analysis, constructing LDA (Latent Dirichlet Allocation) topic models, and establishing LSTM (Long Short-Term Memory) sentiment classification models, the topics and sentiment distribution of the review set are obtained, and the evolution trends of topics and sentiments are analyzed according to different periods. The research finds that in the overall review set, consumers' focus is concentrated on five aspects: functional features, quality and cost-effectiveness, usage effectiveness, post-purchase support, and design and assembly. In terms of changes in review sentiments, the negative proportion of the topics of functional features and usage effects is still relatively high. Given the above, this study integrates the 4P and 4C theories to propose strategies for enhancing the marketing capabilities of cross-border e-commerce in the context of digital cross-border operations, providing theoretical and practical marketing insights for cross-border e-commerce enterprises.

Keywords: online reviews; LDA; LSTM; cross-border e-commerce; marketing strategy



Academic Editors: Chienliang Lin,
Wuke Zhang, Sung-Wen Yu and
Xu Hanqing

Received: 25 April 2025

Revised: 26 May 2025

Accepted: 27 May 2025

Published: 2 June 2025

Citation: Liu, C.; Chen, T.; Pu, Q.; Jin, Y. Text Mining for Consumers' Sentiment Tendency and Strategies for Promoting Cross-Border E-Commerce Marketing Using Consumers' Online Review Data. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 125. <https://doi.org/10.3390/jtaer20020125>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As one of the three key drivers propelling the national economy, foreign trade holds pivotal significance. Cross-border e-commerce, as an emerging form of foreign trade, has demonstrated steady growth in total trade volume in recent years, accompanied by the rise of global platforms such as Temu, SHEIN, and TikTok Shop. Furthermore, the Guiding Opinions on Promoting Stable Scale and Optimized Structure of Foreign Trade issued by the General Office of the State Council emphasizes the need to foster healthy and sustainable innovation in cross-border e-commerce, supporting foreign trade enterprises in expanding sales channels and cultivating independent brands through new business models like cross-border e-commerce. This sector has become a critical force in stabilizing foreign trade.

With the rapid increase in internet accessibility and continuous advancements in information technology, online shopping has become indispensable for fulfilling consumer

demands. According to the China E-Commerce Report released by the Ministry of Commerce, China's online retail sales reached CNY 13.79 trillion in 2022, marking a 4% year-on-year increase. E-commerce, acting as a "stabilizer" for China's economic recovery, further highlights its role in driving consumption upgrades, promoting employment growth, and enhancing industrial competitiveness. At the same time, enterprises also favor overseas markets, and cross-border e-commerce has become a popular way for enterprises to explore new platforms and obtain new traffic. Data from the General Administration of Customs indicates that China's cross-border e-commerce import and export volume exceeded CNY 2.11 trillion in 2022, reflecting a 9.8% annual growth. The booming e-commerce market has generated massive volumes of customer reviews. As an integral component of online shopping, post-purchase reviews play a crucial role, particularly amid challenges such as sluggish external demand, diminishing platform dividends, and rising traffic acquisition costs. A deep understanding of overseas consumers' concerns is significant for businesses formulating internationalization strategies and building global brands.

Based on the post-purchase review data of the best-selling products of cross-border e-commerce Company L, this paper constructs LDA and LSTM sentiment classification models to analyze review topics and sentiment tendencies. It also divides the data into different time periods to complete the analysis of the evolution trend of topic sentiment. This study enriches theoretical research on converting fragmented reviews into structured marketing insights by integrating 4P and 4C theories in digital cross-border operations. It also offers actionable marketing insights for cross-border e-commerce enterprises.

2. Related Research

2.1. Information Asymmetry Theory, Signaling Theory, and Online Review Research

Information asymmetry is prevalent in the process of commodity transactions. Since sellers have a more comprehensive understanding of product information, they often focus on promoting the advantages of products while avoiding information such as product defects during the sales process. Buyers may make decision-making behaviors that are unfavorable to themselves due to this lack of information. In the field of cross-border e-commerce, this information gap is further intensified, as the product information provided by online merchants is often the sole guide for consumers [1]. Additionally, consumers in different countries and regions differ in culture, language, consumption habits, and other aspects, which makes it more difficult for them to obtain and understand product information [2]. Such differences in information transmission can significantly affect consumers' perceptions and, in turn, their satisfaction with online merchants [3].

Based on information asymmetry, signaling theory focuses on how the party with the information advantage transmits information to the party with the information disadvantage by sending signals, thereby influencing the latter's decision-making behavior. In the current prosperous era of e-commerce shopping, consumers often leave post-purchase reviews after shopping through e-commerce platforms to express their experiences of the consumption process and product use. These reviews are rich and diverse, covering multi-dimensional information like product performance, logistics, service quality, and consumers' emotional feedback, with both positive and negative evaluations. Compared with merchants' one-way promotion, these reviews constitute high-quality signals with greater authenticity and reference value, which can intuitively convey the operational status of online merchants and the actual standards of products or services to potential consumers. These reviews directly reflect consumers' core concerns and serve as key references for potential buyers' purchasing decisions. Positive reviews often signal excellent product quality and performance, effectively boosting consumers' willingness to buy. Conversely, negative reviews can raise concerns and dampen purchasing intent [4]. Meanwhile,

these reviews can also provide clear guidance for cross-border e-commerce to optimize marketing strategies, helping cross-border e-commerce enterprises accurately grasp the marketing direction.

Therefore, the establishment of an online review mechanism can effectively bridge the information gap and provide consumers with a channel to obtain information for making purchasing decisions. Currently, research on online reviews mainly focuses on the following three aspects:

Firstly, research on the usefulness of online reviews. Previous studies have identified various signals influencing review helpfulness, including factors related to the review itself and the reviewer [5,6]. Some studies have constructed multiple linear regression models and found that the number of attached images significantly impacts the perceived usefulness of reviews. Specifically, the more images included in negative reviews, the more effectively they help other consumers visually understand the problems with products or services, thereby enhancing the usefulness of the reviews [7]. Yan, using review data from social media platforms and e-commerce platforms as samples, found that reviews with emojis have obvious advantages over pure text reviews in conveying emotions and enhancing review attractiveness, and are more likely to be recognized as useful information by other users. However, when the number of emojis is excessive, it may make the reviews appear too casual or affect the effective transmission of information, thereby reducing the usefulness of the reviews [8]. Based on the interdisciplinary theory of psychology and marketing, scholars like Zhu [9] explored the interactive effects, influence mechanisms, and boundary conditions between online reviews' linguistic style and regulatory focus on review usefulness. This provides a theoretical basis for merchants to optimize review management and consumers to obtain useful review information more effectively. As the beneficial impact of online consumer reviews on corporate sales has become widespread, some enterprises have resorted to unethical practices, such as deleting negative reviews, creating fake positive reviews, and even posting negative reviews to tarnish competitors' reputations [10]. This further demonstrates the useful characteristics of online reviews for consumers and their significant impact on commerce.

Secondly, research on the influence of online reviews. Against the backdrop of rapid digital communication development, scholars analyzing consumer review data from multiple channels such as social media and e-commerce platforms have found that online consumer reviews are increasingly replacing traditional word-of-mouth communication [11]. This shift demonstrates the critical impact of online reviews on business operations, indicating that marketers need to carefully manage online reviews to shape consumers' positive perceptions of their corporate brands and products. Guo and Li [12] analyzed consumer online review data from e-commerce platforms and found that during public health emergencies, consumers' perceived crisis risks significantly increase. In such cases, consumers rely more on online reviews to evaluate product or service safety and reliability, thus amplifying the impact of online reviews on their online purchase decisions. This provides important references for enterprises to formulate marketing strategies, optimize product services, and respond to changes in consumer decision-making during special periods.

Thirdly, the practical application of online reviews. Some scholars have studied the application of online reviews in the field of personalized recommendations [13]. There are also applications in the field of user satisfaction [14,15]. Some scholars have combined multiple industries: for example, online texts were applied to the tourism field, and used to promote the precise marketing of scenic areas through sentiment analysis and topic clustering [16]. Some scholars used the analysis of online reviews to drive product design, so as to improve user satisfaction and product competitiveness [17].

2.2. Reflections on the Integration of 4P and 4C Marketing Theories

As a foundational framework for marketing mix, the traditional 4P theory (Product, Price, Place, Promotion) [18] has provided referable business concepts for cross-border e-commerce marketing activities from an enterprise perspective, especially amid the booming internet economy and the emergence of various online marketing approaches. For example, Thailand's Hill Tribe cocoa enterprise has enhanced product attractiveness through precise brand positioning and promotional strategies, and achieved market expansion by optimizing cross-border logistics channels [19]. Jingdong Mall has built competitive barriers and strengthened market share via standardized product line layout and self-built logistics systems [20]. However, in the context of digital cross-border operations, the limitations of the 4P theory have gradually emerged. Excessive focus on products and prices has caused homogeneous competition, and its supply-side perspective fails to adapt to the fragmented trend of consumer demands. Studies show that enterprises relying solely on 4P strategies lag in user stickiness and market response speed. Particularly in the data-driven era of precision marketing, the lack of in-depth analysis of consumer behavior has become their major shortcoming [21].

In contrast, the 4C theory (Consumer, Cost, Convenience, Communication) is user-centered and emphasizes optimizing the entire shopping process from demand insight to purchase [22]. For example, the Mia platform collects user feedback through its "grass planting machine" feature to optimize product selection and reduce decision-making costs. It also enhances convenience through easy APP ordering and logistics tracking, while using social interactions to strengthen user engagement [23]. However, implementing the 4C theory relies on big data technology and substantial effective customer feedback. Although some cross-border e-commerce enterprises recognize the importance of communication (such as social media marketing), they struggle to achieve precise targeting through effective data mining, leading to wasted promotional resources [24].

Post-purchase reviews from cross-border e-commerce consumers hold significant data mining value, as enterprises can obtain valuable insights from such customer feedback to improve products and optimize marketing strategies [25]. Through topic mining, companies can accurately identify consumer concerns about product quality, functionality, and other aspects, thereby extracting product improvement suggestions (Product) and matching personalized promotion methods (Promotion) to shape and convey brand images, increasing the likelihood of cross-border products going viral in online contexts. Given the global nature of online shopping, factors like tariffs and transportation make consumers more carefully consider the cost risks of online purchases. Analyzing cost-related feedback in reviews can help optimize pricing structures to reduce consumers' overall costs (Cost) and enhance purchase intent. Additionally, convenience factors reflected in reviews—such as information access, logistics, and usage methods (Convenience)—play a critical role in attracting potential consumers and increasing user retention, also driving enterprises to build multi-channel entry points and other improvements [26].

However, existing research underutilizes the critical data point of "consumer online reviews" and mostly isolates studies on 4P and 4C. It fails to systematically analyze how to integrate 4P and 4C strategies via review data or convert fragmented reviews into structured marketing insights. Therefore, this paper integrates the 4P and 4C theories, synthesizes the perspectives of enterprises and consumers, and proposes cross-border e-commerce marketing improvement strategies based on consumer reviews from the following four aspects: defining products through user needs (Product $\times$ Consumer), comprehensive cost optimization (Price $\times$ Cost), convenient purchase process experience (Place $\times$ Convenience), and two-way value transmission (Promotion $\times$ Communication).

2.3. Research on Cross-Border E-Commerce Strategies

With the rapid development of the cross-border e-commerce industry, several distinct characteristics have emerged: platform diversification, refined operations, matured regulation, and fierce competition. Against this backdrop, scholars have started actively exploring new strategic trends in cross-border e-commerce.

According to existing research, from a macro perspective, scholars have studied cross-border e-commerce from different perspectives and the key links or current situations in the development of cross-border e-commerce, providing theoretical guidance and policy recommendations. Ding used a game theory model to analyze how consumers' preferences for product price, quality, and features influence merchants' development of differentiated competition strategies [27]. This provides a theoretical basis for cross-border e-commerce merchants to achieve precise positioning and formulate marketing strategies in competition. Based on a cross-cultural perspective, Cai pointed out that the application of AI technology in cross-border e-commerce marketing activities can significantly enhance user engagement and purchase conversion, but its cross-cultural communication effects are significantly influenced by cultural differences. Therefore, they proposed that cross-border e-commerce marketing activities need to build a dual engine of AI technology and cultural insight [28], which provides a reference practice path for cross-border e-commerce to use AI to break through cultural barriers in marketing. Xiong [29] explored the differences in path construction between platform-based models from the perspective of independent website establishment, providing reference suggestions for cross-border e-commerce enterprises to select appropriate marketing models according to their own development strategies and resource conditions. There are also scholars who, in combination with the strategic background of the "the Belt and Road Initiative", have discussed issues such as the development of cross-border e-commerce between China and ASEAN [30], the optimization of China's cross-border export strategies [31], and the development of cross-border agricultural product trade [32]. These studies reveal the adaptation mechanisms between marketing strategies and market environments, cultural differences, and channel characteristics from multiple perspectives, providing references for cross-border e-commerce enterprises to enhance the effectiveness and adaptability of marketing strategies in dynamic competition.

Secondly, at the micro level, many research results are based on the marketing level for analysis. For example, some scholars notes that small and medium-sized enterprises (SMEs) enhance competitiveness through patented technologies or unique designs to adapt to the "blockbuster product" marketing logic in cross-border e-commerce, providing multi-level references for cross-border e-commerce practices based on enterprise scale differences [33]. Some scholars propose that live-streaming e-commerce can reduce consumers' perceived quality risks of cross-border products through scenario-based marketing and real-time interaction, while leveraging public domain traffic from social media and private community operations to increase product exposure [34]. Fan et al. use online review data as the core data source for consumer needs, mine keywords from consumer reviews, classify evaluation indicators into the five categories of the Kano model using algorithms, construct an evaluation index system for cross-border e-commerce brand internationalization, and propose differentiated marketing strategies, providing quantifiable marketing decision-making tools for cross-border e-commerce brands [35]. There are also researchers who have analyzed marketing strategies of typical cross-border e-commerce platforms and enterprise cases. For example, Jiang H et al. studied TikTok marketing, analyzed its influence mechanism on users' purchase behavior, and found problems such as the insignificant influence of some marketing methods and consumers' impulsive consumption tendency to overlook product features. They proposed that short video platforms use big data for

content pushing, merchants focus on videos' entertainment value and product promotion, and platforms enhance user attitudes through promotional activities [36].

In summary, the academic community has carried out many beneficial explorations and studies in the fields of cross-border e-commerce and online reviews, achieving certain results. However, as a highly complex business form, cross-border e-commerce involves multiple complex challenges in its operational chain, such as policies and regulations of multiple countries, cultural differences, logistics systems, and payment and settlement. These factors are bound to influence consumers' behaviors from pre-purchase decision-making to post-receipt and post-use review behaviors. Existing research still has significant room for expansion in different application scenarios and consumer group segmentation.

Hofstede's cultural dimension theory [2] explaining individual behavioral differences highlights the significant impact of cultural disparities in transnational services, especially in today's context of economic globalization. Han's study [37], based on Hofstede's cultural dimension framework, shows that cultural differences among online consumers affect their perceptions of service quality and online review sentiment. Therefore, the impact of cultural differences on cross-border e-commerce review sentiment warrants further investigation. In addition, unlike domestic e-commerce, cross-border logistics involve complex processes—interconnected stages from goods collection in exporting/importing countries to customs clearance—regulated by governments [38]. Current research has minimally addressed issues like delayed logistics [39] or product quality failures during transportation and their link to customer negative reviews. Exploration of the relationship between logistics service quality and consumer negative evaluations remains limited in cross-border e-commerce contexts [40]. Additionally, international transportation—particularly tariffs and import taxes—may increase total purchase costs, affecting consumers' evaluation of cross-border products and, consequently, their purchase decisions and post-purchase review behaviors [41].

Overall, driven by advancements in technology, the growth of cross-border e-commerce has impacted the entire retail sector by offering customers incredible levels of global market access. Meanwhile, the increasing number of consumer post-purchase online reviews also reveals various issues in the development of cross-border e-commerce, which largely influence potential consumers' purchase decision-making. Most researchers have focused on the impact of online review statistics, such as the number of reviews [42], ratings [43], and text length [44]. However, with e-commerce's rapid development, consumers no longer rely only on numerical ratings but prefer to understand and trust online reviews' textual content [45]. Moreover, negative reviews more directly reflect the pain points where cross-border e-commerce fails to meet consumer needs and require resolution.

Aiming at the gap in existing research—lacking integration of the 4P and 4C theories in the digital cross-border context to transform fragmented reviews into structured marketing insights and propose targeted marketing improvement strategies for cross-border e-commerce—this study analyzes post-purchase review data from cross-border e-commerce enterprise L. Through text mining to explore consumer focus themes and combining theoretical perspectives of both enterprises and consumers, it puts forward marketing improvement strategies for cross-border e-commerce in the digital cross-border era. This research provides theoretical and practical insights for the better development of cross-border e-commerce enterprises and enriches the research content in the fields of online review analysis and cross-border e-commerce.

3. Research Design

3.1. Research Framework

To address the above research questions, social network analysis (SNA) can construct a keyword co-occurrence network from reviews, identify core nodes in negative reviews, and visually present the relevance of consumer pain points, providing a focused direction for subsequent analysis in this study [46]. In cross-border e-commerce scenarios, consumers have multilingual and multicultural backgrounds, and their reviews are diverse and scattered, and contain multiple emotional polarities. In this context, the LDA model can effectively transform massive fragmented reviews into structured and interpretable thematic frameworks [37], while the LSTM model can leverage its sequential learning advantages to achieve precise emotional polarity classification in multilingual environments. Notably, the thematic distributions output by the LDA model can map to the 4P–4C theory of this study, providing a solid theoretical basis for strategic recommendations.

The comprehensive application of the SNA–LDA–LSTM methodological combination not only follows the technical logic of text mining from structuring to emotional analysis but also aligns with the complex needs of cross-border e-commerce's multicultural, multi-stage, and dynamic market environment. By integrating methods with theory, this approach offers a rigorous deductive pathway to construct marketing improvement strategies for cross-border e-commerce in the digital cross-border era. Specific applications are as follows:

Firstly, the review texts from cross-border e-commerce sales platforms are crawled as analysis corpora. Through text cleaning and preprocessing, a dataset for subsequent analysis is formed. Next, social network analysis tools are used to screen out the negative reviews in the review collections as the analysis basis for mining marketing pain points. Then, an LDA topic model is constructed to find the optimal number of topics, classify and summarize the topics, and analyze the topics that consumers are concerned about. Subsequently, the review texts are transformed into structured data, and the word vector representations obtained through the transformation are used for emotional classification. Then, deep learning algorithms such as LSTM are adopted, and the word vectors are used as input data for training. The parameters are adjusted to obtain the optimal effect, and the evaluation of the emotional tendencies of consumers' reviews is completed. Finally, marketing strategy suggestions for cross-border e-commerce are put forward based on the model results. The research framework of this study is shown in Figure 1.

3.2. Model Introduction

3.2.1. LDA Model

The LDA (Latent Dirichlet Allocation) topic model, also known as the latent Dirichlet distribution model, was proposed by Blei and colleagues in 2003. It is a commonly used topic generation model in the fields of machine learning and text mining. It adopts an unsupervised learning approach and can automatically learn the topic distribution without providing labels, which greatly reduces the labor and time costs. The model has a three-level structure: document layer, topic layer, and feature word layer. Given a document library and the number of topics, it can extract the distribution information of topic words and present it as probabilities. Subsequently, text clustering can be completed to form a set of categories with substantial significance, helping readers quickly understand the document information. The model structure is illustrated in Figure 2.

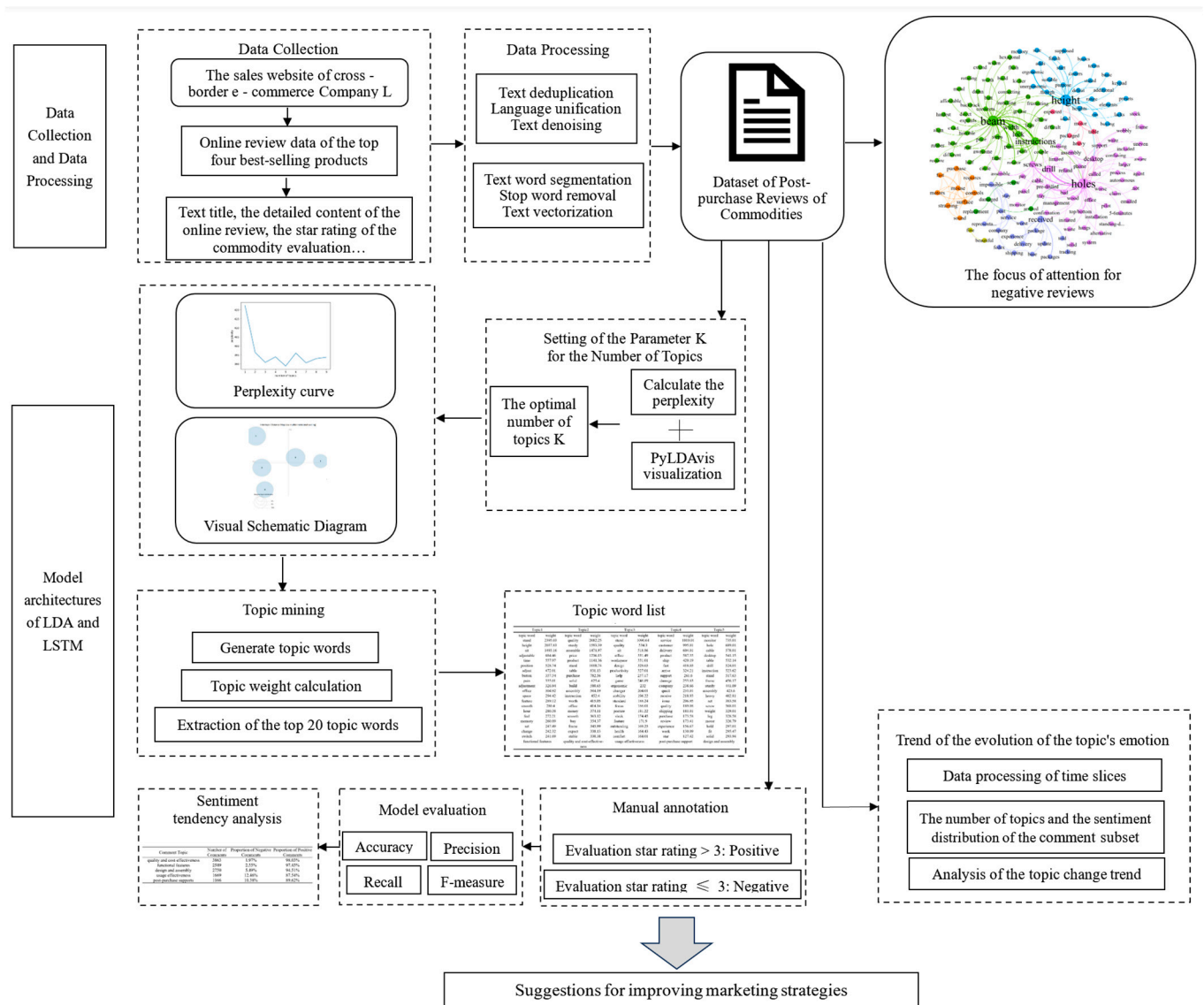


Figure 1. Schematic diagram of the research framework.

In the figure, N denotes the number of words in a document, M represents the total number of documents, and K stands for the number of topics. Both α and β are key parameters of the LDA model, and they both follow the Dirichlet distribution. The training and optimization process of the entire LDA model can also be regarded as the process of optimizing the parameters α and β . The specific generation process is as follows: First, α is used for random initialization to generate the multinomial distribution $\theta(\theta \sim Dirichlt(\alpha))$ of topics corresponding to documents, and then the document topics Z are randomly generated based on the topic multinomial distribution θ . Second, β is used to randomly generate the multinomial distribution Φ of words corresponding to topics. Then, by combining the document topics Z and the word distribution $\phi(\phi \sim Dirichlt(\beta))$, the words W are generated. Finally, by continuously repeating the above two steps, M documents with K topics are generated.

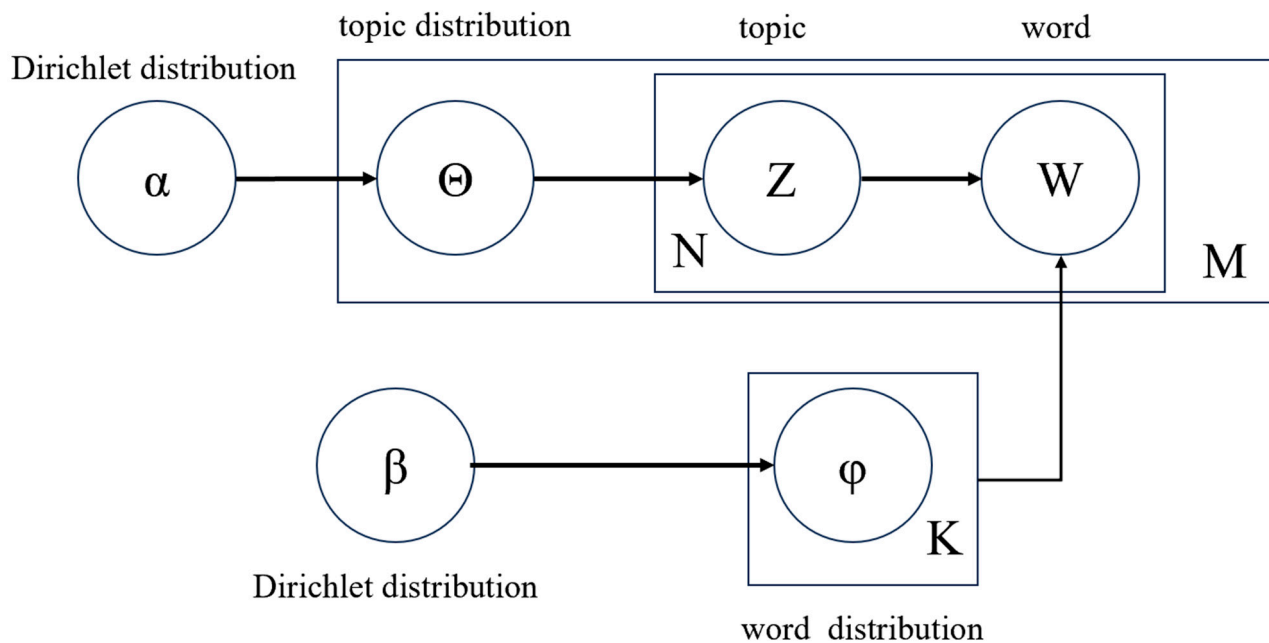


Figure 2. The principle of the LDA model.

3.2.2. LSTM Model

The judgment of the emotional tendency of text language is an important research direction in the field of natural language processing. In the digital economy era, this technology is widely applied in social media, news, commodity shopping, and other domains. It enables us to quickly understand the macroscopic emotional feedback of people towards a certain thematic event or commodity. Deep learning models usually do not require manual feature extraction in text analysis tasks. Instead, they can autonomously learn and retain the main feature information in the original text, and combine the context semantics to achieve more accurate emotional classification results.

Currently, the popular neural network models mainly fall into two categories: CNN (Convolutional Neural Network) and RNN (Recurrent Neural Network). The LSTM (Long Short-Term Memory) neural network model, in full, is an advanced structure developed on the basis of the traditional RNN. Due to the standard RNN model's inability to retain long-term information, LSTM incorporates structural modifications which effectively address the issues of gradient disappearance or explosion that arise when traditional RNN models process lengthy texts. Considering that review texts can be regarded as a kind of long-sequence data, and as a typical deep learning model, LSTM has the advantage of handling data with time-series characteristics when processing texts and can deal with semantic information more comprehensively. Therefore, this study selects it as the main classifier for training. The model structure is shown in Figure 3.

In the structure, the LSTM introduces a series of gating units to enrich its hidden layer: the Input Gate, Forget Gate, and Output Gate. All three are transformed into 0–1 values via the Sigmoid activation function, acting as weight parameters to control information flow. Among them, the Input Gate is responsible for controlling whether the input information at the current moment enters the memory unit; the Forget Gate is responsible for controlling which information in the memory unit is forgotten or retained; and the Output Gate controls how the information in the memory unit affects the current output. At the same time, the memory unit (Cell), as a key part of the model, runs through the entire computational pathway. It updates its state through the outputs of the Input Gate and the Forget Gate, enabling the effective information of each node to be retained.

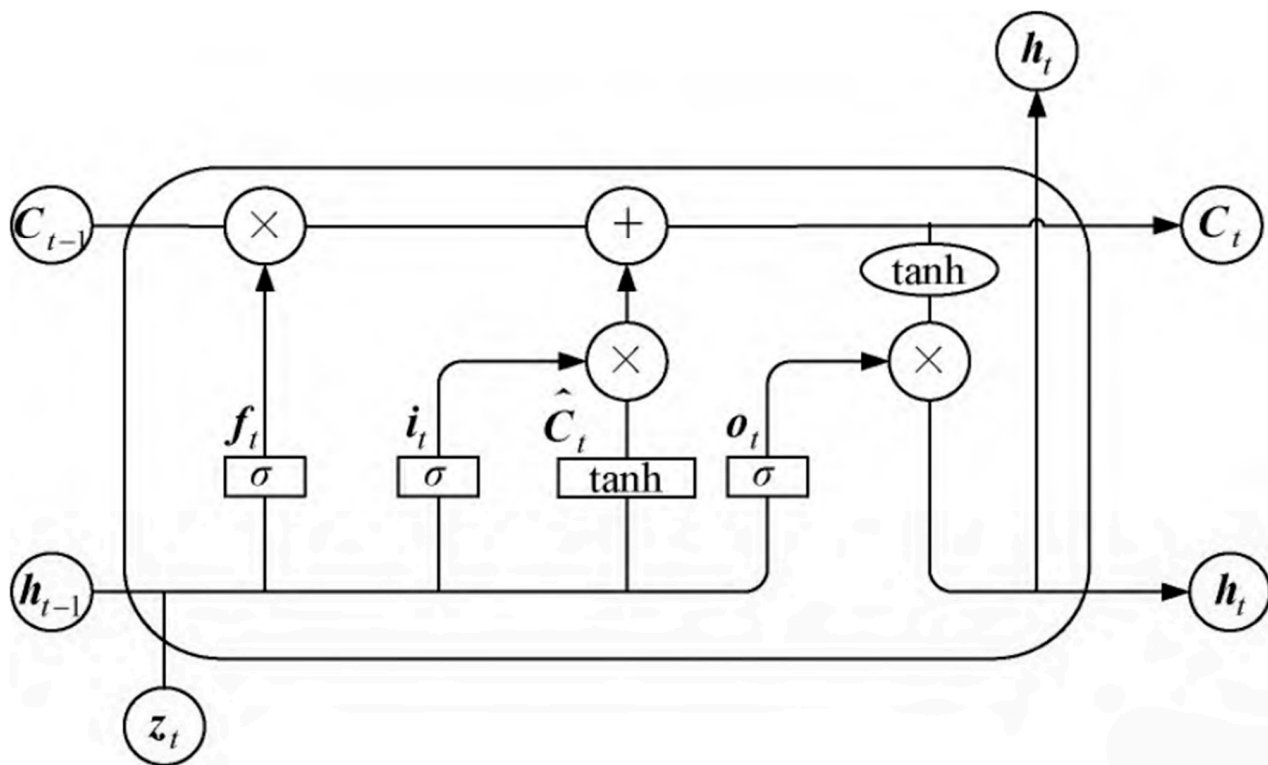


Figure 3. The structure of the LSTM model.

3.3. Data Source

This study uses review data from Company L's independent website as the data source, selecting review data of the top four best-selling products in its standing desk category as the dataset. As a listed company with over a decade of experience in cross-border e-commerce and a leading enterprise in ergonomic products, standing desks are its core products. Not only do they sell well on its independent website, but they also frequently top sales rankings on major cross-border e-commerce platforms such as Amazon, Walmart, and Yahoo. Therefore, its best-selling products have a large user base, and their review data can more comprehensively cover the usage experiences and feedback of different types of consumers, ensuring the richness and representativeness of research data. Additionally, these four products differ in design, price, and other aspects, meeting the needs of different consumer groups, further enriching data diversity. This helps to more comprehensively mine consumers' emotional tendencies and product pain points, making the research scheme on its product marketing strategies have a certain universality and providing useful references for other small and medium-sized cross-border e-commerce sellers to improve products and enhance marketing strategies.

A total of 14,078 pieces of review text data from February 2019 to December 2024 were collected through the Octoparse crawler. The original evaluation fields included text title, reviewer, detailed evaluation content, review date, detailed review content, and evaluation star rating information (1–5 stars, where 1 star indicates the least satisfaction and 5 stars indicates the highest satisfaction). Some of the collected information is shown in Table 1. Subsequent analysis was mainly based on the text format fields and the rating fields.

Table 1. Some collected information.

Title	Date	Content	Rating
I made the right choice!!	18 November 2023	Unbelievable quality, and assembly was easier and quicker than expected. The desk far exceeds my expectations and is amazingly stable. Highly recommended!!!	5
Support is amazing	16 November 2023	Experience was so good that I bought a different product. Best part to me was support.	5
Bad Installation Process	2 January 2024	Table top and legs work and look good. Installation could be much better. Table top should have threaded inserts instead of tiny pilot holes. The pilot holes didn't even line up correctly with the base.	3

4. Empirical Analysis

4.1. Data Processing

To avoid the impact of issues such as duplicates, meaningless comments, mismatches between data content and fields, and garbled characters in the original dataset on the accuracy and stability of subsequent models, this study performed data cleaning steps including text deduplication, language unification, and text denoising on 14,078 original review text data to ensure data quality. A total of 11,997 cleaned English review datasets were obtained.

Before proceeding with the next step of analysis, it was also necessary to convert them into structured data that can be recognized and processed by a computer. Therefore, the cleaned data were further processed, including text word segmentation, stop word removal, text vectorization using the word2vec model, etc. For text word segmentation, as the text type in this study was English with spaces as obvious delimiters, the Python third-party library NLTK was used as the word segmentation tool to obtain word sequence results by splitting text at spaces. For stop word removal, this study adopted a commonly used English stop word list publicly available on the CSDN website as the initial stop word list. On this basis, combined with the characteristics of the review data, high-frequency but semantically irrelevant phrases such as “desk” and “online” were added to construct a stop word library suitable for mining cross-border sales reviews of standing desks. Additionally, to ensure research generality, standing desk brand names were also included in the stop word list.

4.2. Social Network Analysis of Negative Review Concerns

4.2.1. Construction of Co-Occurrence Network for Negative Review Concerns

Due to the scattered nature of consumer review perspectives, this study aimed to gain an overall understanding of word co-occurrence and semantic relationships in negative reviews and explore key concerns in consumer negative feedback. Negative evaluations were defined as reviews with a rating of 3 stars or below. A semantic network analysis method was introduced to construct a word co-occurrence matrix. The co-occurrence window was set as the entire review, meaning that all the words appearing in the same review are considered co-occurring, preserving associations between all the words within the review. Word pairs with a co-occurrence frequency ≥ 2 were filtered, and the top 300 high-frequency word pairs were selected. The results were visualized using Gephi software (version:0.10.1) to obtain the semantic co-occurrence network shown in Figure 4. Each node in the figure represents a keyword—the larger the node, the higher the co-

Table 2. Network centrality indicators and node rankings.

Centrality Indicator	Top 10 Nodes by This Indicator
Degree centrality	beam, holes, height, instructions, drill, screws, width, received, lock, desktop
Closeness centrality	instructions, beam, holes, height, screws, desktop, couple, missing, table, legs
Betweenness centrality	holes, beam, height, instructions, received, desktop, screws, table, couple, service

Based on the ranking results of the three centrality indicators and the semantic co-occurrence network diagram, the co-occurrence results of negative reviews focus on core key nodes such as “beam”, “holes”, and “height”, intuitively reflecting the central topic of user discussions and indicating product manufacturing defects to a certain extent. Nodes like “instructions” and “received” also reveal shortcomings in merchants’ services and other aspects.

The above nodes reveal the current deficiencies of the product. The presentation of the problems is relatively detailed. Overall, the information reflected by the nodes is relatively single. Therefore, it is necessary to further sort out the reviews at a macro level and establish a global view of the product. In particular, by focusing on analyzing consumers’ key concerns about the product and their degrees of praise and criticism, enterprises can clearly identify and consolidate the existing competitive advantages and highlights of the product, maintain market competitiveness, and accurately locate the negative issues that affect the user experience, so as to implement targeted improvement strategies.

4.3. Topic Mining Based on LDA

In order to further grasp consumers’ feedback on products and understand the marketing status quo and deficiencies of cross-border e-commerce products, this study uses the LDA model to conduct in-depth mining of review data, providing guidance for the refinement direction of the cross-border e-commerce marketing route.

4.3.1. Optimization of the Number of Topics

This study uses the sklearn library in Python to perform LDA modeling. Since the LDA model cannot set expected results based on general cognition, the perplexity calculation method is adopted to determine the number of topics [47]. A smaller perplexity value indicates better clustering of the LDA model and a stronger ability to predict or represent information in the dataset. As shown in Figure 5, when the number of topics is five, the perplexity of the LDA topic model reaches a local minimum. As the number of topics increases, the model’s perplexity rapidly decreases, then slightly rises, and hits the lowest value at five topics, generally showing an L-shaped trend. This suggests that when the number of text topics is five, the model achieves the best classification effect and accuracy.

To further verify the rationality of topic number division, the model data with five topics is visualized using the pyLDAvis library, as shown in Figure 6. In the figure, the categories numbered with digits represent five different topics. The size of the category reflects the importance of the topic, that is, the larger the category, the higher the universality of the topic in the review document set. At the same time, the relative distance between categories reveals the correlation between topics. The farther the distance, the weaker the correlation. It can be observed that the five categories do not overlap in the entire two-dimensional vector space, indicating clear distinctions between topics and a good model training effect. Therefore, the optimal number of topics K for the LDA model is determined as five for subsequent analysis. Additionally, the prior parameters are set as $\alpha = 0.1$, $\beta = 0.01$, and the maximum number of iterations is 50.

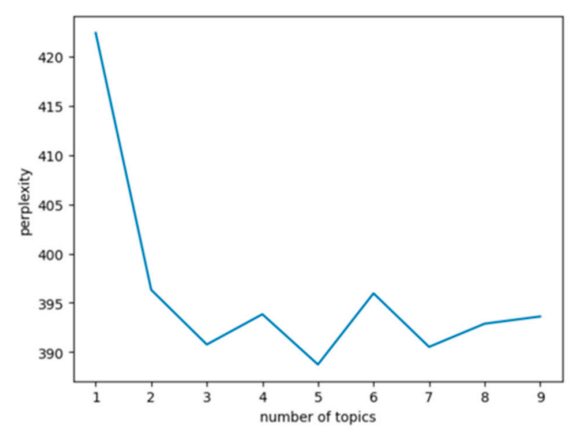


Figure 5. Perplexity curve.

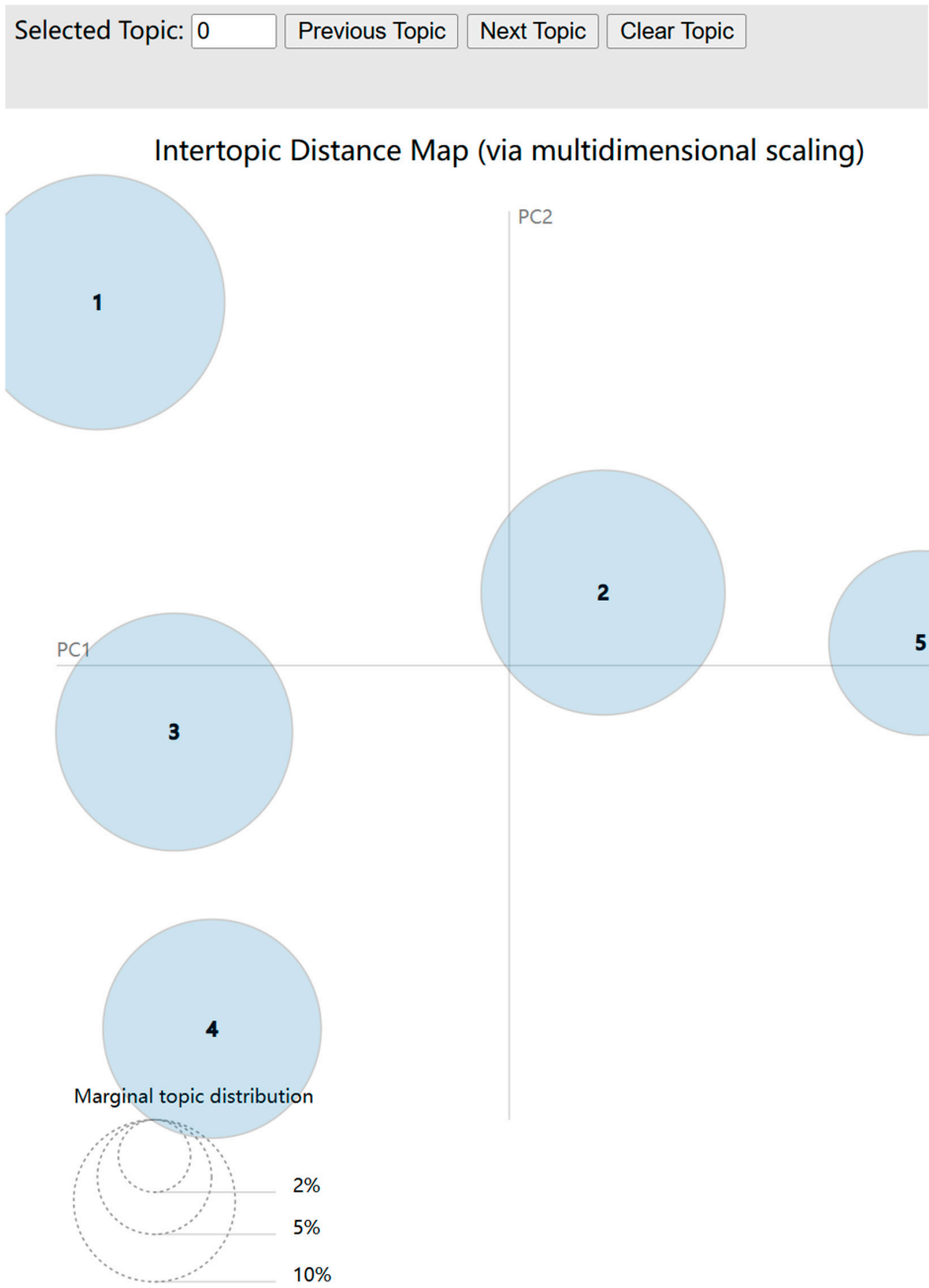


Figure 6. Visualization results of LDA.

4.3.2. Analysis of Topic Mining Results

Based on the above model calculations, the top 20 feature words under each of the five topics are summarized by weight ranking, as shown in Table 3.

Table 3. Results of some topic words.

Topic 1		Topic 2		Topic 3		Topic 4		Topic 5	
topic word	weight	topic word	weight	topic word	weight	topic word	weight	topic word	weight
stand	2395.03	quality	2082.25	stand	1090.64	service	1010.01	monitor	735.01
height	2037.83	sturdy	1593.39	quality	534.3	customer	995.01	hole	689.01
sit	1493.16	assemble	1471.97	sit	518.86	delivery	604.01	cable	578.01
adjustable	684.46	price	1236.13	office	331.49	product	587.35	desktop	541.15
time	557.97	product	1141.36	workspace	331.01	ship	420.19	table	532.14
position	528.74	stand	1018.74	design	329.65	fast	418.03	drill	524.01
adjust	472.01	table	831.13	productivity	327.01	arrive	324.21	instruction	523.62
button	357.54	purchase	782.56	help	257.17	support	261.6	stand	517.63
pain	335.01	solid	625.4	game	248.89	damage	253.83	frame	456.17
adjustment	326.94	build	580.65	ergonomic	232	company	238.66	sturdy	441.09
office	304.92	assembly	504.19	changer	204.01	quick	231.01	assembly	423.6
space	294.42	instruction	452.4	stability	196.22	receive	218.93	heavy	402.81
feature	289.12	worth	415.05	standard	188.24	issue	206.95	set	383.58
smooth	280.4	office	414.14	focus	186.01	quality	189.08	screw	368.01
hour	280.38	money	374.11	posture	181.22	shipping	181.01	weight	329.01
feel	272.21	smooth	363.12	sleek	174.45	purchase	173.58	leg	328.58
memory	260.89	buy	354.37	feature	171.9	review	173.41	motor	326.79
set	247.49	frame	345.99	outstanding	169.25	experience	156.67	hold	297.01
change	242.32	expect	338.13	health	164.43	week	130.09	fit	295.47
switch	241.69	stable	330.38	comfort	164.01	star	127.42	solid	293.94
functional features		quality and cost-effectiveness		usage effectiveness		post-purchase support		design and assembly	

In Topic 1, the high-weight words mainly focus on the core functions of the sit–stand desk (“stand”, “height”, “sit”, “adjustable”, etc.) and the office scenarios (“office”, “time”, “space”, etc.). These words reflect consumers’ concerns about the product’s height adjustment and the improvement of the office experience. Therefore, this topic is summarized as the functional features of the product. Functional features focus on consumers’ evaluations of the product’s inherent functions, technical parameters, additional features, etc., focusing on “what the product has”, such as its lifting function and height memory function.

In Topic 2, the high-weight words mainly concentrate on the product quality (“quality”, “sturdy”, “solid”, etc.) and the purchase value (“price”, “purchase”, “worth”, etc.), highlighting consumers’ emphasis on product quality and pricing. Thus, this topic is related to the product’s quality performance and cost-effectiveness.

In Topic 3, some high-weight words overlap with those in Topic 1, indicating that the relevant words have a relatively high frequency of occurrence in these two topics. However, words such as “productivity”, “help”, “game”, “ergonomic”, etc. show differences from other topics. Especially for office workers, gamers, and other groups, ergonomic products enable consumers to pay attention to the adjustment of their postures, relieve physical discomfort caused by sitting for a long time, and improve work efficiency and the quality of entertainment. Therefore, this topic is summarized as consumers’ feelings about the product’s usage effectiveness. Usage effectiveness centers on consumers’ evaluations of actual usage experiences, performance, and need satisfaction, addressing “how the product works”—such as whether it caters to office, study, or gaming scenario demands (e.g., height

adjustment for different body types, convenience of switching from sitting to standing), and whether health benefits are perceived (e.g., relief from cervical and lumbar pressure).

In Topic 4, the high-weight words mainly focus on customer service (“service”, “customer”, “support”, etc.) and the logistics support and transportation speed (“delivery”, “ship”, “fast”, “arrive”, etc.). Therefore, it is speculated that this topic can be summarized as the post-purchase support. Given the independent operation mode of Company L’s independent website business, its customer service system and logistics distribution need to be managed independently. The words discussed by consumers also reflect, from an indirect perspective, Company L’s control ability in independent channels.

In Topic 5, the high-weight words mainly concentrate on the design details (“hole”, “cable”, “drill”, etc.) and the assembly process (“instruction”, “assembly”, “screw”, etc.). The online shopping consumption scenario combined with the product characteristics often requires users to assemble the product by themselves. Therefore, consumers show a high level of concern for the assembly friendliness, and at the same time, it highlights the importance of the product’s detailed design.

Based on the LDA model, the probability distribution of each comment associated with the five topics is calculated, and the topic to which the comment belongs is determined according to the highest probability distribution. Finally, each comment is assigned to its most relevant topic, and the topic distribution results of the entire comment set are shown in Figure 7.

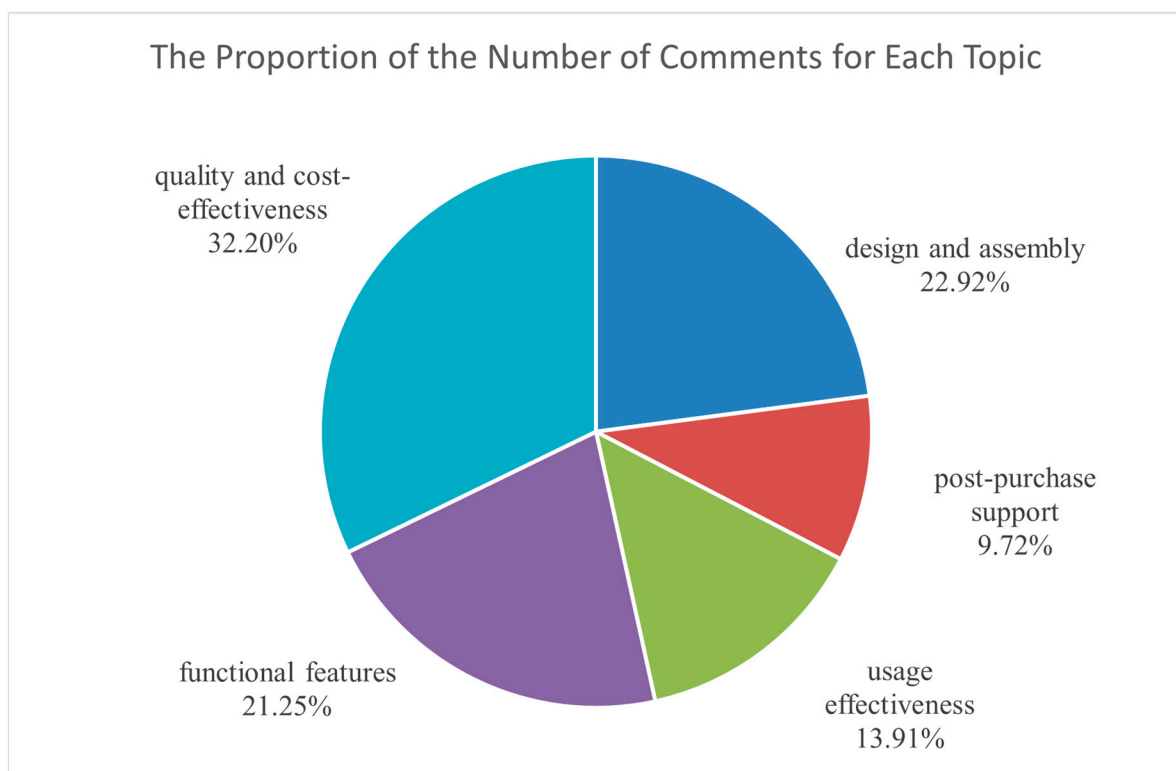


Figure 7. Topic distribution of the comment set.

4.4. Sentiment Classification Model Based on LSTM

After acquiring comment topic communities and their subdivided keywords, a sentiment classification model is applied to capture sentiment signals in the comment set and complete the discrimination of the comment set’s sentiment tendency. In this way, enterprises can not only grasp the current trends of product public opinion but also have an

in-depth understanding of consumers' emotional feedback, thus providing strong support for product optimization and marketing strategy adjustment.

4.4.1. Modeling Process

In the collected consumer evaluation text dataset, one dimension of the data type is the star rating given by users to the products, which is distributed in five levels from 1 star to 5 stars. In this study, this part of the data is manually labeled. In accordance with the general habit of online shopping evaluations, reviews with a rating of 3 stars or below were defined as negative evaluations, with a label value set to 0, while consumer reviews with a rating of 4 stars and 5 stars were defined as positive evaluations, with a label value set to 1, to construct a sentiment binary classification model.

Since the data cleaning part of the LDA model is the same as that of the LSTM, the data preprocessing will not be repeated here. After pre-training word vectors with Word2Vec, this study constructed and trained an LSTM model using the Keras library. Of the original dataset, 80% was used as the training set. After training, the sentiment polarity of the model was predicted based on the 20% test set, and finally, the model's performance was tested. The specific parameters of the model are shown in Table 4.

Table 4. The parameters of LSTM model.

Parameter Name	Parameter Value
LSTM units	128
vector size	100
seq_len	150
activation function	sigmoid
loss function	binary_crossentropy
optimizer	adam
epochs	10
batch_size	64
metrics	accuracy

4.4.2. Model Evaluation Indicators

A series of key indicators are often used when evaluating text classification models. For example, Accuracy represents the proportion of the number of comments that are correctly classified to the total number of comments; Precision is the proportion of comments that are actually positive among the comments predicted to be positive in terms of sentiment; Recall is the proportion of all comments with positive sentiment labels that are correctly predicted as positive comments by the model; The F-measure is often used to comprehensively consider Precision and Recall, as there are sometimes contradictory situations where an increase in the value of one leads to a decrease in the value of the other. Essentially, it is a weighted harmonic mean of Precision and Recall. These indicators all come from the confusion matrix, and the confusion matrix results of the LSTM model in this study are shown in Table 5.

Table 5. Confusion matrix table.

True Label	Predicted Label	
	Negative	Positive
Negative	80	18
Positive	158	2143

It clearly defines four classification results: True Positive (TP, that is, the number of correctly classified positive comments), True Negative (TN, that is, the number of correctly classified negative comments), False Positive (FP, that is, the number of negative comments misclassified as positive), and False Negative (FN, that is, the number of positive comments misclassified as negative). In this framework, P represents the comment text with a positive sentiment tendency, N corresponds to the comment text with a negative sentiment tendency, and T and F mark the correctness (True) and incorrectness (False) of the classification results, respectively.

4.4.3. Analysis of Sentiment Classification Results

After training the LSTM model, the evaluation results shown in Table 6 were obtained. The evaluation results indicate that the model's classification accuracy exceeds 92%, demonstrating excellent classification performance.

Table 6. Model evaluation results.

	Accuracy	Precision	Recall	F-Measure	Duration
LSTM	0.9266	0.9917	0.9313	0.9605	242.671 s

Therefore, by integrating the results of the LDA and LSTM models, review clusters distributed in the form of themes were sorted out and their sentiment classification was completed to observe differences in the quantity and proportion of positive/negative reviews across different thematic categories, as shown in Table 7.

Table 7. Details of comment topics.

Comment Topic	Number of Comments	Proportion of Negative Comments	Proportion of Positive Comments
quality and cost-effectiveness	3863	1.97%	98.03%
functional features	2549	2.55%	97.45%
design and assembly	2750	5.49%	94.51%
usage effectiveness	1669	12.46%	87.54%
post-purchase supports	1166	10.38%	89.62%

From the classification results, it can be seen that most users have given high positive evaluations of the product's quality and cost-effectiveness, functional features, and assembly instructions, indicating that consumers have a high degree of recognition in these three aspects. In terms of the number of comments, the number of comments on cost-effectiveness and functionality is in the first echelon, reflecting consumers' main concerns. In terms of usage effectiveness and post-purchase supports, although most of the comments are positive, the proportion of negative comments is also relatively prominent, indicating that there is still room for improvement in this regard.

4.5. The Evolution Trend of Topic Sentiment Based on Time Series

4.5.1. Trend of Thematic Sentiment Changes over the Years

The organization of the overall comment set has clarified the main focus areas of comments and intuitively presented the sentiment distribution under different topics, providing data support for a deeper understanding of consumer preferences and the optimization of operational strategies. In order to further understand the dynamic changes in consumers' key concerns about the product over time, the comment set is divided into

five subsets with a one-year time step, and the changing trends of topic sentiment in different time periods are explored respectively.

First, the perplexity is calculated respectively and combined with visualization to determine the number of topics in the comment set for each time period. After investigation, there is no overlap between the topic categories in each time period, and the perplexity is also at the lowest value. The numbers of topics in the five time periods are 2, 5, 2, 3, and 4, respectively, as shown in Figure 8. Subfigures (a)–(e) in Figure 8 show the topic perplexity of online comment subsets for each time period.

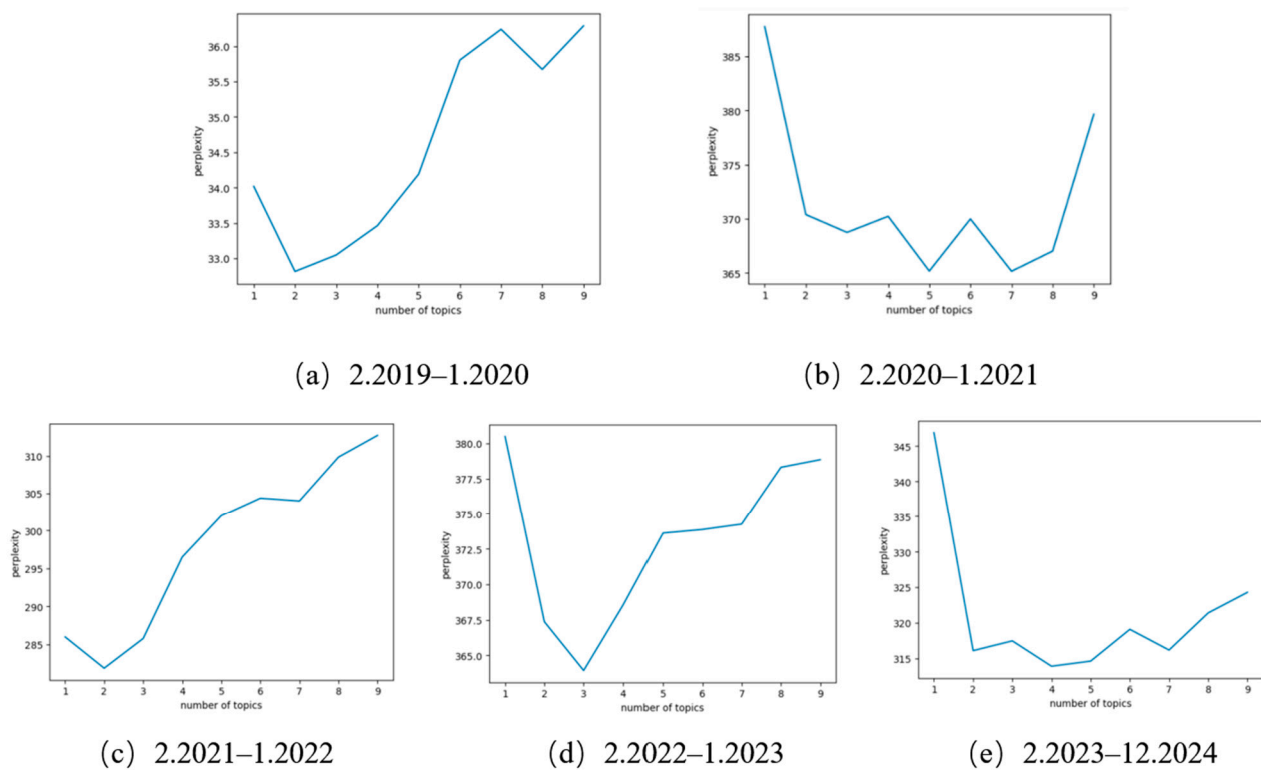


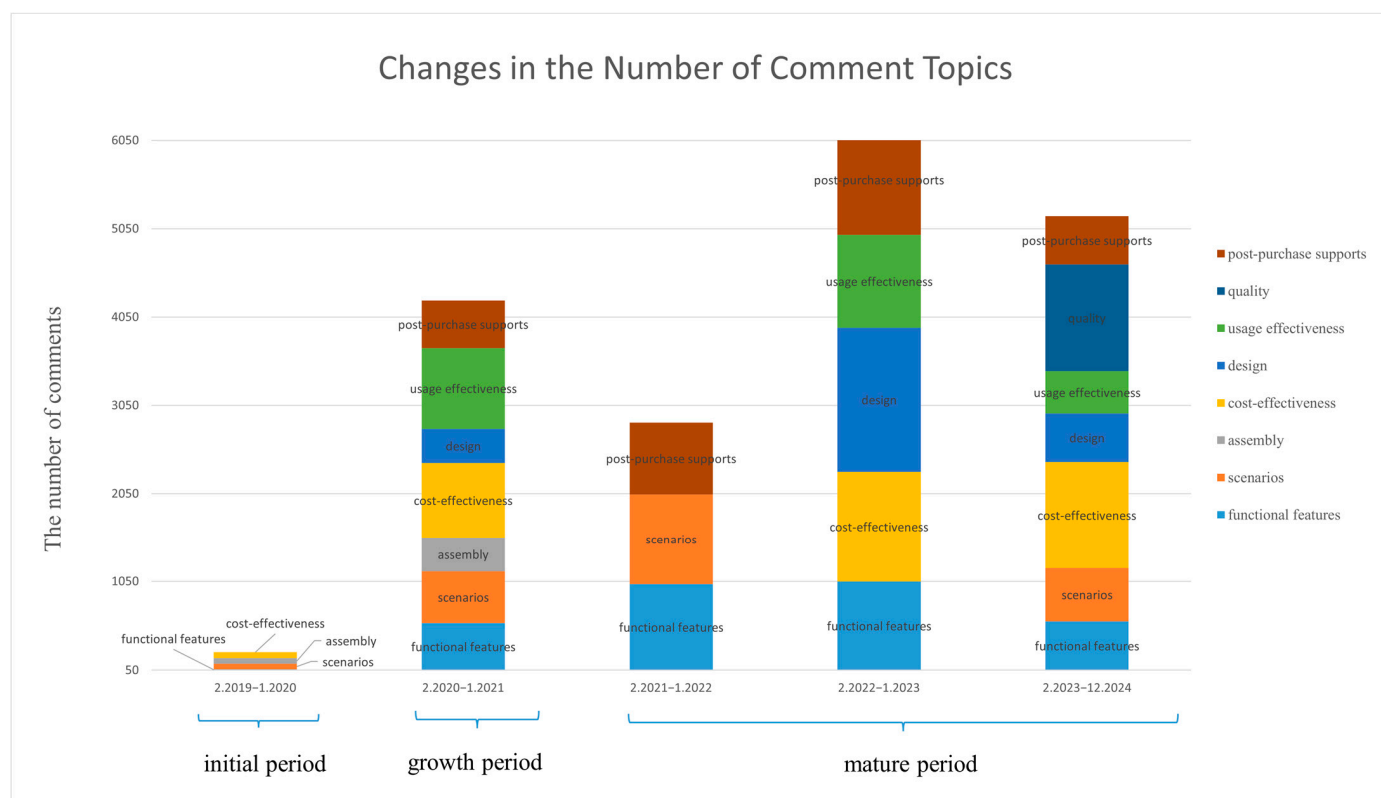
Figure 8. Selection of the number of topics in the comment subset.

Then, based on the sentiment classification results, the proportion of positive and negative comments for each topic is counted, the reputation changes of each topic across different time periods are observed, and the corresponding results are synthesized to obtain Table 8. To facilitate the collation of comment changes, the five time spans are divided into three stages according to the number of comments. The period from 2019 to 2020 is defined as the initial stage of the product, the period from 2020 to 2021 is defined as the growth stage of the product, and the period from 2021 to 2024 is defined as the mature stage of the product.

Table 8 was sorted based on the number of changes in comment topics, and the results are presented in the following Figure 9. In the topic categories of each year, some topic titles are named in a parallel form to summarize the characteristic words of the current topic category. However, the parallel-form topics lack temporal coherence, which hinders the intuitive presentation of topic trends and fails to highlight comment features. Therefore, the parallel topic titles are split, and the original number of comments and the proportion of sentiments are shared. There is a cumulative part in the number of comments, which is only for reference in showing the topic changes.

Table 8. Overview of the topic sentiment distribution.

Time Period	Stage	Topics	Number of Comments	Negative	Positive
2.2019–1.2020	initial stage	functional features & scenarios	62	0%	100%
		assembly & cost-effectiveness	64	0%	100%
2.2020–1.2021	growth stage	cost-effectiveness	849	2.24%	97.76%
		assembly & design	386	0.78%	99.22%
		functional features & scenarios	582	0.52%	99.48%
		usage effectiveness	915	2.19%	97.81%
		post-purchase supports	537	17.69%	82.31%
2.2021–1.2022		functional features & scenarios	1021	4.31%	95.69%
		post-purchase supports	815	6.87%	93.13%
2.2022–1.2023	mature stage	cost-effectiveness & post-purchase supports	1249	6.00%	94.00%
		design	1631	4.84%	95.16%
		functional features & usage effectiveness	1050	9.05%	90.95%
		usage effectiveness	479	4.18%	95.82%
2.2023–12.2024		quality & cost-effectiveness	1208	0.99%	99.01%
		design & post-purchase supports	548	1.82%	98.18%
		functional features & scenarios	601	4.83%	95.17%

**Figure 9.** The changing trend diagram of the number of topics.

From the perspective of the changing trend of the number of topics, during the initial accumulation stage of the product, consumers' comments are very limited, mainly focusing on four subdivided topics: product functionality, scenarios, cost-effectiveness, and assembly, emphasizing the experience during the usage process. As sales increase and the

product enters the growth stage, consumers' hot topics of concern about the product are further deepened and expanded, growing to seven subdivided topics. Based on the original foundation, it further highlights product effectiveness, design details, and post-purchase support. During the spread of the global pandemic, the consumer demand for online shopping was greatly activated. Coupled with the fact that working from home has become the norm, it jointly promoted the rapid growth of sales and prompted the product to quickly enter the mature stage. In this stage, consumers mainly evaluate the product in terms of four core aspects: functional features, cost-effectiveness, design details, and post-purchase support, and the corresponding topics also have a certain continuity in time. In terms of commonalities, topics such as functional features, design details, cost-effectiveness, applicable scenarios, and post-purchase supports run through the last two development stages of the product and become the main focus of consumers' comments. In the last part of the statistical scope, a new topic about product quality emerged, reflecting consumers' latest attention preferences.

By sorting the above topics according to the proportion of negative comments, the negative sentiment trend is obtained, as shown in the following Figure 10.

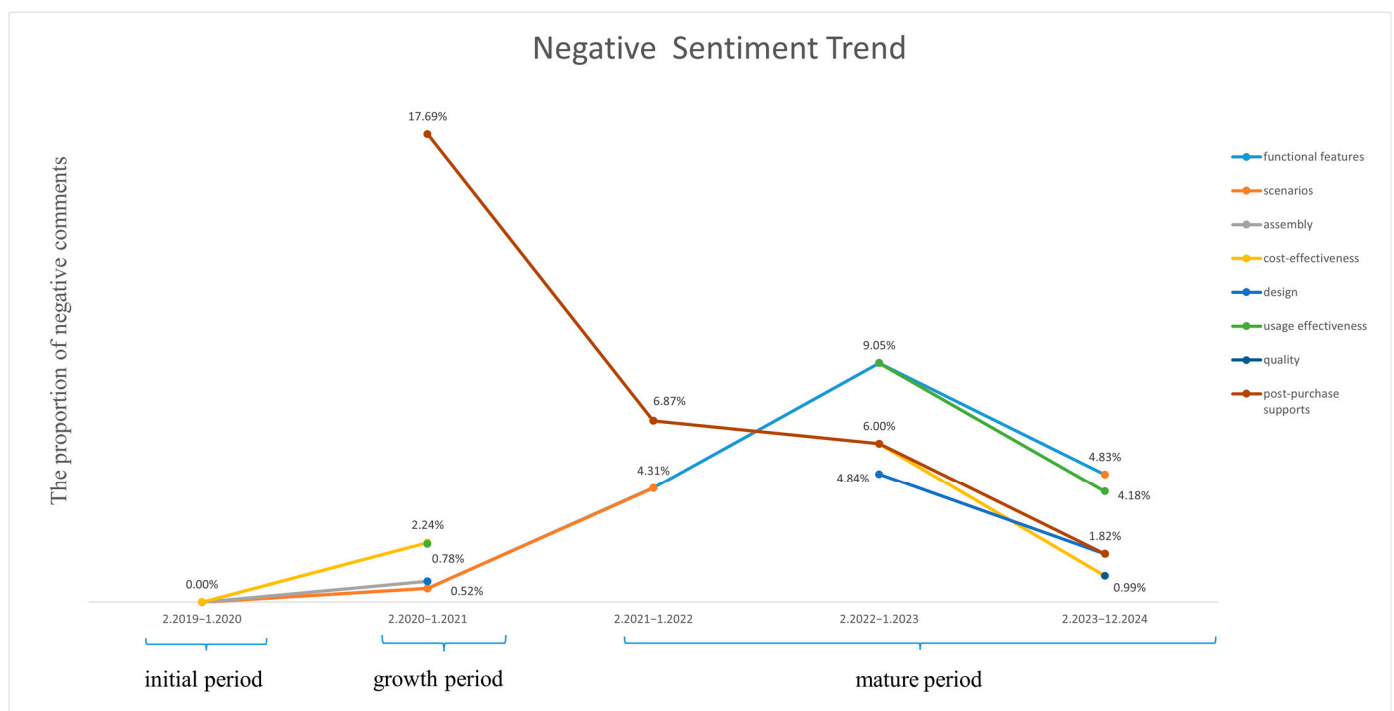


Figure 10. Negative sentiment trend.

From the perspective of the changing trend of sentiment tendency, due to the small number of comments in the initial stage, consumers have a high level of satisfaction with all aspects of the product, and there is less negative emotion. After entering the growth stage, the number of negative comments increases, with post-purchase support issues being the most prevalent, and the proportion of negative emotions reaches its peak in this stage. After entering the mature stage, all aspects of the product are being examined by consumers, and the types of exposed problems increase significantly. Among them, the proportions of negative emotions in the four types of topics, namely service, functional features, product effectiveness, and cost-effectiveness, are relatively high. However, there is a significant decline after 2023, indicating that merchants have targeted and optimized relevant issues according to users' comment feedback, effectively improving the pain

points and optimizing consumers' experiences in all aspects from the product itself to the shopping process. Therefore, the proportion of negative comments decreases.

4.5.2. Trend of Thematic Sentiment Changes During and After the Pandemic

Considering that the dataset's time range (February 2019–December 2024) partially overlaps with the COVID-19 pandemic period, the widespread adoption of large-scale home-office models has increased consumer demand, usage frequency, and attention toward height-adjustable desks. Therefore, this study divides the corpus into two time periods: February 2019–May 2023 and June 2023–December 2024 (as the World Health Organization declared an end to the global COVID-19 emergency in May 2023)—to further explore changes in consumer review themes during the pandemic period and the post-epidemic stage when the market gradually returned to normal.

Following the same procedure of determining the optimal number of topics based on one-year time increments as described above, the number of topics identified for the pandemic period and the post-pandemic period are four and two, respectively. Furthermore, the distribution of themes and negative emotions in these two periods is shown in Figure 11.

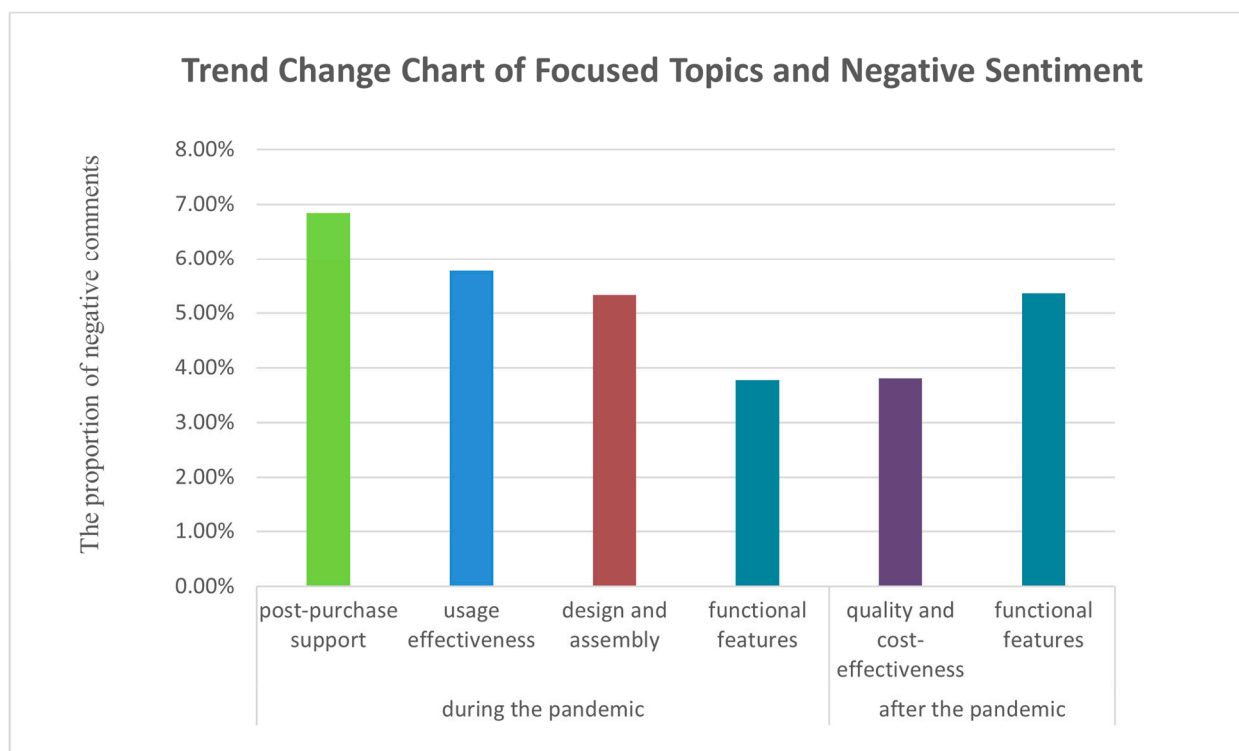


Figure 11. Topics and negative sentiment trends.

As indicated by the figure, there are some changes in the themes consumers focused on during the pandemic period and after its conclusion. During the pandemic, consumers primarily paid attention to four themes: design and assembly, functional features, post-purchase support, and usage effectiveness. After the pandemic, the main themes of focus shifted to functional features and quality and cost-effectiveness. Notably, functional features remained a common concern in both phases, reflecting consumers' continuous emphasis on product functionality. Additionally, negative comments on the themes of design and assembly and post-purchase support—prominent during the pandemic—did not reappear post-pandemic, suggesting a substantial reduction in consumer dissatisfaction in these areas.

In terms of emotional trends, post-purchase support had the highest proportion of negative reviews during the pandemic. This was primarily due to global logistics disruptions at the time, including delayed cross-border shipping and increased rates of package loss or damage, which hindered timely responses to consumers' after-sales repair and return requests. Pandemic control policies worldwide (such as lockdowns and customs clearance delays) further complicated after-sales processes—for instance, requiring extra documents or quarantine procedures for returns. After the pandemic, negative comments regarding functional features increased, likely because consumers transitioned from “sole home-office use” to “hybrid office-home use”, leading to higher demands for product durability and versatility. Products not optimized for these new scenarios were more prone to negative evaluations. Furthermore, the emergence of negative comments on quality and cost-effectiveness post-pandemic indicates that with the abundance of market supply, consumers had more choices among similar products and tended to compare quality and price differences across brands, influencing their post-purchase evaluations.

Synthesizing the above trends in review themes and sentiment changes, it is clear that although consumer focus on the product varies across different years, core themes such as functional features, design details, product efficacy, and service experience are consistently reflected in different development stages of the product, forming common elements of consumer evaluations. Therefore, enterprises need to continue prioritizing these aspects. Whether divided by year or by pandemic period, the theme distribution in the final time segment shows that the “quality” theme has garnered close consumer attention, indicating a shift in focus from mere functional requirements to the pursuit of higher quality. Additionally, in terms of sentiment changes, the proportions of negative comments in the themes of “functional features” and “usage efficacy” remain relatively high, requiring attention and corresponding measures to address them.

5. Conclusions

This study analyzes the product review data on cross-border e-commerce websites in applying the SNA, LDA, and LSTM models. The findings suggest that, across the entire review dataset, consumers' primary focuses are centered on five dimensions: functional features, quality and cost-effectiveness, post-purchase support, design and assembly, and usage effectiveness. When segmented according to the time series, the theme of quality has become a new focus of consumers' preferences. In terms of the changes in review sentiment, the negative sentiment ratios for the themes of functional features and usage effectiveness remain relatively high. Based on the above, this study integrates the 4P and 4C theories to propose the following improvement suggestions for the marketing strategies of cross-border e-commerce in the context of digital cross-border operations:

Firstly, define products based on user needs and build a data-driven demand response ecosystem (Product $\times$ Consumer). In digital cross-border scenarios, AI-driven sentiment analysis can be used to deeply analyze global post-purchase reviews, capture high-frequency negative terms and latent expectation terms, and establish dynamic demand maps to reconstruct product logic from the origin of consumer needs. Product functions and designs should be promptly optimized for changes in specific usage scenarios, as these aspects are prone to triggering negative consumer evaluations. For example, in response to the high environmental requirements in European and American markets, cross-border e-commerce enterprises can not only prefer the selection of eco-friendly materials but also dynamically display carbon footprint data to align with ESG consumption trends. Additionally, setting up a “Product Lab” section on independent websites invites core users to participate in function voting and co-creation of appearance designs, transforming user ideas into limited-edition products. Meanwhile, a cross-border quality traceability system

can be established, leveraging blockchain technology to enable consumers to inquire about raw material origins, production processes, quality inspection reports, etc., to strengthen “perceivable quality” through transparent production.

Secondly, for comprehensive cost optimization, dynamic pricing and full-link cost deconstruction should be implemented ($\text{Price} \times \text{Cost}$). Cross-border e-commerce can use dynamic algorithms for real-time price adjustment—for example, predicting demand peaks based on Google Trends data to apply “stepwise price increases” for upcoming popular products (such as locking in traffic with early-bird prices 30 days in advance and gradually adjusting to market average prices as the peak season approaches), triggering “smart clearance” for long-tail products through inventory turnover rates and competitor price-comparison APIs, and categorizing products into clearly labeled entry-level, advanced, and premium tiers to strengthen internal differentiation and meet diverse consumer needs and budgets. In optimizing cross-border circulation costs, enterprises can collaborate with local logistics providers to build a “smart pre-positioned warehouse system” that forecasts regional demand hotspots and deploys high-unit-price or fragile items to bonded warehouses in target markets in advance to reduce logistics costs. Additionally, developing a “group order cost calculator” to automatically match users in the same region, share international shipping discounts, and lower consumer purchase costs is recommended.

Thirdly, to improve the purchase process convenience, intelligent reconstruction of omni-channel touchpoints is essential ($\text{Place} \times \text{Convenience}$). Redefine channels as a “consumer touchpoint network” and optimize the entire cross-border shopping process with convenience as the core. At the traffic entrance stage, implement a dual-track strategy of “search engines + local platforms” for different markets: strengthen Google SEO in English-speaking markets by enhancing organic rankings through “long-tail keyword matrices + video structured data”; in non-English-speaking markets, access local search engines and localize keywords according to local grammatical habits. On the shopping experience layer, introduce immersive technologies such as AR fitting and VR showrooms to reduce the “physical cognition gap” in cross-border shopping. In the payment phase, integrate multi-currency wallets and offer “green logistics options” (such as degradable packaging and carbon offset credits) to meet the needs of environmentally conscious consumers.

Finally, bilateral value transmission is indispensable ($\text{Promotion} \times \text{Communication}$). Cross-border e-commerce enterprises should break through the one-way communication of traditional promotions and establish a value symbiosis system between brands and consumers through communication as the link. At the content marketing level, launching a “Global Experience Officer Program” to invite KOCs (Key Opinion Consumers) from different countries to participate in product internal testing, releasing content in the form of “unboxing vlogs + localized scenario challenges” on platforms such as TikTok and Instagram Reels, and automatically generating multilingual versions of UGC content to reduce cross-cultural communication costs. In crisis communication scenarios, developing a “negative review response robot” that automatically triggers a “2 h rapid feedback” process for comments with strong negative sentiment, while simultaneously generating multilingual solutions (such as return and exchange guidance videos) to transform negative word-of-mouth into service value-added points.

This study builds a closed-loop ecosystem of “demand insight–product co-creation–experience optimization–value symbiosis” based on post-purchase review data of hot-selling products from cross-border e-commerce Company L, through text mining and deep integration of the 4P and 4C theories. This approach maintains supply chain efficiency on the enterprise side while strengthening emotional identification on the consumer side, providing references for cross-border e-commerce enterprises to enhance marketing effectiveness in global markets.

However, the study has certain limitations: From a product perspective, selecting products from a single brand may have unique brand styles and product characteristics, meaning its online post-purchase reviews may not fully represent product conditions across the entire industry. From a review perspective, ignoring high-star reviews may lead to an incomplete understanding of product advantages and overall consumer satisfaction, potentially missing some positive consumer feedback and the product's competitive advantages in the market. Future research could break through the limitations of single products; cover a wide range of product categories; deeply explore differences in consumer groups, product attributes, and application scenarios across categories; analyze corresponding review texts; and form more universally applicable business management recommendations.

Author Contributions: Conceptualization, T.C. and Q.P.; methodology, Q.P.; software, C.L. and Q.P.; validation, C.L.; formal analysis, C.L. and T.C.; investigation, C.L. and Q.P.; resources, C.L. and Y.J.; data curation, C.L. and Y.J.; writing—original draft preparation, C.L.; writing—review and editing, T.C. and Y.J.; visualization, C.L.; supervision, T.C.; project administration, T.C.; funding acquisition, T.C. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the National Social Science Foundation of China (23BJY073) and the Merchants' Guild Economics and Cultural Intelligent Computing Laboratory Research Project of Ningbo University (CGEC2024Z04).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: We confirm that the data used in this study were collected from publicly accessible websites, and that the process of scraping consumer online review data for certain products complied with the platform's terms of service. During the data collection process, we strictly adhered to ethical guidelines and refrained from collecting or storing any personally identifiable information, and the data were solely utilized for this academic research, ensuring the protection of the privacy and rights of all relevant users.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Fathy, E.A.; Salem, I.E.; Zidan, H.A.Y.; Abdien, M.K. From plate to post: How foodstagramming enriches tourist satisfaction and creates memorable experiences in culinary tourism. *Curr. Issues Tour.* **2024**, 1–20. [[CrossRef](#)]
2. Hofstede, G. *Culture's Consequences: International Differences in Work-Related Values*; Sage Publications: Thousand Oaks, CA, USA, 1984.
3. Elshaer, I.A.; Azazz, A.; Fayyad, S.; Mohamed, S.A.; Fouad, A.M.; Fathy, E.A. From Data to Delight: Leveraging Social Customer Relationship Management to Elevate Customer Satisfaction and Market Effectiveness. *Information* **2025**, 16, 9. [[CrossRef](#)]
4. Wang, X.; Guo, J.; Wu, Y.; Liu, N. Emotion as signal of product quality: Its effect on purchase decision based on online customer reviews. *Internet Res.* **2020**, 30, 463–485. [[CrossRef](#)]
5. Choi, H.S.; Leon, S. An empirical investigation of online review helpfulness: A big data perspective. *Decis. Support Syst.* **2020**, 139, 113403. [[CrossRef](#)]
6. Hong, H.; Xu, D.; Wang, G.A.; Fan, W. Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decis. Support Syst.* **2017**, 102, 1–11. [[CrossRef](#)]
7. Ruijuan, W.; Peiyu, L.; Yan, L. The Effect of Image Number in Online Consumer Reviews on Review Helpfulness—An Empirical Study Based on Negative Reviews. *Manag. Rev.* **2022**, 34, 157.
8. Yan, H.; Liao, Q.; Xiong, H. Emotional or non-emotional? The impact of emojis on the usefulness of online restaurant reviews. *Tour Trib* **2024**, 39, 145–160.
9. Zhu, Z.; Fang, X.; Shan, M. The effect of language style on consumers' perceived usefulness of online reviews: From the perspective of regulatory focus. *Nankai Bus. Rev.* **2024**, 27, 234–246.
10. Kong, J.; Lou, C. Do cultural orientations moderate the effect of online review features on review helpfulness? A case study of online movie reviews. *J. Retail. Consum. Serv.* **2023**, 73, 103374. [[CrossRef](#)]

11. Park, J.; Park, H. Understanding the Impact of Inconsistency on the Helpfulness of Online Reviews. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 80. [\[CrossRef\]](#)
12. Guo, L.; Li, J. Impact of Consumers' Perception of Crisis Risks on the Effect of Online Reviews. *J. Northeast. Univ. (Nat. Sci.)* **2022**, *43*, 1662.
13. Flavián, C.; Akdim, K.; Casaló, L.V. Effects of voice assistant recommendations on consumer behavior. *Psychol. Mark.* **2023**, *40*, 328–346. [\[CrossRef\]](#)
14. Riswanto, A.L.; Ha, S.; Lee, S.; Kwon, M. Online Reviews Meet Visual Attention: A Study on Consumer Patterns in Advertising, Analyzing Customer Satisfaction, Visual Engagement, and Purchase Intention. *J. Theor. Appl. Electron. Commer. Res.* **2024**, *19*, 3102–3122. [\[CrossRef\]](#)
15. Wang, B.; Zhao, Q.; Zhang, Z.; Xu, P.; Tian, X.; Jin, P. Understanding the Heterogeneity and Dynamics of Factors Influencing Tourist Sentiment with Online Reviews. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 22. [\[CrossRef\]](#)
16. Zhang, C.; Liu, Y.; Zhao, B.; Chai, J.; Jiang, F. Research on destination image mining of “tang culture” based on online text sentiment analysis. *China J. Econom.* **2023**, *3*, 387–407.
17. Cheng, F.M.; Wang, J.; Chen, C.; Hu, G.R.; Cao, Z.J. Product design improvement method driven by online product reviews. *Sci. Rep.* **2025**, *15*, 10252. [\[CrossRef\]](#)
18. McCarthy, E.J.; Shapiro, S.J.; Perreault, W.D. *Basic Marketing*; Irwin-Dorsey: Toronto, ON, Canada, 1979.
19. Yu, Y.; Suteeca, R. A Study of Factors Affecting the Free Trade of Thai Cocoa Products in China's E-commerce Market of Hill Tribe Cocoa Cof Co., Ltd. In Proceedings of the 2024 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON), Chiang-mai, Thailand, 31 January 2024; IEEE: New York, NY, USA, 2024; pp. 336–341.
20. Lan, Y.Y.; Qu, X.T. Research on B2C E-commerce Marketing Strategies Using Jingdong Mall as an Example. In Proceedings of the 3rd Annual International Conference on Management, Economics and Social Development (ICMESD 17); Atlantis Press: Bedford Park, IL, USA, 2017; pp. 557–563.
21. Li, L.; Mao, Z.; Ren, Y. E-commerce Precision Marketing Based on the Advantages of Big Data Technology. *For. Chem. Rev.* **2021**, *70*–79.
22. Lauterborn, B. New marketing litany: Four Ps passé: C-words take over. *Advert. Age* **1990**, *41*, 26.
23. He, Y.; Li, J. On Marketing Modes of ICBECE Based on 4C Principles-A Case Study of Mia. In Proceedings of the 2017 International Conference on Applied Mathematics, Modelling and Statistics Application (AMMSA 2017); Atlantis Press: Bedford Park, IL, USA, 2017; pp. 270–274.
24. Lei, S. Online Marketing Factors Influencing Shopping Decisions Through Cross-Border E-commerce Platform. In Proceedings of the International Academic Multidisciplinary Research Conference in Geneva 2022, Geneva, Switzerland, 3–6 October 2022; pp. 25–29.
25. Maidar, U.; Ra, M.; Yoo, D. A Cross-Product Analysis of Earphone Reviews Using Contextual Topic Modeling and Association Rule Mining. *J. Theor. Appl. Electron. Commer. Res.* **2024**, *19*, 3498–3519. [\[CrossRef\]](#)
26. Goldman, S.P.; van Herk, H.; Verhagen, T.; Weltevreden, J.W. Strategic orientations and digital marketing tactics in cross-border e-commerce: Comparing developed and emerging markets. *Int. Small Bus. J.* **2021**, *39*, 350–371. [\[CrossRef\]](#)
27. Feng, D.; Jun, C.; Chao, C.; Jia-zhen, H. Study on cross-border e-commerce competitive differentiation strategy. *Oper. Res. Manag. Sci.* **2019**, *28*, 33–40.
28. Cai, Y.; Liu, X. AI-Driven Social Media E-commerce Advertising: A Cross-Cultural Communication Study from the Perspective of Yiwu's Trade and Commerce. *Sociol. Philos. Psychol.* **2024**, *1*, 20–32. [\[CrossRef\]](#)
29. Xiong, J. Research on the Construction Strategy of Independent Stations for Cross-border E-commerce Enterprises. *J. Jiaozuo Univ.* **2021**, *35*, 84–87.
30. Li, X.; Ke, Q. Strategic Research on the Development of China-ASEAN Cross-border E-commerce under the Background of the Belt and Road Initiative. *Pract. Foreign Econ. Relat. Trade* **2019**, *9*, 33–36.
31. Liu, Y. Optimization of Export Marketing Strategies for Cross-border E-commerce under the Belt and Road Initiative. *Pract. Foreign Econ. Relat. Trade* **2019**, *2*, 36–39.
32. Wu, Z.; Zhang, S. Research on the Strategies of China's Agricultural Products Cross-border E-commerce under the Background of the Belt and Road Initiative. *Bus. Econ.* **2021**, *7*, 79–81.
33. Chen, W.H.; Lin, Y.C.; Bag, A.; Chen, C.L. Influence factors of small and medium-sized enterprises and micro-enterprises in the cross-border e-commerce platforms. *J. Theor. Appl. Electron. Commer. Res.* **2023**, *18*, 416–440. [\[CrossRef\]](#)
34. Wang, G.; Zhang, Z.; Li, S.; Shin, C. Research on the influencing factors of sustainable supply chain development of agri-food products based on cross-border live-streaming e-commerce in China. *Foods* **2023**, *12*, 3323. [\[CrossRef\]](#)
35. Fan, M.; Tang, Z.; Qalati, S.A.; Tajeddini, K.; Mao, Q.; Bux, A. Cross-Border e-commerce brand internationalization: An online review evaluation based on Kano model. *Sustainability* **2022**, *14*, 13127. [\[CrossRef\]](#)

36. Jiang, H.; Cai, J.; Lin, Y.; Wang, Q. Understanding the effect of TikTok marketing on user purchase behavior: A mixed-methods approach. *Electron. Commer. Res.* **2024**, 1–36. [\[CrossRef\]](#)
37. Han, L.; Han, X. Improving the service quality of cross-border e-commerce: How to understand online consumer reviews from a cultural differences perspective. *Front. Psychol.* **2023**, 14, 1137318. [\[CrossRef\]](#)
38. Giuffrida, M.; Mangiaracina, R.; Perego, A.; Tumino, A. Cross-border B2C e-commerce to China: An evaluation of different logistics solutions under uncertainty. *Int. J. Phys. Distrib. Logist. Manag.* **2020**, 50, 355–378. [\[CrossRef\]](#)
39. Zhang, Y.; Huang, H. Unraveling how poor logistics service quality of cross-border E-commerce influences customer complaints based on text mining and association analysis. *J. Retail. Consum. Serv.* **2025**, 84, 104237. [\[CrossRef\]](#)
40. Huang, L.; Ma, L. A protective buffer or a double-edged sword? Investigating the effect of “parasocial guanxi” on consumers’ complaint intention in live streaming commerce. *Comput. Hum. Behav.* **2024**, 151, 108022. [\[CrossRef\]](#)
41. Guo, X.; Ma, S.; Zhang, H. Does cross-border e-commerce improve China’s imported product quality?—A quasi-natural experiment based on the establishment of cross-border e-commerce comprehensive pilot zones. *J. Int. Trade Econ. Dev.* **2024**, 1–24. [\[CrossRef\]](#)
42. Duan, W.; Gu, B.; Whinston, A.B. Do online reviews matter?—An empirical investigation of panel data. *Decis. Support Syst.* **2008**, 45, 1007–1016. [\[CrossRef\]](#)
43. Ramanathan, U.; Ramanathan, R. Investigating the impact of resource capabilities on customer loyalty: A structural equation approach for the UK hotels using online ratings. *J. Serv. Mark.* **2013**, 27, 404–415. [\[CrossRef\]](#)
44. Ghasemaghahi, M.; Eslami, S.P.; Deal, K.; Hassanein, K. Reviews’ length and sentiment as correlates of online reviews’ ratings. *Internet Res.* **2018**, 28, 544–563. [\[CrossRef\]](#)
45. Zhou, R.; Zhu, Y.; Zheng, R.; Zhou, J. Differences of consumer reviews and consumer satisfaction on cross-border e-commerce platforms: A text mining analysis based on cross-cultural perspective. *J. Organ. Comput. Electron. Commer.* **2025**, 1–23. [\[CrossRef\]](#)
46. Huang, Y.; He, Z.; Lv, H.; Min, J. Research on Mining Negative Online Reviews on E-commerce Platforms Based on Social Network Analysis and LDA Model. In *Intelligent Management of Data and Information in Decision Making: Proceedings of the 16th FLINS Conference on Computational Intelligence in Decision and Control & the 19th ISKE Conference on Intelligence Systems and Knowledge Engineering (FLINS-ISKE 2024)*; World Scientific Publishing: Singapore, 2024; pp. 177–185.
47. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, 3, 993–1022.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.