Journal of
Theoretical and Applied Electronic Commerce Research

MDPI

Article

User Requirements Analysis for Audiovisual Products Based on User Review Data

Chuchu Liu [1,2,*], Xin Zhang [3], Mengsi Cai [2] and Zheng Han [1]

[1] School of Economics and Management, Changsha University of Science and Technology, Changsha 410114, China; zhenghan0417@163.com
[2] College of Systems Engineering, National University of Defense Technology, Changsha 410073, China; caimengsi18@163.com
[3] College of Economics, Shenzhen University, Shenzhen 518060, China; 2500123038@mails.szu.edu.cn
* Correspondence: liuchuchu15@nudt.edu.cn; Tel.: +86-13548692656

Abstract

This study analyzed online review data to examine user requirements for audiovisual products and to compare requirement salience and satisfaction across traditional and emerging product contexts. We collected 86,213 Chinese-language reviews of Skyworth TVs, Xiaomi TVs, and Xiaomi projectors from JD.com. LDA topic modeling was used to identify major user requirement areas, and Logistic Regression, Random Forest, and Support Vector Machine (SVM) models were compared for sentiment classification, with the tuned SVM model retained for downstream analysis. The results show that user discussions primarily concern audiovisual experience, cost performance, service quality, design aesthetics, and intelligent operation. Skyworth TVs receive particularly strong evaluations for picture and sound quality (97.89% positive sentiment), whereas Xiaomi TVs are more strongly associated with cost-effectiveness and smart features (94.05% positive sentiment). Xiaomi projectors attract attention for portability but receive lower satisfaction ratings on core audiovisual performance and intelligent operation. These findings suggest that traditional manufacturers should continue strengthening core performance while improving service responsiveness, whereas emerging brands should build on their technological advantages while further enhancing their product reliability and user experience.

Keywords: user requirements analysis; online reviews; audiovisual products; topic modeling; sentiment analysis

Check for updates

Academic Editor: Jorge Bernardino

Received: 25 January 2026
Revised: 14 May 2026
Accepted: 18 May 2026
Published: 20 May 2026

Copyright: © 2026 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.

1. Introduction

The rapid advancement of information technology has profoundly shaped the digital economy, while the widespread adoption of the Internet and smart devices has significantly transformed consumer behavior, particularly in the audiovisual entertainment sector [1,2]. Products such as televisions and projectors have evolved from traditional household appliances into integral components of modern digital lifestyles, serving as key media for home entertainment and information acquisition. Consequently, consumer expectations regarding product performance, user experiences, and service quality have escalated.

Traditionally, competition in the traditional television market mainly focused on hardware features, such as picture quality, sound performance, and aesthetic design. However, since 2014, the integration of intelligent technologies with the entry of Internet-based enterprises, such as Xiaomi and LeTV, into the television industry has intensified market competition, accelerated technological innovation, and fundamentally reshaped consumer

expectations [3,4]. This paradigm shift has transformed televisions from electronic media to network terminals, from shared family viewing to personalized media centers, and from hardware-focused devices to multifunctional digital platforms. This transformation has redefined television from a unidirectional medium into an interactive service hub and from a tangible device into an unlimited content ecosystem. As a result, competition between traditional audiovisual manufacturers and emerging smart device brands has become increasingly fierce and the profitability of traditional television leaders has declined due to the rapid rise of Internet-based smart television enterprises, where features such as smart operating systems, voice control, and content ecosystems have enriched user experiences and diversified the value propositions of audiovisual products. Consequently, consumers now value not only hardware performance but also the overall user experience and perceived value-added services.

In this competitive landscape, understanding the requirements of audiovisual experience users, defined here as individuals who access audiovisual content and interactive experiences via televisions and projectors, is crucial. Online reviews, as a primary form of user-generated content (UGC), provide a rich and scalable data source for capturing such evolving user requirements. Unlike structured survey data, review texts simultaneously encode functional evaluations, emotional responses, and contextual usage experiences. A growing body of literature demonstrates that online reviews not only influence demand and product diffusion but also serve as valuable inputs for requirement identification and product improvement [5,6]. Moreover, recent advances in text mining and machine learning have enabled researchers to extract fine-grained insights from large-scale unstructured data, bridging the gap between user expression and managerial decision-making.

This study employed machine learning and natural language processing (NLP) techniques to analyze online reviews of TVs and projectors deeply. It aimed to (1) identify key user requirements (e.g., picture quality, sound performance, cost-effectiveness), (2) compare satisfaction levels between traditional brands (e.g., Skyworth) and emerging brands (e.g., Xiaomi), (3) perform sentiment analysis to evaluate user feedback, and (4) provide actionable recommendations for product improvement. By using an interpretable review-mining workflow, this study aimed to address a substantive gap in the audiovisual product context. Specifically, this study identified the requirement dimensions discussed in reviews, linked those dimensions to sentiment outcomes, and compared the resulting patterns across Skyworth TVs, Xiaomi TVs, and Xiaomi projectors.

To achieve these objectives, we collected 86,213 user reviews from JD.com. We employed Latent Dirichlet Allocation (LDA) for topic modeling and requirements identification and utilized supervised machine learning models, including Support Vector Machine (SVM), Logistic Regression (LR), and Random Forest (RF), for sentiment classification. This integrated approach facilitated a systematic exploration of audiovisual experience users' requirements and satisfaction disparities between traditional and emerging brands.

2. Related Work

With the rapid expansion of the digital economy, the Internet has become an essential medium for individuals to acquire information, express opinions, and engage in social interaction. Online reviews, as a key form of user-generated content (UGC), serve not only as rich sources of information for consumers but also as valuable data assets for firms to gain insights into user requirements, improve products, and optimize services. The authenticity and credibility of online reviews have long been a focal topic of academic inquiry. Harrison-Walker and Azer Jaylan [7,8] demonstrated that consumers tend to trust information provided by third parties, perceiving it as more reliable than promotional communications. However, not all reviews are authentic, and the ability of consumers

to detect fake reviews depends on multiple factors, such as linguistic style, emotional tone, and reviewer characteristics. Regarding the influence of online reviews on consumer behavior, Cho et al. [9] empirically revealed the interactive effects between star ratings and textual sentiment on product requirement, highlighting their complementary nature and managerial relevance for review management.

Regarding methods for user requirements elicitation, advancements in big data analytics and artificial intelligence have inspired a range of data-driven approaches, including STandR-BUI [10], the BERTopic text mining model [11], the ELMo language model [12], the STM method [13], the Bi-LSTM model [14], and the distributional semantic model [15]. Cai et al. [16] provided a comprehensive review of automatic requirements extraction from UGC, emphasizing the roles of text clustering, topic modeling, and machine learning techniques in requirements engineering. Addressing the challenge of sparse review data during early product iterations, Cong et al. [17] proposed a small-sample-based requirements extraction method that integrates pre-trained language models (e.g., ERNIE) with an improved SIFRank algorithm to effectively identify critical user requirements. Similarly, Zhang et al. [18] developed a framework for modeling and tracking user requirements based on online reviews, classifying product attributes into functional, structural, and parametric dimensions. By combining sentiment analysis with evolutionary modeling, their approach provides actionable insights for guiding product upgrades and innovation.

In the domain of text-based sentiment analysis, the proliferation of social media platforms has generated vast volumes of user-generated text data, offering rich material for understanding user emotions and attitudes [19–26]. Chen and Mankad [27] introduced a Structural Topic and Sentiment Discourse (STS) model capable of jointly capturing topic distributions, sentiment orientations, and contextual variables, which significantly improved the model's performance during text regression tasks. From a multimedia perspective, Li et al. [28] found through panel data analysis that the proportion of user-generated images in online reviews positively affects the subsequent review volume and length, while the emotional incongruence between text and images exhibits a complex nonlinear impact on user engagement. Bauman and Tuzhilin [29] proposed a novel method for parsing contextual information from reviews to enhance recommendation accuracy. Moreover, by using a natural experiment design, Cao et al. [30] demonstrated that integrating heterogeneous information sources could impose a cognitive load and reduce the user contribution, providing managerial insights for content platform design.

Prior studies have made important progress in automatic requirement extraction and text-based sentiment analysis. However, two issues remain insufficiently addressed in the present research setting. First, most existing studies do not focus on audiovisual products, where product evaluation simultaneously involves audiovisual performance, intelligent interaction, design, service experience, and cost performance. Second, prior review-mining studies often treat topic extraction and sentiment classification as separate analytical tasks, whereas fewer studies explicitly connect requirement salience, sentiment polarity, brand type, and product category within the same empirical framework. Therefore, the contribution of this study is not the development of a new algorithm; rather, it lies in constructing an interpretable workflow that links LDA-based requirement identification, Doc2Vec-based document representation, supervised sentiment classification, and cross-category statistical comparison. This workflow extends prior online review-mining research by showing how unstructured review data can be transformed into domain-specific evidence on user requirements and satisfaction differences across Skyworth TVs, Xiaomi TVs, and Xiaomi projectors.

3. Data and Methods

3.1. Data Collection

This study used online user review data from JD.com, one of China's major e-commerce platforms for consumer electronics, as the focal data source for text analysis. JD.com provides a large volume of structured, time-stamped, and product-specific review data, which makes it suitable for requirement mining and sentiment analysis in a comparable product context. The dataset comprised user reviews for three representative audiovisual products—Skyworth TVs, Xiaomi TVs, and Xiaomi projectors—which were selected to capture the contrasts between traditional and emerging brands in the audiovisual market. Specifically, as a benchmark for conventional television manufacturers, Skyworth was founded in 1988 and currently has 183 million active users. Skyworth TVs ranks among the top three in the Chinese television market, where it is characterized by deep industry experience and broad market penetration, thus reflecting consumer perceptions of established brands. In contrast, Xiaomi represents an innovative entrant emphasizing intelligent design, affordability, and integration with the Internet of Things (IoT) ecosystem. Xiaomi TVs were launched in 2013, positioned as "the first TV for young people", and were the top-selling TV in China for three consecutive years from 2022 to 2024. Its products, particularly smart TV and portable projectors, exemplify new-generation audiovisual devices that cater to consumers' growing requirements for intelligent, interactive, and high-value experiences. Moreover, analyzing reviews for Xiaomi projectors, a rapidly emerging segment within home entertainment, offers insights into consumer expectations for portable, multifunctional, and smart viewing devices.

The three focal products were selected to support two complementary comparisons. Skyworth TVs and Xiaomi TVs enable a cross-brand comparison within the same product category, contrasting a traditional manufacturer with an emerging Internet-oriented brand under a broadly comparable usage scenario. Xiaomi projectors were included to extend the analysis from within-category brand differences to cross-category differences within the broader audiovisual market, especially for portable and smart viewing devices. Such cross-product and cross-brand analysis provides a comprehensive understanding of evolving user requirements and perceptions within the audiovisual product market. By comparing user reviews of Skyworth TVs with those of Xiaomi TVs and projectors, this study aimed to uncover the differentiated user requirements, satisfaction levels, and behavioral tendencies across product types and brand categories.

Data were collected from JD.com using the Octopus Collector web-scraping tool (https://www.octoparse.com/). A total of 86,213 reviews posted between January 2023 and December 2024 were obtained, comprising 40,002 reviews for Skyworth TVs, 34,139 for Xiaomi TVs, and 12,072 for Xiaomi projectors. Each record contained multiple attributes, including user ID, user level, review content, star rating, product size, review date, and review location, which formed a structured dataset suitable for text mining and sentiment analysis. This dataset supported a robust within-context analysis of differences across brands and product categories; however, because the corpus was drawn from a single platform, the findings should not be interpreted as fully generalizable to all e-commerce environments.

3.2. Data Preprocessing

Given that raw online review data often contain duplicate entries, invalid records, and irrelevant information, a comprehensive data cleaning procedure was conducted prior to the analysis. The preprocessing process involved three major steps: (1) removing duplicate reviews; (2) eliminating trivial or non-informative expressions such as "good review" or "recommended", which carry limited analytical value; (3) filtering out invalid records, including those with missing or placeholder content, such as "this user did not

provide a review". After completing the cleaning and normalization procedures, including deduplication and removal of mechanical or compressed text, 33,546 reviews of Skyworth TV, 25,251 reviews of Xiaomi TVs, and 8245 reviews of Xiaomi projectors were retained for subsequent analysis.

For text segmentation, the Jieba tokenizer was applied in precise mode to segment the cleaned Chinese review texts into lexical tokens. This segmentation process facilitated the construction of meaningful textual features required for downstream modeling. To further enhance processing accuracy and analytical efficiency, a combined stopword list that included a customized stopword dictionary and the standard list provided by Baidu was employed to remove high-frequency words lacking semantic relevance. This ensured that only linguistically and contextually meaningful words were preserved for topic modeling and sentiment classification.

Preparing the training data was a critical step for ensuring the robustness and generalizability of the machine learning models. To guarantee representativeness, 20% of the total dataset was randomly sampled from each product category to construct the labeled dataset for sentiment analysis. Manual annotation was conducted to assign sentiment labels to individual reviews (Table 1). Each review was carefully examined based on its linguistic context, expression, and emotional tone, and was categorized as positive (1) or negative (0). The manual labeling process produced a reliable gold-standard dataset for supervised learning. For model evaluation, the labeled dataset was randomly divided into training and testing subsets following an 80/20 split. A fixed random seed (random_state = 42) was applied to ensure the consistency and reproducibility of experimental results across runs. The resulting preprocessed dataset provided a solid empirical foundation for the subsequent phases of topic modeling and sentiment analysis, ensuring the data quality, analytical precision, and experimental reproducibility.

Table 1. Number of manual annotations.

Number of Positive Reviews	Number of Negative Reviews	Total
11,436	1971	13,407
85.20%	14.80%	100%

All reviews analyzed in this study were originally written in Chinese. Accordingly, Chinese word segmentation was performed using the Jieba tokenizer. For international presentation, the keywords, topic labels, and figure annotations reported in the manuscript were translated from Chinese into English while preserving their substantive meanings.

3.3. Experimental Procedure

The overall experimental framework is shown in Figure 1. The workflow comprised three stages: text vectorization, topic modeling, and sentiment analysis. First, the review texts were segmented with Jieba and converted into dense document representations using Doc2Vec. Next, LDA was applied to identify latent requirement-related topics. Finally, LR, SVM, and RF classifiers were trained and compared, and the best-performing model was used for downstream sentiment classification.

3.3.1. Text Vectorization

Although traditional bag-of-words (BoW) representations are intuitive, they often fail to capture complex semantic and contextual relationships. To address this, we employed the Doc2Vec model to generate distributed representations of review documents. Originally proposed by Mikolov et al. [31], Doc2Vec extends the Word2Vec framework to the

document level, learning continuous vector spaces that encode both semantic similarity and contextual relevance.

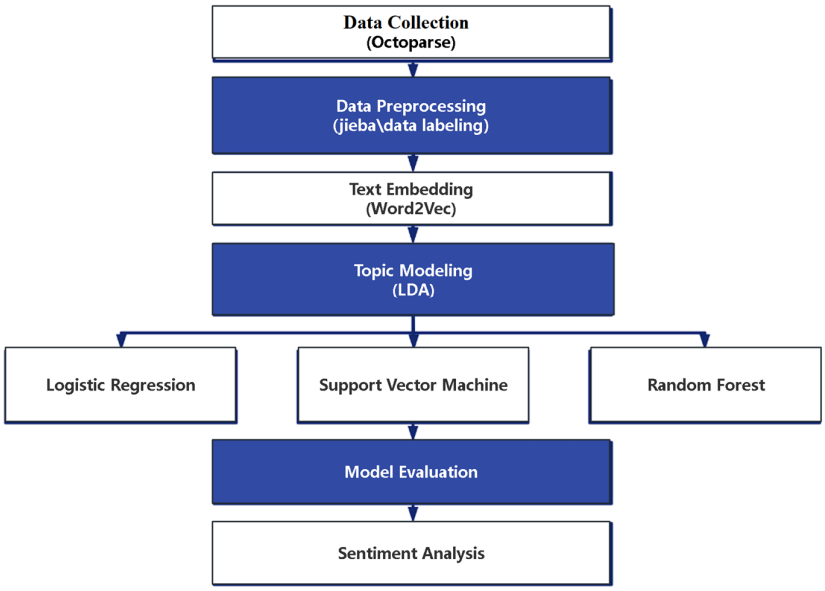


Figure 1. Overall research framework, including the comparison of LR, SVM, and RF models.

Doc2Vec was selected for this study primarily for its methodological parsimony and robust compatibility with classical classifiers (e.g., LR, SVM, and RF) within a reproducible machine-learning pipeline. Compared to transformer-based architectures like BERT, Doc2Vec offers lower computational overhead and a more transparent integration process, making it particularly suitable for large-scale corpora where an interpretable baseline workflow is prioritized over algorithmic complexity. While acknowledging that transformer-based models may capture richer nuances in short or ambiguous texts, this study utilizes Doc2Vec as an efficient, validated representation method that aligns with our core research objective of establishing a domain-specific analytical framework.

In this study, the Doc2Vec model was trained on preprocessed review data. Each review was converted into a TaggedDocument object after tokenization and stopword filtering. The model was initialized with a vector size of 10, learning rate $\alpha = 0.025$, and minimum word frequency min_count = 1. The model was trained for 100 iterations to ensure convergence and stability. Through careful hyperparameter tuning and iterative optimization, high-quality document embeddings were obtained.

3.3.2. Topic Modeling

To uncover users' latent interests and core concerns, this study employed Latent Dirichlet Allocation (LDA) for topic modeling. LDA is a generative probabilistic model based on Bayesian inference that assumes each document is a mixture of multiple latent topics, and each topic is represented by a distribution over words. By analyzing the co-occurrence patterns of words across documents, LDA can uncover the hidden thematic structure of textual data.

Formally, for a given document d and word ω, the probability of observing ω under the LDA model can be expressed as

$$\mathrm{P}(\omega|\mathrm{d}) = \sum_{\mathrm{z}=1}^{\mathrm{T}} \mathrm{P}(\omega|\mathrm{z}) \cdot P(\mathrm{z}|\mathrm{d}) \quad (1)$$

where T denotes the total number of latent topics, $p(z|d)$ is the topic proportion of topic z in document d, and $p(\omega|z)$ is the probability of word ω under topic z.

To determine the topic structure, we first estimated a 10-topic LDA model for each product category. The 10-topic setting was used at the modeling stage to retain sufficient thematic granularity and avoid prematurely merging relatively small but still interpretable themes. After model estimation, each review was assigned to its dominant topic according to the highest posterior topic probability, and the topics were then ranked by the number of supporting documents.

For reporting purposes, the six dominant topics with the largest supporting-document counts were retained (Table 2) because they captured the main discussion structure of the corpus while avoiding over-fragmented and low-support topics. The LDA model was estimated with num_topics = 10, passes = 20, and iterations = 400. During the model tuning, coherence and perplexity were used as diagnostic criteria, and the interpretability of the retained topics was further checked through keyword inspection and pyLDAvis-based visualization of the inter-topic distance and term relevance.

Table 2. Clustering results of online review topics.

Audiovisual Products	Topic Ranking	Number of Documents	Subject Headings
Skyworth TVs	1	9164	Speed, Screen, Appearance, Size, Sound Effects, Shape, Grand, Clear, Beautiful, Smooth
	2	6338	Quality, Price, Purchase, Packaging, Speed, Affordable, Cost-Effectiveness, Service, Cheap, Service
	3	5894	Installation, Technician, Delivery, Service, Professional, Patience, Service Attitude, Door-to-Door Delivery, Distribution
	4	5011	Clear, Picture Quality, Image, Sound Quality, Sound, Cost-Effectiveness, Operation, Smooth, Color, Effect
	5	3105	Voice, Function, Intelligent, Effect, Display, Screen Casting, Eye Protection, Experience, Color, Mode
	6	1722	Brand, Home, Activity, Quality, Cost-Effectiveness, Price, Worthwhile, Trade-in, Parents, Elderly
Xiaomi TVs	1	7144	Cost-Effectiveness, Clear, Price, Quality, Image, Picture Quality, Price Reduction, Brand, Affordable, Activity
	2	7141	Screen, Speed, Appearance, Operation, Size, Sound Effects, Effect, Size, Shape, Clear
	3	3976	Installation, Technician, Service, Professional, Patience, Speed, Responsible, Distribution, Experience, Enthusiastic
	4	1892	Elderly, Advertisement, Voice, Startup, Home, Xiao Ai, Operation, Screen Casting, Intelligent, Speaker
	5	1420	All-in-One, Delivery and Installation, Service, Ready to Install, Trade-in, Service Attitude, Speed, New Device, Gift, Attentive
	6	1338	Customer Service, Price Protection, Quality, Attitude, Baby, Installation Fee, Packaging, Professional, Accessories, Genuine
Xiaomi projectors	1	3591	Clear, Effect, Image, Cost-Effectiveness, Picture Quality, Satisfied, Smooth, Experience, Sound, Screen Casting
	2	1324	Operation, Brightness, Appearance, Sound Quality, Color, Shape, Clear, Compact, Effect, Exquisite
	3	1061	Price, Quality, Brand, After-Sales, Cost-Effectiveness, Affordable, Activity, Good Quality and Low Price, Quality, Home
	4	551	Simple, Price, Experience, Screen, Screen Casting, Mi Home, Image, Installation, Cost-Effectiveness, Design
	5	512	Customer Service, Quality, Logistics, Speed, Service, Packaging, Delivery, Seller, Patience, Cheap
	6	471	Automatic, Focus, Image, Trapezoidal, Clear, Brightness, Screen, Focus, Intelligent, System

3.3.3. Sentiment Analysis

The final analytical stage involved supervised sentiment classification to quantify user attitudes and satisfaction levels. Three representative machine learning algorithms

were compared: Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF). The performance of these models was evaluated, and the best-performing one was selected for further analysis.

LR is a linear classification model commonly used for binary sentiment tasks. It employs a sigmoid transformation to map linear outputs onto probabilities within the [0, 1] interval, indicating the likelihood of a review expressing positive sentiment. LR provides computational efficiency and probabilistic interpretability but may underperform on nonlinear decision boundaries.

SVM, a robust supervised learning method, was used for linear and nonlinear classification. SVM identifies the optimal separating hyperplane that maximizes the margin between classes. To handle nonlinear patterns, kernel functions were incorporated. In this study, the Radial Basis Function (RBF) kernel was selected, with optimized parameters $C = 100$ and $\gamma = 0.1$, based on empirical validation.

RF, an ensemble learning algorithm, constructs multiple decision trees and aggregates their predictions through majority voting to improve classification stability and accuracy. Randomness is introduced through feature and sample selection, enhancing model generalization. In this study, the RF configuration included 287 estimators, the minimum samples split was set to 5, the minimum leaf samples was 1, max_features was "sqrt", and there was no restriction on maximum depth.

Hyperparameter tuning plays a crucial role in optimizing machine learning model performance. To optimize the model performance, GridSearchCV was employed for hyperparameter tuning across all selected machine learning algorithms. This systematic approach allowed for the identification of optimal parameter configurations and improved the reliability of the sentiment classification outcomes.

Regarding model performance evaluation, we selected a comprehensive set of metrics to ensure a robust assessment of the classifiers. Sentiment distribution in e-commerce reviews is often characterized by a significant positivity bias; for instance, on platforms like JD.com, positive reviews typically constitute the vast majority of the data due to social norms and system-default praise mechanisms. In such imbalanced settings, Accuracy can be a misleading indicator, as a model may achieve a deceptively high score by simply favoring the majority class. To mitigate this risk, we expanded our evaluative framework beyond traditional accuracy, precision, recall, F1-score, and AUC by incorporating balanced accuracy and macro-F1. Balanced Accuracy provides an equitable measurement by weighting the sensitivity and specificity of both classes equally, while Macro-F1 calculates the unweighted mean of class-specific F1-scores. This multi-dimensional metrics ensures that the model's ability to identify minority-class (negative) sentiments, which often contain critical information for product improvement, is accurately captured.

4. Results

4.1. User Requirements Analysis Based on the LDA Model

In the digital era, consumers frequently share their experiences and evaluations of products on e-commerce platforms and social media after purchase and use. These online reviews not only contain direct assessments of product performance but also implicitly reflect users' actual requirements and expectations. To gain a deeper understanding of user requirements, firms must extract valuable insights from large volumes of unstructured review data. In this study, the Latent Dirichlet Allocation (LDA) model was applied to identify latent topics within user reviews of audiovisual products. The identified topics were then analyzed to uncover underlying user requirements and behavioral patterns, thereby providing empirical evidence to guide product improvement and marketing strategies.

4.1.1. Analysis of High-Frequency Words in Online Reviews

User-generated reviews play a crucial role in modern e-commerce ecosystems. They serve as an essential medium for consumers to express satisfaction or dissatisfaction and as a valuable source of market intelligence for firms seeking to enhance product quality and customer experience. To uncover consumers' focal points and requirements characteristics, this study conducted a high-frequency word analysis on the user reviews of three representative products: Skyworth TVs, Xiaomi TVs, and Xiaomi projectors. The results of this analysis provided an essential foundation for the subsequent topic modeling and sentiment analysis.

To further visualize user perceptions and experiential emphases, Figure 2 presents word clouds depicting the most frequently mentioned terms across the three product categories. For Skyworth TVs, the word cloud highlights user attention toward operating speed, design, installation service, screen and audio quality, and size. Users generally praised the display clarity, satisfactory sound performance, affordable pricing, and responsive service. The Xiaomi TVs word cloud emphasizes users' positive perceptions of processing speed, visual design, screen sound quality, installation service, and price-performance ratio, revealing its strong reputation and engagement among consumers. Meanwhile, the Xiaomi projectors word cloud demonstrates the users' focus on audiovisual effects, operational simplicity, sound quality, auto-focus functionality, and value for money, which collectively shape overall user satisfaction and influence purchase intentions.

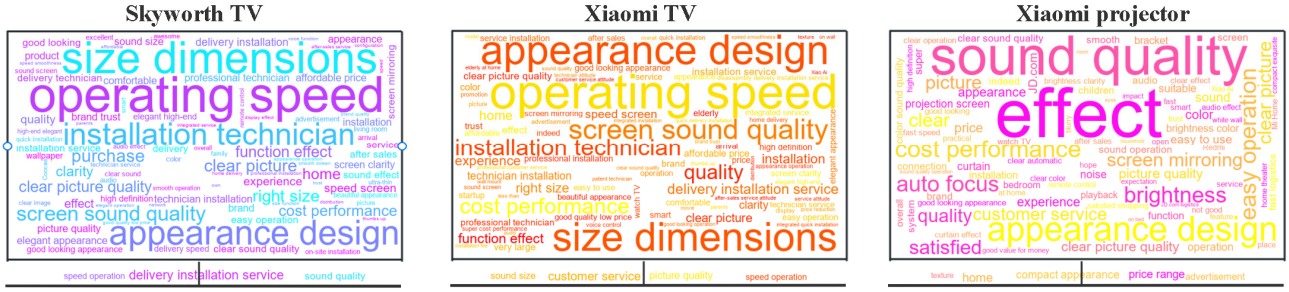


Figure 2. Word clouds of high-frequency review terms for Skyworth TVs, Xiaomi TVs, and Xiaomi projectors after Chinese word segmentation and English translation. The size of each word corresponds to its frequency in user comments.

These results indicate that although users of traditional and emerging brands share common concerns (e.g., display quality, ease of use), their distinct attention patterns reflect differing product expectations and experience priorities.

4.1.2. LDA Topic Modeling Results

The Latent Dirichlet Allocation (LDA) model was applied to the online review datasets of Skyworth TVs, Xiaomi TVs, and Xiaomi projectors to identify latent thematic structures. For each product category, an initial 10-topic solution was estimated. To improve readability and focus on the main discussion patterns, Table 2 reports the six dominant topics regarding supporting-document counts, together with the top 10 high-probability keywords for each topic.

Table 2 shows that the topic clustering results clearly reveal the major focal points and discussion patterns among users for each product type. Each identified topic reflects distinct aspects of user attention, ranging from product performance and installation services to customer experience and price–value perception. These thematic distinctions capture the multidimensional nature of user evaluations and provide a valuable foundation for further analysis of user requirements and product improvement strategies. For international presentation, the terms shown in Figures 2 and 3 were translated from Chinese into English.

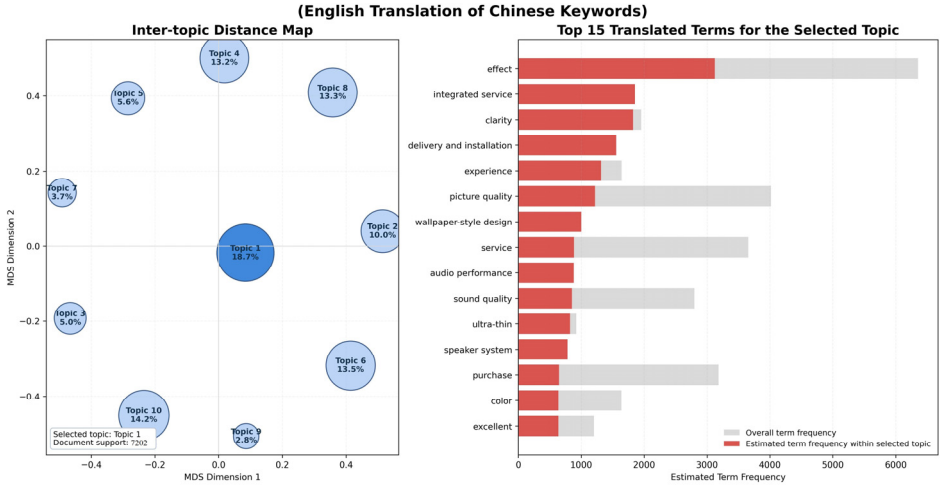


Figure 3. LDA topic modelling results for Skyworth TV reviews. The left panel illustrates the prevalence and inter-topic distances of various topics, while the right panel displays the representative topic keywords extracted from Skyworth TV reviews.

To further assess topic interpretability, Python 3.11.7's pyLDAvis module was used together with Gensim 4.3.0 for interactive visualization. Figure 3 shows that the visualization presents topic prevalence, inter-topic distance, and keyword relevance for the extracted topics. This step was used to inspect whether the dominant topics were reasonably separated in semantic space and whether their most relevant keywords formed coherent semantic clusters. Together with the supporting document coverage reported above, these checks provide additional evidence that the retained topics adequately reflect the main discussion structure of the review corpus.

Overall, the topic identification results reveal that user discussions center on functional and experiential aspects. Functional topics predominantly emphasize installation efficiency, display clarity, and operational reliability, whereas experiential topics focus on brand perception, service quality, and overall satisfaction. These insights provide a foundation for linking textual themes to user sentiment, supporting subsequent analyses of satisfaction patterns and brand differentiation in the audiovisual product market.

4.1.3. User Requirements Analysis

Based on the online review data for Skyworth TVs, Xiaomi TVs, and Xiaomi projectors, and informed by prior studies on review-based requirement extraction and user-requirement analysis [16,18,32,33], this study synthesized the LDA results to identify the common user-requirement areas shared across the three product types. Figure 4 shows that these requirements can be broadly grouped into five major areas.

First, audiovisual experience constitutes the most salient user concern. Consumers frequently comment on clarity, picture quality, color reproduction, and sound effects, reflecting their expectations for high-resolution displays and vivid, lifelike color performance.

Second, users emphasize cost performance, often expressing sensitivity to the price–value ratio. Terms such as cost-effectiveness and price discount appear frequently, indicating consumers' expectations of fair pricing relative to functionality.

Third, service quality is another critical factor. Users place importance on pre-sale consultation, installation service, and after-sales support, expecting timely, professional, and responsible service experiences.

Fourth, aesthetic design is a recurring theme, as consumers pay attention to appearance, size, and design harmony with home environments. Product aesthetics are perceived as a means of enhancing household decor and emotional satisfaction.



Figure 4. Five common user-requirement areas extracted from online reviews: audiovisual experience, cost performance, service quality, design aesthetics, and intelligent operation.

Finally, intelligent operation has emerged as a distinctive requirement in the context of smart home adoption. Users increasingly expect products to feature voice control, screen casting, and other intelligent functionalities that enhance convenience and interactivity.

These five areas collectively reflect consumers' comprehensive expectations toward functional performance and service experience. The findings provide practical insights for manufacturers and marketers to optimize the product design, marketing strategies, and after-sales service models in alignment with user requirements.

Drawing on the LDA results and the common requirements framework, the three focal products exhibit distinctive user-requirement profiles (Figure 5). Skyworth and Xiaomi, as traditional and well-established brands, benefit from a high level of brand trust and long-term market reputation. Users frequently mention keywords such as brand, quality, and durability, indicating strong confidence in product reliability and longevity. Additionally, Skyworth and Xiaomi are perceived as particularly suitable for family environments, with notable appeal among elderly user groups, reflecting the brand's strength in household adaptability and product stability.

In contrast, Xiaomi TVs stand out for their integration within the smart ecosystem. Beyond audiovisual quality and cost efficiency, users value Xiaomi's seamless connectivity with other smart home devices. Keywords such as Xiao Ai (AI assistant) and Mi Home highlight consumer interest in intelligent home control and ecosystem integration. Furthermore, features such as a gaming mode and motion detection enhance the interactivity and entertainment experience, demonstrating the brand's focus on technological innovation and experiential engagement.

Xiaomi projectors, meanwhile, attract users who prioritize portability and automation. Frequent mentions of the compact design, auto-focus, and keystone correction underscore the importance of mobility and ease of use. The minimalist aesthetic and lightweight structure resonate strongly with modern consumers seeking flexibility, efficiency, and convenience. These features not only enhance usability but also align with users' expectations of technologically sophisticated and user-friendly products.

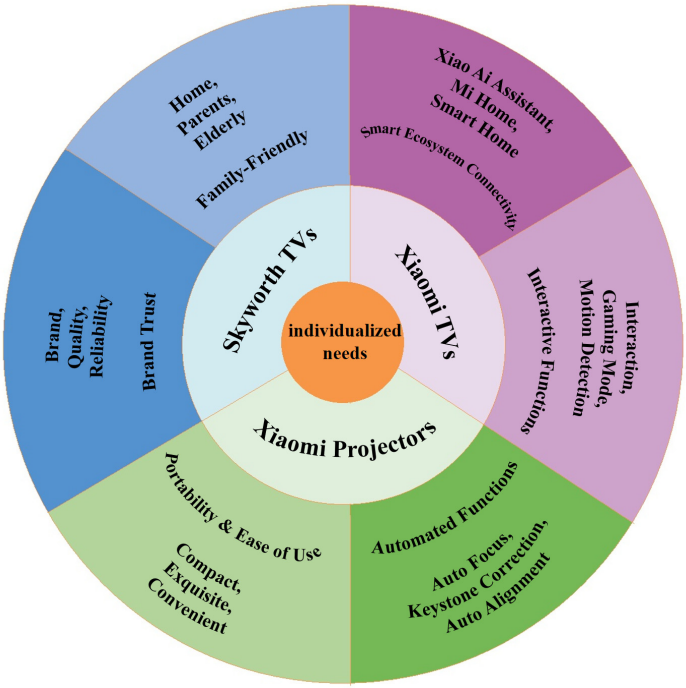


Figure 5. Brand- and product-specific user requirement profiles.

Although Skyworth TVs, Xiaomi TVs, and Xiaomi projectors share common user concerns such as audiovisual quality, cost performance, service experience, design aesthetics, and intelligent functionality, each brand differentiates itself through distinct functional emphases and experiential positioning. Understanding these distinctions allows firms to refine market segmentation, product development, and customer service strategies, enabling them to better address the requirements of diverse user groups and strengthen their competitive advantage.

A deeper comparison between traditional and emerging brands, as well as between TVs and projectors, reveals further nuances in consumer preferences and requirements dynamics.

In the comparison between traditional and emerging brands, distinct differences in user expectations were observed. Traditional brands, such as Skyworth, rely on brand credibility, stability, and reliability, emphasizing product durability and family integration. Users of traditional TVs tend to focus on consistent quality, long-term usage experience, and dependable after-sales service, preferring products validated by market experience and consumer trust.

Conversely, emerging brands such as Xiaomi emphasize technological innovation and intelligent functionality. These brands integrate the latest advancements in smart voice control, Internet connectivity, and interoperability with smart home systems, appealing to users seeking modern, connected living experiences. Consumers of new-generation Xiaomi TVs expect not only basic performance but also advanced smart features that enhance convenience and entertainment value.

In the comparison between TVs and projectors, differences are primarily driven by usage scenarios and functional expectations. Xiaomi TVs remain the mainstream choice for home entertainment, offering a comprehensive large-screen and immersive audiovisual experience. Users emphasize display quality, audio effects, and smart connectivity as key evaluation dimensions. Projectors, on the other hand, are valued for their portability and multifunctionality, catering to home, business, and educational applications. Users of projectors pay closer attention to mobility, image clarity, and multimedia adaptability, favoring devices that can be flexibly deployed across different environments.

These distinctions reflect how consumer preferences vary across product types, use contexts, and technological adoption stages. The findings underscore that while Xiaomi dominates in household integration and visual fidelity, projectors appeal to consumers seeking versatility and mobility. Recognizing these differentiated user requirements enables firms to develop targeted product strategies, enhance user satisfaction, and foster personalized consumption experiences.

4.2. Sentiment Classification Based on Machine Learning Models

Although online reviews on platforms such as JD.com include user-assigned star ratings, these ratings provide only a summary-level measure of satisfaction and fail to capture the granularity of emotional expression embedded in textual comments. For instance, a five-star rating typically represents positive feedback, while a one-star rating indicates dissatisfaction. However, intermediate ratings (2 to 4 stars) may contain mixed or context-dependent sentiments.

Unlike numerical ratings, sentiment analysis extracts emotion directly from the text, revealing the polarity and strength of opinions expressed toward specific product attributes. Therefore, sentiment classification was implemented to achieve a nuanced understanding of users' emotional tendencies and satisfaction levels.

4.2.1. Model Training and Evaluation

Three supervised classification models, including Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF), were trained and optimized to identify the best-performing model for sentiment classification. Because the labeled dataset was highly imbalanced, model performance was evaluated using both conventional and imbalance-aware metrics, including accuracy, precision, recall, F1-score, AUC, balanced accuracy, and macro-F1. These evaluation metrics ensured a balanced assessment of each model's predictive power.

As shown in Table 3, the SVM model demonstrated superior overall performance compared to other classifiers. And the tuned SVM achieved the most robust balance across multiple key metrics, specifically attaining an accuracy of 0.9370, an F1-score of 0.9637, and an AUC of 0.9590. Notably, its high balanced accuracy (0.8484) and macro-F1 (0.8627) underscore its effectiveness in handling the skewed sentiment distribution typical of e-commerce data. By providing a more reliable equilibrium between precision and recall than its counterparts, the Doc2Vec + SVM framework was retained as the optimal pipeline for downstream sentiment classification.

Table 3. Model evaluation.

Model	Parameter	Accuracy	Precision	Recall	F1-Score	AUC Value	Balanced Accuracy	Macro-F1
LR model	Default parameters	0.896	0.9044	0.9831	0.9421	0.6690	0.6690	0.7142
	After adjusting the parameters	0.9314	0.9474	0.9745	0.9607	0.8192	0.8192	0.8447
SVM model	Default parameters	0.8975	0.9043	**0.9853**	0.943	0.8511	0.6687	0.7155
	After adjusting the parameters	**0.937**	**0.9565**	0.971	**0.9637**	**0.959**	**0.8484**	**0.8627**
RF model	Default parameters	0.8904	0.9071	0.9723	0.9386	0.8702	0.6770	0.7150
	After adjusting the parameters	0.9288	0.9424	0.9771	0.9594	0.9476	0.8030	0.8348

Note: Bold indicates the optimal value.

4.2.2. Sentiment Analysis of the Three Product Categories

The optimized SVM model was subsequently employed to classify user reviews of Skyworth TVs, Xiaomi TVs, and Xiaomi projectors into positive and negative sentiment categories. Descriptive statistics of sentiment polarity across the three product categories are summarized in Table 4. Substantial variation in the proportion of positive reviews

was observed, with Skyworth TVs attaining the highest positive review rate (92.87%, 95% Wilson CI [92.59%, 93.14%]), followed by Xiaomi TVs (85.87%, 95% Wilson CI [85.43%, 86.29%]) and Xiaomi projectors (72.82%, 95% Wilson CI [71.85%, 73.77%]).

Table 4. Positive and negative review volumes of the three product categories.

	Skyworth TVs	Xiaomi TVs	Xiaomi Projectors
Number of positive reviews	31,153	21,682	6004
Number of negative reviews	2393	3569	2241
Ratio	13.02:1	6.08:1	2.68:1
Positive review rate (%)	92.87%	85.87%	72.82%
95% Wilson CI	[92.59–93.14%]	[85.43–86.29%]	[71.85–73.77%]

To formally evaluate the association between the product category and review sentiment, a Pearson chi-square test of independence was conducted. The results revealed a statistically significant association, indicating that the distribution of positive and negative reviews was not homogeneous across the three categories, $\chi^2(2, N = 67{,}042) = 2612.74$, $p < 0.001$. To identify the specific loci of these differences, pairwise comparisons of positive review proportions were performed using two-proportion z-tests with Bonferroni correction to maintain the family-wise error rate. Statistically significant differences in the proportion of positive reviews emerged for all three pairwise comparisons: Skyworth TVs versus Xiaomi TVs ($z = 27.84$, $p < 0.001$), Skyworth TVs versus Xiaomi projectors ($z = 51.94$, $p < 0.001$), and Xiaomi TVs versus Xiaomi projectors ($z = 27.16$, $p < 0.001$). Collectively, these inferential findings offer robust empirical support for the proposition that user satisfaction varies systematically as a function of the product category within the analyzed corpus.

Regarding practical significance, Skyworth TVs exhibited the most favorable sentiment profile, with a positive-to-negative review ratio of 13.02:1, which is indicative of a high degree of consumer endorsement and brand loyalty. Xiaomi TVs followed, with a ratio of 6.08:1, suggesting a predominantly favorable user perception. Conversely, Xiaomi projectors presented a markedly less favorable sentiment distribution, yielding a positive-to-negative ratio of 2.68:1. This pattern underscores notable divergences in consumer sentiment and points to potential avenues for improvement in product reliability, technological refinement, and after-sales service for the projector category.

4.2.3. Sentiment Analysis of Common Requirements

To further investigate user perceptions, sentiment analysis was conducted at the level of specific user requirements derived from the LDA topic-modeling results. The five identified user-requirement areas, namely, audiovisual experience, cost performance, service quality, design aesthetics, and intelligent operation, were used to categorize review content, and the positive sentiment rate was calculated for each area.

Table 5 shows that Skyworth TVs achieved the highest satisfaction across nearly all requirement areas. The audiovisual experience area received a positive sentiment rate of 97.89%, reflecting Skyworth's strength regarding the core technical performance and design quality. High satisfaction was also observed regarding the cost performance, service quality, design aesthetics, and intelligent operation. Nevertheless, minor issues such as hardware failures and delayed after-sales responses were mentioned in some reviews, suggesting areas for continued improvement.

Xiaomi TVs were highly praised for their cost-effectiveness and intelligent system design, with positive sentiment rates of 94.05% and 93.99%, respectively. The design aesthetics category received a particularly high positive sentiment rate of 99.03%, highlighting users' appreciation of Xiaomi's modern visual appeal and user-friendly interface. However, some users noted remote control difficulties and minor functional inconsistencies, indicating

potential enhancements for future product iterations. Xiaomi's proactive feedback-driven approach and frequent system updates have helped maintain strong consumer engagement and satisfaction.

Table 5. Positive sentiment rates across five common user-requirement areas.

User Requirement Area	Audiovisual Products	Total Reviews	Positive Reviews	Positive Sentiment Rate
Audiovisual experience	Skyworth TVs	15,967	15,630	97.89%
	Xiaomi TVs	8057	7805	96.87%
	Xiaomi projectors	4142	2825	68.20%
Cost performance	Skyworth TVs	6312	6141	97.29%
	Xiaomi TVs	5466	5141	94.05%
	Xiaomi projectors	1615	1184	73.31%
Service quality	Skyworth TVs	12,748	12,104	94.95%
	Xiaomi TVs	9759	8668	88.82%
	Xiaomi projectors	474	345	72.31%
Design aesthetics	Skyworth TVs	8776	8715	99.30%
	Xiaomi TVs	5354	5302	99.03%
	Xiaomi projectors	1582	1019	64.41%
Intelligent operation	Skyworth TVs	5489	5244	95.54%
	Xiaomi TVs	3462	3254	93.99%
	Xiaomi projectors	1184	578	48.82%

As an emerging category, Xiaomi projectors received lower satisfaction levels regarding the audiovisual experience (68.20%) and intelligent operation (48.82%) compared with the television products. Despite this, users valued the projector's portability and practical performance in outdoor or mobile scenarios, and many acknowledged its favorable cost–value ratio. Some critical reviews mentioned brightness limitations and noise issues, which are common challenges in projector technology. These insights suggest that further technological innovation and engineering refinement are required to improve the performance consistency.

In summary, the three product categories display distinctive strengths and weaknesses. Skyworth TVs excel in audiovisual quality but need improvement in after-sales service. Xiaomi TVs perform well in design and smart functionality but can benefit from enhanced service responsiveness. Xiaomi projectors, while competitive in portability and affordability, still lag behind in user satisfaction due to performance constraints. These findings provide actionable guidance for manufacturers seeking to optimize product design, feature development, and customer service strategies in the evolving audiovisual market.

5. Discussion

5.1. Summary of Findings

This study deepens the understanding of user requirements in the audiovisual product market through a large-scale analysis of online reviews for Skyworth TVs, Xiaomi TVs, and Xiaomi projectors. By integrating LDA topic modeling with a Doc2Vec + SVM sentiment classification framework, the study identified and compared the main user requirement areas discussed across the three product groups.

The findings reveal five primary user-requirement areas: audiovisual experience, cost performance, service quality, design aesthetics, and intelligent operation. Skyworth TVs, representing a traditional manufacturer, are especially recognized for audiovisual perfor-

mance and design quality, whereas Xiaomi TVs, representing an emerging Internet-oriented brand, are more strongly associated with cost-effectiveness and intelligent functionality. Xiaomi projectors receive positive feedback for portability but show lower satisfaction in core audiovisual performance and intelligent operation.

Sentiment analysis supports these patterns. The overall positive review rates were 92.87% for Skyworth TVs, 85.87% for Xiaomi TVs, and 72.82% for Xiaomi projectors. These differences were statistically significant across product categories ($\chi^2 = 2612.74$, df = 2, $p < 0.001$), indicating that the satisfaction patterns differ meaningfully across the three product groups within the present review corpus. However, these results should be interpreted as evidence from JD.com rather than as universally generalizable conclusions for all consumers or all online retail settings.

5.2. Practical Contributions

The contribution of this study is primarily empirical and contextual rather than algorithmic. Prior research has shown that online review data can support requirement extraction and sentiment analysis. Building on this literature, the present study shows that user requirements in the audiovisual market can be organized into five recurring areas: audiovisual experience, cost performance, service quality, design aesthetics, and intelligent operation.

A second contribution is that the study links requirement identification with sentiment evaluation within the same analytical setting. This makes it possible to examine not only what users discuss but also how positively or negatively they evaluate different user-requirement areas. The comparison across Skyworth TVs, Xiaomi TVs, and Xiaomi projectors further shows that requirement salience and satisfaction patterns vary across product positions and product categories. In this sense, the study extends prior review-mining research by providing domain-specific evidence on competitive differentiation in the audiovisual sector.

From a practical perspective, the results help firms translate large-scale review data into concrete improvement priorities. Skyworth should maintain its strengths in audiovisual performance while addressing after-sales service issues; Xiaomi TVs should continue refining intelligent functions and service responsiveness; and projector manufacturers should prioritize core performance issues such as brightness, noise, and operational stability.

5.3. Limitations and Future Research

Despite its contributions, this study is subject to several limitations that open avenues for future research.

First, all review data in this study were collected from a single e-commerce platform, namely, JD.com. Although JD.com provides a large volume of authentic user-generated content, platform-specific factors such as the user composition, review-generation mechanisms, product assortment, and interaction design may influence the salience of user requirements and the distribution of expressed sentiments. Therefore, the findings should be interpreted as evidence derived from the JD.com review context rather than as conclusions that can be directly generalized to the entire audiovisual product market. Future research could incorporate data from multiple e-commerce and social media platforms and further compare whether user requirements and satisfaction patterns remain stable across different platform environments.

Second, this study used Doc2Vec as an efficient and interpretable document representation method. This choice supports computational feasibility and transparent integration with classical classifiers, but it may not capture contextual semantics as fully

as transformer-based embeddings. Future studies could compare Doc2Vec with BERT, RoBERTa, or domain-adapted transformer models to examine whether richer contextual representations improve requirement identification and sentiment classification in audiovisual product reviews.

Finally, the categorization of user sentiments in this study involved a degree of researcher judgment, which may introduce subjectivity. Future work could explore automated classification and clustering algorithms to improve the objectivity, robustness, and reproducibility of the results.

Author Contributions: Methodology, C.L.; validation, M.C.; investigation, X.Z.; writing—original draft preparation, X.Z. and Z.H.; writing—review and editing, C.L.; visualization, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Nature Science Foundation of China (72401039), Natural Science Foundation of Hunan Province (2024JJ6069), and Humanities and Social Sciences Research Project of the Ministry of Education (24YJC630128).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be made available upon request.

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

References

1. Adner, R.; Puranam, P.; Zhu, F. What is different about digital strategy? From quantitative to qualitative change. *Strat. Sci.* **2019**, *4*, 253–261. [CrossRef]
2. Zhou, Z.; Zheng, L.; Qi, Y. Bilateral interaction between Mainland China and Hong Kong audiovisual products under the perspective of cross-cultural communication. *J. Mod. Soc. Sci.* **2025**, *2*, 82–93. [CrossRef]
3. Huang, L.; Jia, Y. Innovation and development of cultural and creative industries based on big data for industry 5.0. *Sci. Program.* **2022**, *2022*, 2490033. [CrossRef]
4. Nambisan, S.; Lyytinen, K.; Majchrzak, A.; Song, M. Digital innovation management: Reinventing innovation management research in a digital world. *MIS Q.* **2017**, *41*, 223–238. [CrossRef]
5. Meng, F.; Jia, Y. Data-Driven Prioritization of User Requirements in Health E-Commerce: An Explainable Machine Learning Study. *J. Theor. Appl. Electron. Commer. Res.* **2026**, *21*, 104. [CrossRef]
6. Davoodi, L.; Mezei, J.; Heikkilä, M. Aspect-based sentiment classification of user reviews to understand customer satisfaction of e-commerce platforms. *Electron. Commer. Res.* **2026**, *26*, 1417–1459. [CrossRef]
7. Harrison-Walker, L.J.; Jiang, Y. Suspicion of online product reviews as fake: Cues and consequences. *J. Bus. Res.* **2023**, *160*, 113780. [CrossRef]
8. Azer, J.; Anker, T.; Taheri, B.; Tinsley, R. Consumer-Driven racial stigmatization: The moderating role of race in online consumer-to-consumer reviews. *J. Bus. Res.* **2023**, *157*, 113567. [CrossRef]
9. Cho, H.S.; Sosa, M.E.; Hasija, S. Reading between the stars: Understanding the effects of online customer reviews on product demand. *Manuf. Serv. Oper. Manag.* **2022**, *24*, 1977–1996. [CrossRef]
10. García-Zamora, D.; Dutta, B.; Jin, L.; Chen, Z.-S.; Martínez, L. A data-driven large-scale group decision-making framework for managing ratings and text reviews. *Expert Syst. Appl.* **2025**, *263*, 125726. [CrossRef]
11. Li, Q.; Chen, H.; Long, R.; Huang, Z.; Yang, S.; Sun, Q.; Sun, Y.; Ye, X. Research on data-driven group consensus decision-making of green methanol vehicle evaluation based on BERTopic text mining. *Sustain. Energy Technol. Assess.* **2025**, *80*, 104362. [CrossRef]
12. Oliseenko, V.D.; Eirich, M.; Tulupyev, A.L.; Tulupyeva, T.V. BERT and ELMo in task of classifying social media users posts. In *International Conference on Intelligent Information Technologies for Industry*; Springer International Publishing: Cham, Switzerland, 2022; pp. 475–486.
13. Vanhala, M.; Lu, C.; Peltonen, J.; Sundqvist, S.; Nummenmaa, J.; Järvelin, K. The usage of large data sets in online consumer behaviour: A bibliometric and computational text-mining–driven analysis of previous research. *J. Bus. Res.* **2020**, *106*, 46–59. [CrossRef]
14. Cai, M.; Tan, Y.; Ge, B.; Dou, Y.; Huang, G.; Du, Y. PURA: A product-and-user oriented approach for requirement analysis from online reviews. *IEEE Syst. J.* **2021**, *16*, 566–577. [CrossRef]

15. Scheurwegs, E.; Luyckx, K.; Luyten, L.; Goethals, B.; Daelemans, W. Assigning clinical codes with data-driven concept representation on Dutch clinical free text. *J. Biomed. Inform.* **2017**, *69*, 118–127. [CrossRef]
16. Cai, M.; Yang, W.; Du, Y.; Tan, Y.; Lu, X. Automatic requirements elicitation from user-generated content: A review of data, methods, and representations. *Eng. Appl. Artif. Intell.* **2025**, *156*, 111110. [CrossRef]
17. Cong, Y.; Yu, S.; Chu, J.; Su, Z.; Huang, Y.; Li, F. A small sample data-driven method: User needs elicitation from online reviews in new product iteration. *Adv. Eng. Inform.* **2023**, *56*, 101953. [CrossRef]
18. Zhang, Y.; Guo, W.; Chang, Z.; Ma, J.; Fu, Z.; Wang, L.; Shao, H. User requirement modeling and evolutionary analysis based on review data: Supporting the design upgrade of product attributes. *Adv. Eng. Inform.* **2024**, *62*, 102861. [CrossRef]
19. Lilingpadang, Y.C.; Haddara, M.; Langseth, M. Data-Driven Fashion Trend Identification: Utilizing Text Mining for New Product Development Process. *Procedia Comput. Sci.* **2025**, *263*, 258–267. [CrossRef]
20. Shi, L.; Jia, Z.; Liu, R. Expert-Level Financial Sentiment Analysis via Internal and External Financial Knowledge Fusion. *Inf. Fusion* **2025**, *127*, 103884. [CrossRef]
21. Cui, G.; Chung, Y.; Peng, L.; Zheng, W. The importance of being earnest: Mandatory vs. voluntary disclosure of incentives for online product reviews. *J. Bus. Res.* **2022**, *141*, 633–645. [CrossRef]
22. Moloi, M.; Quaye, E.S.; Saini, Y.K. Evaluating key antecedents and consequences of the perceived helpfulness of online consumer reviews: A South African study. *Electron. Commer. Res. Appl.* **2022**, *54*, 101172. [CrossRef]
23. Sboev, A.; Naumov, A.; Rybka, R. Data-driven model for emotion detection in Russian texts. *Procedia Comput. Sci.* **2021**, *190*, 637–642. [CrossRef]
24. Jin, E.; Oh, J. The role of emotion in interactivity effects: Positive emotion enhances attitudes, negative emotion helps information processing. *Behav. Inf. Technol.* **2022**, *41*, 3487–3505. [CrossRef]
25. Chiu, M.C.; Lin, K.Z. Utilizing text mining and Kansei Engineering to support data-driven design automation at conceptual design stage. *Adv. Eng. Inform.* **2018**, *38*, 826–839. [CrossRef]
26. Braoudaki, A.; Kanellou, E.; Kozanitis, C.; Fatourou, P. Hybrid data driven and rule based sentiment analysis on Greek text. *Procedia Comput. Sci.* **2020**, *178*, 234–243. [CrossRef]
27. Chen, L.; Mankad, S. A structural topic and sentiment-discourse model for text analysis. *Manag. Sci.* **2025**, *71*, 5767–5787. [CrossRef]
28. Li, H.; Liu, H.; Shin, H.H.; Ji, H. Impacts of user-generated images in online reviews on customer engagement: A panel data analysis. *Tour. Manag.* **2023**, *101*, 104855. [CrossRef]
29. Bauman, K.; Tuzhilin, A. Know thy context: Parsing contextual information from user reviews for recommendation purposes. *Inf. Syst. Res.* **2022**, *33*, 179–202. [CrossRef]
30. Cao, Z.; Zhu, Y.; Li, G.; Qiu, L. Consequences of information feed integration on user engagement and contribution: A natural experiment in an online knowledge-sharing community. *Inf. Syst. Res.* **2024**, *35*, 1114–1136. [CrossRef]
31. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781. [CrossRef]
32. Bai, S.; Shi, S.; Han, C.; Yang, M.; Gupta, B.B.; Arya, V. Prioritizing user requirements for digital products using explainable artificial intelligence: A data-driven analysis on video conferencing apps. *Futur. Gener. Comput. Syst.* **2024**, *158*, 167–182. [CrossRef]
33. Zhou, T.; Chen, Z.; Cao, Y.; Miao, R.; Ming, X. An integrated framework of user experience-oriented smart service requirement analysis for smart product service system development. *Adv. Eng. Inform.* **2022**, *51*, 101458. [CrossRef]

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.