Pinnacle Academic Press
Proceedings Series

Vol. 2 2025

Article **Open Access**

Research on the Demand Hierarchy of E-Commerce Products Based on Text Mining and the IPA-KANO Model

Anyi Chen [1,*]

[1] Walmart E-commerce Search Team, Wal-Mart Associates Inc., Bentonville, Arkansas, 72716, USA
* Correspondence: Anyi Chen, Walmart E-commerce Search Team, Wal-Mart Associates Inc., Bentonville, Arkansas, 72716, USA

Received: 16 April 2025
Revised: 25 April 2025
Accepted: 10 May 2025
Published: 14 June 2025

Abstract: In the current era of rapid growth in e-commerce, understanding consumer demand is crucial and also the core of enhancing market competitiveness. This study integrates text mining technology and IPA-KANO model to create a comprehensive system for analyzing the demand hierarchy of e-commerce products. This study uses text mining to collect, process, and analyze consumer evaluation data, uncovering the basic needs of consumers. By integrating the IPA matrix and KANO model to analyze user needs, basic needs, necessary needs, and attractive needs are distinguished. The text analysis results are combined with the IPA-KANO framework to construct a hierarchy of needs and response mechanisms, providing specific data support for product iteration. Research shows that this method helps to uncover users' deep-seated needs, assist e-commerce companies in customizing precise marketing plans and product upgrade paths, and strengthen the e-commerce platform's ability in market competition.

Keywords: e-commerce; text mining; IPA-KANO model; product demand hierarchy

1. Introduction

With the continuous development of information technology, consumer demands have become increasingly diversified and personalized. For e-commerce companies, how to systematically identify and classify consumer needs has become the key to product development and marketing. Text mining technology provides strong support for analyzing consumer evaluations and feedback due to its powerful analytical ability for unstructured data. In addition, using the IPA-KANO model for requirement classification and importance evaluation can help enterprises determine the types and criticality of requirements, allocate resources reasonably, and optimize product strategies. By conducting in-depth mining and research on consumer review texts, revealing the essential laws behind user needs, classifying and ranking the demand for e-commerce products, and dividing their importance, a sustainable demand feedback system is constructed to provide support for e-commerce companies to achieve innovation and improvement under changing market competition conditions.

2. Text Preprocessing Process

In the process of text mining, text preprocessing is crucial, as shown in Figure 1. This stage has a direct impact on the accuracy and effectiveness of subsequent data analysis. The core task of preprocessing is to transform unordered text information into an ordered data structure that is easy to analyze. As the primary step of preprocessing, preliminary

data cleaning is responsible for removing useless characters (such as specific symbols, HTML code) and commonly used invalid words (such as "zhi", "hu", etc.) from comments, reducing the interference of noise. Text segmentation is a key step in natural language processing. For Chinese comment content, word segmentation techniques can be used to implement segmentation operations and ensure the accuracy of industry terminology. Further steps such as part of speech tagging, frequency statistics, and sentiment analysis are taken to extract core information from comments. To enhance the efficiency of data processing, text information can be vectorized, such as using TF-IDF or Word2Vec techniques to convert text data into numerical feature vectors. The preprocessed data needs to be standardized for storage, providing a high-quality data foundation for subsequent IPA-KANO model analysis [1].

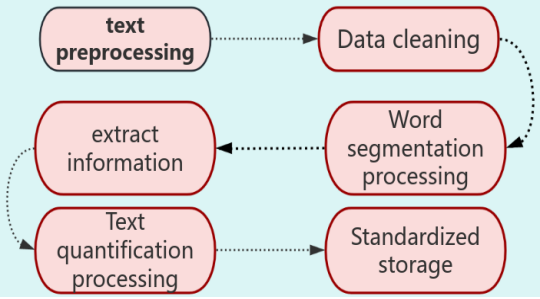

Figure 1. Data Processing Flow.

3. Construction of IPA-KANO Model

3.1. IPA Model Construction

IPA is an analytical tool specifically designed to explore consumer needs. Its core lies in the comparative analysis of the importance and actual performance of product or service characteristics to determine areas for improvement [2]. The purpose of this study is to use the IPA method to explore the main demand characteristics of e-commerce products and provide theoretical support for product optimization design decisions. Under the IPA framework, text mining techniques were used to conduct in-depth analysis of user comments on e-commerce products, extract key vocabulary and sentences, and construct attribute sets $\{A_1, A_2, \ldots, A_n\}$. Basic needs refer to the minimum expectations of consumers, functional needs refer to the clearly expressed needs of consumers, and charismatic needs refer to innovative functions or attributes that exceed consumer expectations and can greatly improve user satisfaction. The specific calculation formula is as follows:

$$I_i = \frac{\sum_{j=1}^{m} W_j \cdot S_{ij}}{m} \quad (1)$$

In formula (1), W_j represents the weight of the comment (evaluated based on user level and comment detail), S_{ij} represents the sentiment score of attribute A_i in the j-th comment, and m is the total number of comments. The performance level (P) is obtained by crawling the rating data of e-commerce platforms and normalizing it, and the formula is as follows:

$$P_i = \frac{\sum_{k=1}^{n} R_{ik}}{n} \quad (2)$$

In formula (2), R_{ik} represents the performance score of attribute A_i in the kth rating, and n represents the overall rating score.

3.2. Classification of Requirements Using KANO Model

The KANO model classifies user needs into five categories: basic needs, performance needs (functional needs), attractive needs, neutral needs, and reverse needs. Basic needs refer to the minimum expectations of consumers, functional needs refer to the clearly ex-

pressed needs of consumers, and Attractive needs refer to innovative functions or attributes that exceed consumer expectations and can greatly improve user satisfaction. The neutrality of customer demand will not significantly affect satisfaction levels, while reverse demand may trigger reverse evaluations. This study uses text mining techniques to identify emotional tendencies and extract specific vocabulary from consumer reviews. Based on the correlation between keywords and demand categories, the KANO model is applied to classify consumer demand. In terms of specific operations, the study integrates the importance of needs and the intensity of emotions, using quantitative methods to distinguish the types of needs [3]. In the specific application process of the model, keywords with high frequency of occurrence were selected, corresponding to five different demand categories, and the classification probabilities of these categories were calculated. The following is the KANO model requirement classification formula. Assuming the emotional weight S_i, importance score I_i, and frequency proportion F_i of the requirement features, the classification probability P_i^k is:

$$P_i^k = \frac{\omega_k \cdot (S_i \cdot I_i \cdot F_i)}{\sum_j^n (\omega_j \cdot (S_j \cdot I_j \cdot F_j))} \tag{3}$$

In formula (3), k represents the classification of requirements (basic, functional, attractive, neutral, reverse); ω_k refers to the weight factor for demand classification, which is configured based on user attributes and industry requirements; N refers to the total amount of demand attributes. By using formulas, it is possible to specifically determine the type of demand to which each attribute belongs, providing support for product improvement and service enhancement.

3.3. Integrating IPA Analysis Results with KAN0 Model

To further explore the primary and secondary importance levels of e-commerce product demand, a strategy of integrating IPA method with KANO model is adopted to comprehensively evaluate the impact of demand attributes on consumer satisfaction. Establish a formula for assessing the overall priority of requirements, and merge the User Satisfaction Index (CS) under the KANO category with the importance rating and performance data of requirements in IPA.

$$P_i = I_i \times (CS_i^+) - P_i \times (CS_i^-) \tag{4}$$

In formula (4), P_i represents the priority of demand, while I_i represents the importance of demand; P_i demand performance value; CS_i^+ positive satisfaction coefficient (weight of expectancy and charm in KANO model); CS_i^- negative Satisfaction Coefficient (the weight of essential and inverse types in the KANO model). The following is the integrated analysis result of the main requirements of an e-commerce platform (Table 1):

Table 1. Calculation of Requirement Integration Priority.

Requirement number	Requirement	I_i	P_i	CS_i^+	CS_i^-	Comprehensive priority P_i
N1	Product search function	0.85	0.95	0.60	0.30	0.485
N2	Personalized recommendation function	0.75	0.80	0.70	0.20	0.465
N3	Voice shopping assistant	0.65	0.70	0.80	0.10	0.442
N4	Social sharing shopping function	0.40	0.50	0.30	0.10	0.115
N5	Optimization of complex registration process	0.30	0.40	0.10	0.60	-0.126

After integration, analysis shows that product search and personalized recommendation functions occupy a leading position in priority ranking, and prioritizing the improvement of these functions will significantly improve customer satisfaction. Although the priority of voice shopping assistants has not yet been elevated to a high level, as a highly attractive demand, its improvement in user experience cannot be ignored. The social shopping sharing function is relatively secondary in importance and performance ratings, with priority given to it. For the complex user registration process, certain changes

may reduce user satisfaction due to their significant negative impact, so caution is recommended when addressing this issue. It is recommended to be extra cautious when dealing with this issue. The quantitative evaluation method combining IPA and KANO models provides a scientific basis for e-commerce platforms to develop targeted demand optimization strategies [4].

4. Research on the Hierarchy of Demand for E-Commerce Products

4.1. Analysis of Text Mining Results for User Comment Data

User comment data is an important source for e-commerce platforms to understand user needs, and text mining techniques can be used to filter key information from numerous comments [5]. The article selected user feedback from a large e-commerce platform as the research sample, and analyzed these feedbacks using natural language processing methods to extract the demand characteristics and emotional tendencies that users are concerned about. The first step of the research is to clean and preliminarily process the original feedback, which involves removing redundant characters, punctuation, and stop words. Then, the feedback text is segmented into keywords using word segmentation technology, and the TF-IDF method is applied to identify frequently mentioned key words by users. By implementing emotion analysis methods, consumer feedback is classified into three emotional attitudes: positive, neutral, and negative, and the satisfaction level of customers at various levels of needs is quantitatively evaluated. To further analyze the hierarchical nature of user needs, this study constructed a needs classification model that combines the importance of needs and performance ratings, aiming to quantify the correlation between the weight of user comments and satisfaction. The frequency of appearance of feature words is represented by F_i, the intensity of emotional expression is represented by S_i, and the importance rating of needs is represented by I_i. The calculation of user demand weight Wi follows the following mathematical formula:

$$W_i = \alpha \cdot \frac{F_i}{\sum_{i=1}^{n} F_j} + \beta \cdot \frac{S_i}{max(S)} + \gamma \cdot \frac{I_i}{max(I)} \tag{5}$$

In formula (5), α, β, and γ represent the influence weights, and the sum is equal to 1. $Max(S)$ and $max(I)$ respectively refer to the highest scores of emotional intensity and attention, used to achieve standardized scoring. N refers to the number of keywords in the comment. This formula integrates three major elements: the frequency of comments, emotional sentiment, and level of attention, providing a theoretical basis for classifying demand levels and planning strategies.

4.2. Classification and Priority Analysis of Essential Requirements

Essential requirements are the most basic expectations of users for e-commerce products and services, which are a category within the KANO model and are usually implied in users' expressions. For e-commerce products, correctly identifying and categorizing requirements is crucial to effectively prioritize development efforts and improve customer satisfaction. Using text mining techniques, organize and study consumer evaluation data to extract multiple core demand elements such as product quality, after-sales support, and delivery speed. Based on the theoretical framework of the IPA model, using importance and performance levels as evaluation criteria, the requirements are divided into four major areas: "urgent optimization", "maintaining investment", "potential observation", and "non key". Research has found that in terms of logistics efficiency and after-sales service, although they have a high level of importance, their actual performance score is significantly lower than other needs, so they are classified as urgently needing optimization. The product performance and appearance design scores are relatively balanced and classified as maintaining investment category. As for the product price and additional features, there are significant differences in user opinions, which are classified as potential observation categories. In order to more accurately determine the priority order of require-

ments, this study developed a priority evaluation formula based on weights, which comprehensively considers the importance of requirements, performance ratings, and the weight ratio of user emotional tendencies. The formula for calculating the demand weight Wi, importance score I_i, performance score P_i, positive sentiment tendency ratio Ei+, and priority R_i is:

$$R_i = W_i \times (I_i + \alpha \cdot ({E_i}^+ - \beta \cdot (1 - P_i))) \quad (6)$$

In formula (6), α represents the moderating factor of positive emotions and regulates the weight of consumer satisfaction. β represents a punitive factor for underperformance, revealing the impact of failure to meet demand standards. W_i is determined based on the weights of different dimensions using the Analytic Hierarchy Process. The formula integrates actual performance data and emotional response of demand presentation, providing theoretical support for resource priority allocation [6].

4.3. Charm Based Requirement Feature Extraction

In the KANO model, the need for charm is seen as an implicit need that has not been directly proposed by users but can improve their satisfaction. This type of demand is manifested in the value-added characteristics of the product or service, such as novel features or unique design elements. Exploring such demands has a significant impact on enhancing the market competitiveness of e-commerce products and improving user interaction experience. This study uses text mining methods to reveal hidden expressions in user comments, focusing on uncovering users' positive emotional tendencies towards content that exceeds expectations through sentiment analysis. The frequency inverse document frequency (TF-IDF) technique is used to identify key expressions, and high-value demand information is selected through the evaluation of emotional tendencies. For example, consumers have shown great recognition for innovative design of product appearance and customized services of intelligent recommendation systems, which reflects that these elements possess core demand attributes that attract consumers. This study developed a composite feature screening algorithm to evaluate the specific contribution of such core needs to consumer satisfaction. This algorithm comprehensively analyzes the frequency of requirement mentions, the intensity of emotions, and the mutual influence between factors, and develops a set of evaluation criteria for feature importance. The following is the specific mathematical expression of the criterion.

$$Q_i = \gamma \cdot \frac{F_i}{\sum_{j=1}^{n} F_j} + \delta \cdot \frac{S_i}{max(S)} + \eta \cdot R_i \cdot log(1 + F_i) \quad (7)$$

In formula (7), γ, δ, and η respectively represent the moderating factors of frequency of use, emotional influence, and correlation (with a total of 1). R_i is derived based on the co-occurrence probability matrix between demand attributes and core requirements, and through the application of the logarithmic function $log(1 + F_i)$, it reduces the excessive weight advantage of high-frequency demands. Formulas can accurately identify and rank those attractive requirements, providing scientific basis for selection in the design process.

4.4. Construction of Continuous Demand Monitoring and Feedback Mechanism

In the constantly changing market environment, it is particularly crucial for e-commerce platforms to maintain their competitive advantage by continuously tracking user needs. Consumer preferences and needs gradually change over time and market conditions, so it is extremely necessary to create a demand monitoring and feedback system that can respond in real time [7]. Based on text mining technology and the IPA-KANO model, a comprehensive cycle system for requirement monitoring and feedback has been designed. The system mainly includes data collection and pre-processing, real-time analysis and evaluation of requirements, and optimization and updating of feedback strategies. In the information gathering stage, a complete user information database was constructed by continuously tracking users' comments, ratings, and social media interactions.

The evaluation of demand evolution utilizes text analysis techniques to regularly track fluctuations in demand hierarchy and reclassify user demands using the IPA-KANO model. The optimization of feedback strategy quickly delivers the results of requirement evaluation to the product development and service departments, in order to make accurate responses to key and attractive requirements. This study conducted a simulation test to verify the actual effectiveness of this mechanism. By continuously monitoring user message data, real-time changes in the demand hierarchy were tracked. The following is a summary and analysis of experimental data.

Observing the data in Table 2, it can be seen that the frequency of use of basic requirements continues to decrease, while functional requirements show an upward trend. This indicates that users' attention to core functions is shifting towards additional functions that enhance the user experience. Although the proportion of attractive demand is not high, it has increased during the product optimization stage, reflecting that companies have successfully attracted consumer interest while meeting basic needs through innovative means. The user satisfaction has increased from the initial 85% to 92% after optimization, which confirms the significant role of feedback mechanisms in improving user satisfaction. This mechanism provides practical operational guidelines for e-commerce websites to continuously improve user experience under ever-changing market conditions.

Table 2. Analysis of Dynamic Changes in Demand Levels in Different Periods (Simulated Data).

Period	Proportion of basic demand frequency	Proportion of expressive demand frequency	Proportion of Charm Type Demand Frequency	Change in demand satisfaction
Initial	60%	30%	10%	85%
Mid term	55%	35%	10%	88%
Post production	50%	40%	10%	90%
Subsequent optimization period	45%	40%	10%	92%

5. Conclusion

By integrating text mining technology and improved IPA-KANO analysis method, a deep exploration was conducted on the user demand structure of e-commerce products, and a systematic mechanism for demand identification and classification was established. Research has found that text mining techniques can efficiently extract hidden demand points from user feedback. The IPA-KANO model evaluates based on the importance of demand and user satisfaction, providing a solid reference for strategic decision-making in enterprises. Especially in discovering and tracking unique and attractive demands, this study has opened up new avenues for product innovation and user experience improvement. With the expansion of e-commerce data scale and the diversified development of user demand, the framework of this study is expected to further develop, integrating big data and artificial intelligence technology to achieve real-time prediction and feedback of demand. This study brings innovative strategies that combine theory and practice to e-commerce enterprises, which have practicality and promotion potential in the ever-changing market context.

References

1. J. Lu, et al., "BSTC: A fake review detection model based on a pre-trained language model and convolutional neural network," *Electronics*, vol. 12, no. 10, p. 2165, 2023, doi: 10.3390/electronics12102165.
2. Q. Peng, C. Wang, and M. Goh, "Green financing strategies in a low-carbon e-commerce supply chain under service quality regulation," *Environ. Sci. Pollut. Res.*, vol. 30, no. 2, pp. 2575–2596, 2023, doi: 10.1007/s11356-022-22329-w.
3. M. Yan, et al., "An empirical investigation of the impact of influencer live-streaming ads in e-commerce platforms on consumers' buying impulse," *Internet Res.*, vol. 33, no. 4, pp. 1633–1663, 2023, doi: 10.1108/INTR-11-2020-0625.

4. T. Kojo and T. Yoshikawa, "Comparison of Location Potential for Customer Attraction of a Central Business District and a Large Suburban Shopping Mall Considering Residents' Behavior When Shopping in Multiple Stores at a Time," *Urban Reg. Plan. Rev.*, vol. 12, pp. 102–125, 2025, doi: 10.14398/urpr.12.102.
5. J. Wang, "Cost control problems and countermeasures of e-commerce enterprises under the background of big data and Internet of Things," *J. Comput. Methods Sci. Eng.*, vol. 23, no. 6, pp. 3135–3145, 2023, doi: 10.3233/JCM-226931.
6. E. Aldhahri, "The Use of Recurrent Nets for the Prediction of e-Commerce Sales," *Eng. Technol. Appl. Sci. Res.*, vol. 13, no. 3, pp. 10931–10935, 2023, doi: 10.48084/etasr.5964.
7. R. E. Widodo and T. A. Napitupulu, "Exploring the Impact of Live Streaming for E-Commerce Business: A Systematic Literature," *J. Theor. Appl. Inf. Technol.*, vol. 101, no. 16, 2023.