

Research Article

Discovery of Product Features for Redesign from User Implicit Feedback

Xiaolong Wang ¹, Yu Wang ^{1,2}, Jiacong Wu,¹ and Sara Shafiee³

¹School of Management, Jinan University, Guangzhou 510632, China

²International Business School, Jinan University, Zhuhai 519070, China

³Department of Mechanical Engineering, Technical University of Denmark, Kongens Lyngby 2800, Denmark

Correspondence should be addressed to Yu Wang; twygs@jnu.edu.cn

Received 20 April 2022; Accepted 29 July 2022; Published 21 August 2022

Academic Editor: Zahid Mehmood

Copyright © 2022 Xiaolong Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Online reviews are key sources to help manufacturers design and improve their products. However, using online reviews for product redesign is challenging because many online reviews contain useless information, and customers rarely provide explicit information about the products. The few methods available to recognize the customers' demands from online reviews have several limitations, including failing to consider the redundancy of useless reviews, confusion in providing targeted information among different aspects of product design, and disregarding the manufacturing limitation. Hence, considering customer needs and manufacturing constraints for product improvement, this study proposes a new model called DPFR to discover the potential product features for redesign. In DPFR, it first analyzes and predicts helpfulness of reviews for more effective decision of product design and improvement, which is neglected in previous research. Then, a structured model based on topic modeling is conducted. Overlooked by previous researchers, DPFR also introduces a feature importance matrix to measure the significance of each commented feature to product redesign. Third, a target features selection model is conducted further considering the limitations of manufacturers rather than only focusing on customer satisfaction. The effectiveness of the proposed model was assessed using a case study of data acquired from Bilibili, a large short-video platform in China. Focusing only on customer satisfaction, the results confirm the effectiveness of DPFR. Finally, including the constraints of product redesign, DPFR balances customer satisfaction and the manufacturing constraints such as redesign cost, redesign time, and carbon emissions. This study provides more practical suggestions for manufacturers to discover product redesign features that result in strategic product improvement decisions.

1. Introduction

Following a dynamic market environment, redesign is widely required to improve the existing products [1]. Most product improvements are based on the redesign of existing products [2]. In the product design and improvement process, according to a survey, 70% of manufacturers choose to redesign rather than design from scratch [1]. In the process of product redesign, it is vital to meet customers' needs by discovering the product features for the redesign [3, 4]. Identifying potential product features for redesign can help manufacturers formulate product improvement strategies [5].

Researchers and practitioners have investigated the identification of product features for redesign. However, most previous works relied on surveys and interviews to identify the product features for the redesign [5–7]. Guided by the questions in the surveys and interviews, customers can express their demands and preferences explicitly. However, surveys and interviews require a significant investment in manpower and resources [6]. In recent years, with the rapid development of information technology, the social media developed rapidly. We have witnessed a convenient platform for customers to share opinions and experiences about the products. Furthermore, online reviews about products have also become an essential source for

manufacturers to understand the customers' demands [8–11].

Considering recent progress in social media, researchers have used online reviews to identify product features for the redesign [5, 7, 12]. Although online reviews conveniently allow customers to express their opinions about products, customers rarely provide explicit information. Thus, the product features suitable for the redesign are hidden in these online reviews [13]. The primary challenge is finding an effective solution to utilize online reviews for product redesign. Several prior studies have focused on mining product features for redesign. Conjoint analysis with the KANO model is a widely used method for identifying potential product features for redesign [5, 8, 12, 14]. The goal of conjoint analysis is to determine the weight of all potential product features eligible for the redesign. The KANO model is used to measure the level of importance of potential product features [15, 16]. Nevertheless, conjoint analysis still has limitations. Hence, the main gaps in the current literature are discussed below.

First, online reviews contain a lot of useless information, which brings redundancy to the process of selecting potential product features for redesign. However, high-quality online reviews positively affect the consumer purchase intentions [17]. Manufacturers are therefore challenged with capturing customer feedback from helpful online reviews to improve product design [18]. However, previous studies have neglected how helpful reviews influence decisions surrounding mining product features for redesigns. Second, online reviews present different *product topics* discussed by customers. Considering online reviews of electronic products as an example, these reviews usually include different product features, such as CPU, speaker, screen, and battery, which belong to the hardware dimension of the product. In this paper, the different dimensions of products are called *product topics*. The *product topics* can be identified by the topic modeling of online reviews but have been neglected by the researchers. In selecting the potential features for the redesign, the level of importance or weight of each product topic also needs to be considered [4]. Third, previous studies have focused on selecting product features for redesign based on online reviews only to meet the customer satisfaction factors [12, 19]. However, focusing on product redesign, we must consider not only the customer needs but also the constraints of product redesign from the manufacturers' perspective.

To overcome the limitations of existing studies, this study aims to answer the following research questions:

- (1) How can we benefit from helpful reviews to explore potential product features that are suitable for the redesign?
- (2) How can we extract the customers' needs based on the helpful reviews on the *product topics*?
- (3) How can we construct a model to select product features for redesign, considering both the customer needs and the limitations of manufacturers?

To address the above questions, this paper presents the DPFR method to identify potential product features for the

redesign. The DPFR consists of three models. The first model discovers helpful online reviews from the manufacturers' perspective. Second, a structured model based on topic modeling is constructed to cluster helpful online reviews into different *product topics* and calculate the importance of each topic. Hence, an index is subsequently introduced from the structured model to measure the priority of all potential product features for the redesign. Finally, the product features selection model for the redesign is constructed, considering both customers' needs and the constraints of product redesign.

The contributions of this study can be summarized as follows. First, this study explains how helpful reviews contribute to product redesign and improvement. Second, this study investigates the influence of *product topics* that the customers discuss on social media. Third, this study considers not only customer satisfaction but also product redesign constraints, which help to develop product redesign strategies. Fourth, DPFR can be used as a real-time customer monitoring tool to help product designers cope with the rapid changes in customers' needs and formulate plans for product redesign in time.

The rest of the paper is organized as follows: The next section briefly describes the related works. Section 3 introduces the proposed method for identifying product features for the redesign. Section 4 uses the case study of Huawei MatePad 11 and the data acquired from Bilibili to verify the feasibility and validity of the proposed method. Finally, the discussion and conclusions are presented in Section 5.

2. Related Works

2.1. Identification of Helpful Online Reviews. In the Web 2.0 era, customers can share their product experiences and opinions online. A product can easily receive hundreds or even thousands of online reviews from customers ([20–22]. Thereafter, the reviews provide valuable feedback and insights from the customers' perspective for product manufacturers and service providers, enabling them to make sustainable decisions on product innovation and improvements [5]. However, online reviews also contain useless reviews that have nothing to do with product design and functionality. Therefore, it is particularly important to identify helpful product reviews [18, 23]. Previous studies have used various features extracted from customer reviews and classification models to identify helpful reviews, as summarized in Table 1. The classification models used in Table 1 include an artificial neural network, support vector machine (SVM), empirical analysis, naive Bayes (NB), random forest (RF), and K-nearest neighbor (KNN).

The length of a review is a highly influential review factor [29]. The majority of the studies have proven that the length of a review has a positive effect on predicting the level of the review's helpfulness [24–26, 29]. To automatically evaluate the usefulness of online reviews, Kim et al. [27] selected features (e.g., punctuation, review length, and star rating) from Amazon.com product reviews. Then, the support vector regression was trained to learn the helpfulness

TABLE 1: Summary of existing studies for identifying helpful reviews.

Categories	Features	Literatures
Length	Number of words	Martin and Pu [24]; Eslami et al. [25]; Luo et al. [26]; Meng et al. [23]
	Number of sentences	Kim et al. [27]; Martin and Pu [24]; Zeenia Singla and Sushma [28]; Eslami et al. [25]; Meng et al. [23]; Lutz et al. [29]
Sentiment	Positive words	Cao et al. [30]; Chua and Banerjee [31]; Zeenia Singla and Sushma [28]; Liu et al. [32]; Oumayma Oueslati and Sushma [33]; Zeng et al. [34]; Chua and Chen [35]
	Negative words	Cao et al. [30]; Chua and Banerjee [31]; Zeenia Singla and Sushma [28]; Liu et al. [32]; Oumayma Oueslati and Sushma [33]; Zeng et al. [34]; Chua and Chen [35]
	Number of positive words	Kaushik et al. [36]; Li et al. [37]
	Number of negative words	Kaushik et al. [36]; Li et al. [37]
Lexical	Number of nouns	Kim et al. [27]; Kikuchi and Klyuev [38]; Krishnamoorthy [39]; Jiang et al. [18]; Du et al. [40]; Shan et al. [41]
	Number of verbs	Kim et al. [27]; Krishnamoorthy [39]; Jiang et al. [18]; Du et al. [40]; Shan et al. [41]
	Number of adverbs	Kim et al. [27]; Krishnamoorthy [39]; Jiang et al. [18]; Du et al. [40]; Shan et al. [41]
	Number of adjectives	Kim et al. [27]; Kikuchi and Klyuev [38]; Krishnamoorthy [39]; Jiang et al. [18]; Du et al. [40]; Shan et al. [41]

function, subsequently identifying the length of a review as one of the most useful features. Eslami et al. [25] used an artificial neural network to predict the helpfulness of reviews based on review length, review score, and argument frame. The results showed that review length was the most helpful factor in predicting the helpfulness of reviews. In order to select helpful reviews in user-generated content platforms, Luo et al. [26] used empirical studies to explore the moderating effect of reviews features on review helpfulness based on a large-scale online reviews data. Hereby, the results showed that the review length has a significant effect on review helpfulness. Lutz et al. [29] developed a model to test the validity of review length for identifying helpful reviews, and the results showed that the review length was useful for predicting helpful reviews with a fixed emotional orientation.

Sentiment features are another influential factor that is strongly correlated with review helpfulness prediction [33, 34, 42, 43]. Sentiment features reveal the customer experience of products, mainly including negative, positive, and neutral emotional tendencies [30, 34]. Through the analysis of negative and positive features, manufacturers can evaluate customer needs and preferences about their products [28, 33]. Through empirical research, Cao et al. [30] examined the influence of various features (e.g., sentiment features) on review helpfulness and concluded that reviews with positive and negative features are more useful than those with mixed or neutral features. Chua and Banerjee [31] classified sentiment features into positive, negative, and mixed. They indicated that positive features are more related to helpful votes. Zeng et al. [34] examined the impact of sentiment features and product type on review helpfulness predication, and the results revealed that sentiment features are more useful than the product type. Chua and Chen [35] used empirical studies to explore the relationship between useful and false reviews, and the results suggested that sentiment features can help to distinguish helpful false reviews from helpful authentic reviews.

Together with review length and sentiment features, studies have identified that lexical features have an effect on review helpfulness [41, 44, 45]. Nouns, verbs, adverbs, and adjectives are widely used as lexical features [18, 27, 38–40, 46]. Krishnamoorthy [39] used the NLTK Parts-of-Speech (POS) tagger to tag each customer review and reported that lexical features are quite effective in predicting review helpfulness. Jiang et al. [18] proposed a multi-class classification approach based on ensemble learning to identify helpful reviews and verified that the number of lexical features is an important factor in determining the classifier performance. Du et al. [40] selected 30 different features from relevant research papers to test the performance of different feature combination patterns in helpful review predications. They suggested that the feature combination based on nouns, verbs, adverbs, and adjectives leads to visible performance. Shan et al. [41] constructed a machine learning method based on 22 features to identify fake reviews and analyzed the importance of different features, and the results showed that nouns, verbs, and adjectives are important features to identify the fake reviews.

2.2. Topic Modeling. Topic modeling is a text-mining method for topic discovery [47, 48]. Researchers have applied topic modeling in various fields (e.g., online review analysis and medical sciences) [47, 49–51]. Topic modeling based on latent Dirichlet allocation (LDA) is widely used in online review analysis to find business opportunities, such as extracting service quality dimensions [49], identifying customer satisfaction dimensions [12, 52], and identifying latent product topics discussed by customers [4, 53, 54]. To measure healthcare service quality, James et al. [55] proposed a text-mining method to analyze large amounts of unstructured patient feedback using the LDA clustering model. They verified that the topic areas extracted by LDA are useful in evaluating physician performance based on feedback. Nilashi et al. [52] proposed a combined method based on LDA, TOPSIS, and neural networks for customer

segmentation in green hotels. LDA was used to extract different aspects of customer satisfaction. The results showed that LDA is effective in discovering a mixture of topics from online reviews. Jeong et al. [4] presented an approach based on LDA and sentiment analysis to mine opportunities for further product improvement and highlighted the importance of topics discussed by the customers. In order to get the preference of online game community players, Yu et al. [54] used LDA to analyze the reviews collected from Steam game community, and finally obtained the players' favorite topics.

2.3. Identification of Product Features from Online Reviews.

The primary task of product redesign is to capture redesign requirements (e.g., features that need to be improved), which is an important research area for improving product quality and reducing product costs [56, 57]. Traditional product redesigns are mainly based on quantitative, qualitative, or hybrid methods to discover customers' needs, using interviews, laboratory research, and market research [58–61]. These methods often require labor-intensive work, such as organizing tests, finding suitable participants, designing appropriate questionnaires, and testing products [7]. Although traditional methods have proven to be effective, they have limitations due to the small sample size and high costs [6].

With the rapid growth of big data, online reviews have become an important data source for product redesign and innovation in rapidly changing market [11, 19, 62]. Analyzing customer needs and preferences, online reviews have been used in many studies to discover to-be-improved features. The majority of studies have focused on discovering product features that have a great effect on customer satisfaction, as shown in Table 2. Conjoint analysis with the KANO model has been widely used in previous studies [5, 8, 12, 14]. The goal of conjoint analysis is to determine the weight of all potential product features for the redesign [65]. Sentiment analysis is a widely used method in conjoint analysis that calculates the customer satisfaction level of each product feature. The KANO model is then applied to develop suggested product improvement strategies [15]. In general, potential product features are classified into five KANO categories: attractive, one-dimensional, must-be, indifferent, and reverse. Martínez-Torres [66] proposed a method based on conjoint analysis and KANO to extract customer changing expectations for product affordances. Moreover, Wu et al. [16] proposed a framework to discover the priority of product/service improvement based on online reviews by using semantic analysis and the KANO model to capture products/services features. In the framework, sentiment analysis was used to calculate the performance of the features, and then the KANO model was used to measure the importance of each feature's performance. Finally, product features were classified into attractive, one-dimensional, must-be, and indifferent in their study. They reported that their method is effective in analyzing customers' needs. Bi et al. [12] used the KANO model to identify different customer satisfaction dimensions, and the results showed

that manufacturers mainly focus on one-dimensional and attractive categories for product improvement process. Joung and Kim [14] used the KANO model to rank products from the customers' perspective, and the results showed that the KANO model could recognize different types of customer preferences.

As presented, previous studies have discovered that the quality of online reviews is vital for both customers and product manufacturers. Many studies have investigated the key effective features of identifying helpful reviews. However, these studies did not explain how helpful reviews can benefit strategic decisions for the product improvement.

Investigating the potential product features, the majority of researchers focused on mining customer preferences and opinions about products from customers' perspectives by calculating the customer satisfaction level of the product features. However, many studies have neglected the effects of *product topics*. Furthermore, the literature lacks a method for extracting and using customer preferences and opinions about products. Focusing on product redesign, manufacturers must not only consider the customers' need but also the constraints of product redesign.

Accordingly, this study proposes a novel method that aims to discover the product features from users' implicit feedbacks for the product redesign. The research results have broad impacts on product redesign and product improvement strategies using the analytics of online reviews and capturing customer needs.

3. Methodology

Figure 1 shows the framework of the proposed method (DPFR) for identifying to-be-improved product features based on online reviews for product redesign. In this framework, the process of identifying potential product features consists of three steps. The first step is to create a model to extract helpful reviews. Second, a structured model based on topic modeling is constructed to identify customers' needs. Two major tasks in this step include the identification of product topics and the calculation of the importance of each topic. The candidate target features are stored in the structured model. Then, an index called *feature attention (FA)* is subsequently introduced from the structured model to measure the priority of all potential product features for the redesign. In the third step, a product feature selection model for the redesign is created, considering both the customer needs and the constraints of product redesign.

Moreover, we considered three different constraints for our optimal configuration model. Product redesign cost and lead time refer to two very important factors that have been discussed in the previous research [67, 68]. Product redesign costs consist of traditional engineering costs such as materials, labor, and product transportation, as well as environmental protection costs throughout the lifecycle of products [67] and lead time refers to Hammami et al. [68] which cannot exceed the requested customer lead time. Further, low-carbon sustainability has become an important factor in product design [69, 70]. In the rapidly changing market, the lifecycle of products is shortened, which leads to

TABLE 2: Summary of existing research to identify the product features for redesign based on focused areas.

Suggested models/methods	Focused areas	
	Customer satisfaction	The constraints of product improvement
Qi et al. [5]	√	×
Wu et al. [16]	√	×
Bi et al. [12]	√	×
Hou et al. [8]	√	×
Yang et al. [7]	√	×
Zhou et al. [63]	√	×
Jin et al. [64]	√	×
Joung and Kim [14]	√	×
Lee et al. [15]	√	×
This study	√	√

rapid product iteration, contrary to the concept of low-carbon sustainable development [71–73]. Therefore, manufacturers should consider low-carbon sustainability when designing the products.

3.1. Identification of Helpful Reviews. Effectively identifying helpful reviews is important to product design. Review length, sentiment features, and lexical features are used to predict helpful reviews:

- (1) *Review Length* ($F_L = \{f_{L1}, f_{L2}\}$). For each review, we selected the number of words (f_{L1}) and the number of sentences (f_{L2}) as the features.
- (2) *Sentiment Features* ($F_{ST} = \{f_{ST1}, f_{ST2}\}$). For each review, we selected the number of positive words (f_{ST1}) and negative words (f_{ST2}) as sentiment features.
- (3) *Lexical Features* ($F_{SP} = \{f_{SP1}, f_{SP2}, f_{SP3}\}$). For each review, we selected the number of nouns (f_{SP1}), the number of verbs (f_{SP2}), the number of adverbs (f_{SP3}), and the number of adjective (f_{SP4}) as lexical features.

Naive Bayes, KNN, RF, and SVM algorithms that have been widely used to identify helpful reviews were selected as the key algorithms for this study [17, 18, 39, 74].

3.2. The Structured Model. The structured model is constructed to describe the customers' needs. In this phase, the latent product topics discussed by customers are identified based on helpful reviews that identify directions for product improvement. Hereafter, the keywords are extracted that reflect customers' needs and preferences. Constructing the structured preference model consists of two phases. First, topic modeling was used to identify the latent *product topics* from helpful reviews and quantify the importance of each topic. Second, the *FA* index was introduced to calculate the priority and weight of all potential product features for the redesign.

3.2.1. Topic Modeling. Before topic modeling, we need to process and filter the helpful reviews because the vocabularies also contain a large number of irrelevant words. The

negation words, sentiment words, and degree words are filtered out. Then, LDA was selected as the topic model to train the helpful reviews. LDA can obtain the probability distribution of keywords on the topic. The keywords unrelated to the target products need to be manually removed.

LDA is a three-layer Bayesian probability model, shown in Figure 2. The model used in this study is briefly described as follows:

- (1) Let M denotes the set of helpful reviews, K is the number of latent topics, W_p denotes the importance of each topic, and $P \in (1, K)$.
- (2) Each helpful review, m (N_m is the length of m), has its topic distribution. The topic distribution follows a multinomial distribution. The parameters (θ_m) of the multinomial distribution follow the Dirichlet distribution, $\theta_m \sim \text{Dirichlet}(\alpha)$. α is the prior parameter of the Dirichlet distribution.
- (3) Each topic ($P \in (1, K)$) has its word distribution, which follows the multinomial distribution. The parameters (ϕ_k) of the multinomial distribution follow the Dirichlet distribution, $\phi_k \sim \text{Dirichlet}(\beta)$. β is the prior parameter of the Dirichlet distribution.
- (4) For each of the N words $w_{m,n}$, the first step is to choose a topic $Z_{m,n}$, followed by choosing a word $w_{m,n}$ from $p(w_{m,n}|Z_{m,n}, \beta)$, a multinomial probability limits the topic $Z_{m,n}$. This random process is repeated until all helpful reviews have been captured. Finally, we can obtain the topics-reviews matrix (TR Matrix), as shown in Figure 3. The importance of each topic (W_p) can be quantified from the topic mentioned in each review. W_p is calculated as follows:

$$C_p = \sum_{p=1}^K \sum_{i=0}^n \text{TRMatrix}_{p,i}, \quad (1)$$

$$W_p = \frac{C_p - C_{\min}}{C_{\max} - C_{\min}}, \quad (2)$$

where K is the number of topics, $p \in [1, K]$, and i denotes the review, $i \in [1, n]$.

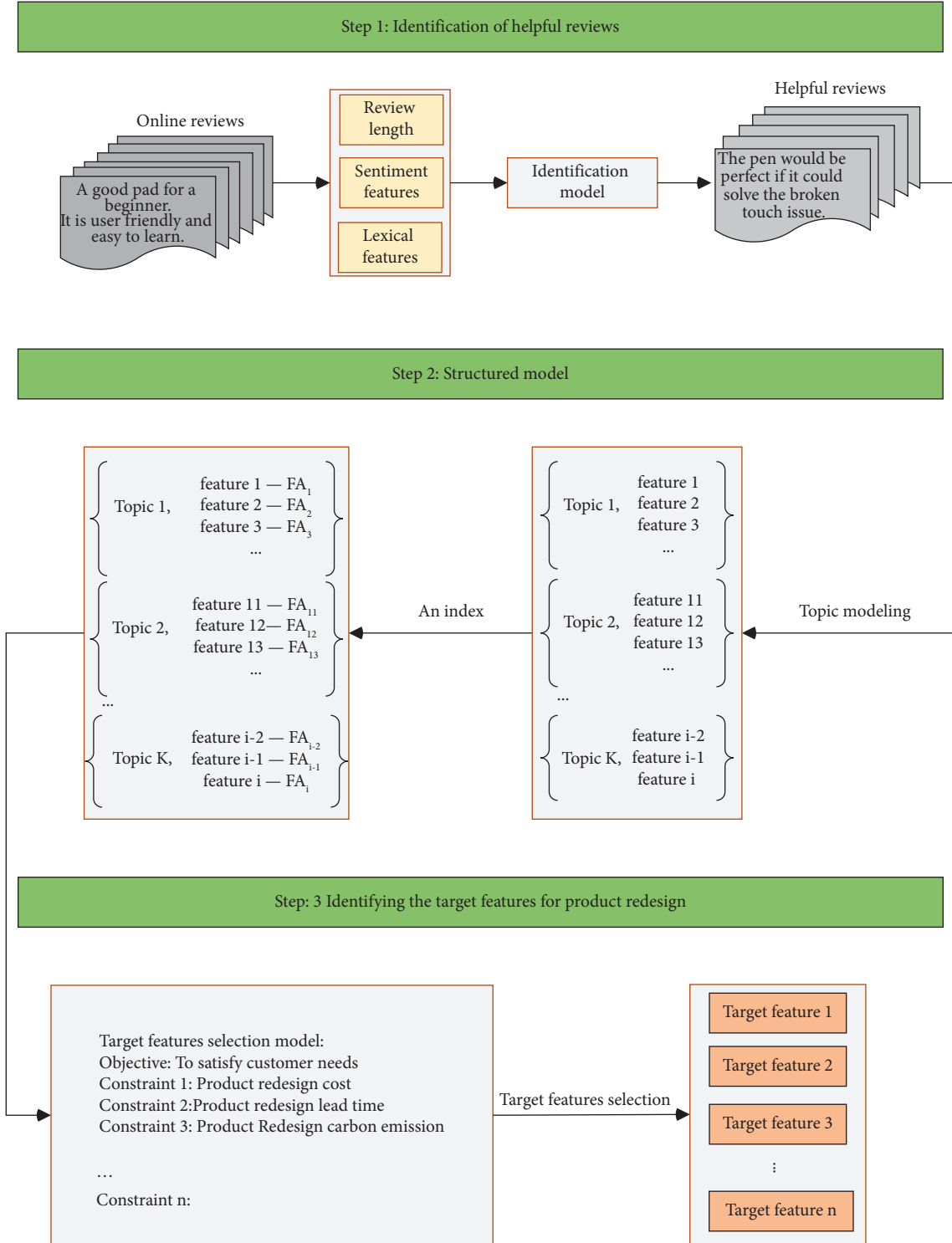


FIGURE 1: The framework of the proposed DPFR model.

Although LDA can identify product topics, it cannot determine the number of product topics. The most popular model for selecting the number of topics is perplexity. Although perplexity reflects the predictive ability of the model, pursuing predictive ability blindly leads to the problem of having a large number of topics. Hence, only similar topics will appear, resulting in a low recognition of the topics

[75, 76]. To solve this challenge, this study proposes a new measure (*PSA*) to select some topics. *PSA* is an extension of perplexity and considers the similarities between topics.

PSA can be defined as follows:

$$\text{Perplexity_var}(M) = \frac{\text{Perplexity}(M)}{\text{var}(T)}. \quad (3)$$

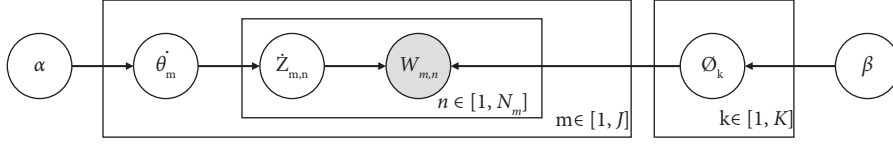


FIGURE 2: The graphical model of the LDA.

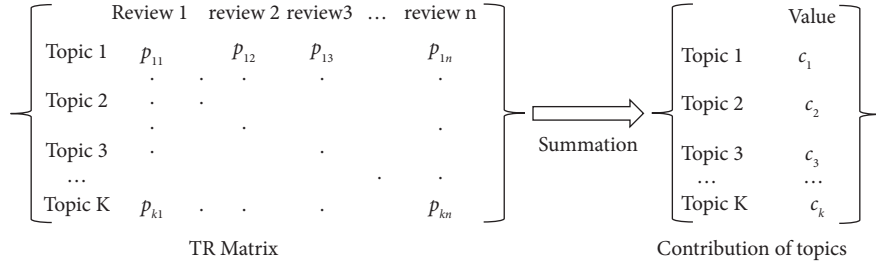


FIGURE 3: Concepts for calculating the importance value of product topics.

Let $\text{Perplexity_var}(M)$ as the measurement index for selecting K . M is denoted as the set of helpful reviews.

$$\text{Perplexity}(M) = \exp \left\{ -\frac{\sum_{m=1}^M \log p(w_m)}{\sum_{m=1}^M N_m} \right\}. \quad (4)$$

Let w_m denote the word vector of m , $m \in [1, M]$. N_m is the number of words in m .

$$\text{var}(T) = \frac{\sum_{T=1}^K [\text{corre}(T) - \text{avg_cos}]^2}{K}, \quad (5)$$

where $\text{var}(T)$ represents the degree of deviation between each topic and the average similarity of the topics. K is the number of latent topics, and $T \in [1, K]$.

$$\text{vg_cos} = \frac{\sum_{i=1}^K \sum_{j=i+1}^K \text{corre}(Z_i, Z_j)}{K * (K - 1)/2}, \quad (6)$$

$$\text{corre}(Z_i, Z_j) = \cos(\mu) = \frac{\sum_{i=1}^n A_i B_i}{\sum_{i=1}^n (A_i)^2 \sum_{i=1}^n (B_i)^2} = \frac{AB}{|AB|}.$$

Let avg_cos denote the average similarity of topics. A and B are n -dimensional topic vectors, $A = (A_1, A_2, \dots, A_n)$, $B = (B_1, B_2, \dots, B_n)$, $\cos(\mu)$ is the cosine value of A and B .

3.2.2. An Index for Product Redesign. 1. *Sentiment Analysis of Potential Features Associated with Each Topic.* Customer sentiment contains positive, negative, and neutral emotions. Moreover, creative ideas and other needs are mainly hidden in positive and negative reviews. Hence, this study mainly focuses on positive and negative reviews.

We define a 3-tuple, F_i , LP_i , and GP_i . In this 3-tuple, F_i denotes each product feature captured from helpful reviews. LP_i denotes the local preference for the feature F_i . GP_i denotes the global preference for the feature F_i .

The local preference can be denoted as a set $LP_i = \{n_{ij}, p_{ij} | j = 1, 2, \dots, j\}$, where n_{ij} , and p_{ij} are the negative and positive opinion frequencies on the feature i , in the helpful review j .

$$n_{ij} = \text{count}(Op_{ij}^-), \quad (7)$$

$$p_{ij} = \text{count}(Op_{ij}^+). \quad (8)$$

The global preference can be denoted as a set $GP_i\{N_i, P_i\}$, where N_i and P_i are the negative and positive opinion frequencies, respectively, on the feature i in all helpful reviews.

$$N_i = \sum_{j=1}^M n_{ij}, \quad (9)$$

$$P_i = \sum_{j=1}^M p_{ij}. \quad (10)$$

2. The Calculation of Feature Attention. Rai [77] defined a quantitative metric based on customers' needs called the local global normalization measure (LGNM). LGNM uses three different types of term weighting: local, global, and normalization. In our study, we used LGNM to evaluate FA :

$$FA_i = \sum_{j=1}^M L_{ij} * G_i * N_j, \quad (11)$$

where M is the total number of helpful reviews. L_{ij} is denoted as local weight for feature i in the helpful review j . G_i is the global weight used to measure the importance of candidate feature i in all helpful reviews. Let N_j denotes the normalization factor to reduce the effect due to the lengths of the helpful reviews.

The local weight L_{ij} is calculated by

$$L_{ij} = \log_2(1 + f_{ij}), \quad (12)$$

where f_{ij} is the number of occurrences of the potential feature i in the helpful review j . It can be calculated as

$$f_{ij} = n_{ij} + p_{ij}. \quad (13)$$

The local weight G_i is calculated as

$$G_i = 1 + \sum_{j=1}^M \frac{(f_{ij}/F_i) \log(f_{ij}/F_i)}{\log M}, \quad (14)$$

where F_i is the number of occurrence of the feature i in all helpful reviews. It can be calculated as

$$F_i = N_i + P_j. \quad (15)$$

The normalization factor N_j is calculated by

$$N_j = \frac{1}{\sqrt{\sum_i (G_i * L_{ij})^2}}. \quad (16)$$

FA can be calculated by the following equation.

$$FA_i = \sum_{j=1}^M [\log_2(1 + n_{ij} + p_{ij})] * \left(1 + \sum_{j=1}^M \frac{(n_{ij} + p_{ij}/N_i + P_j) \log(n_{ij} + p_{ij}/N_i + P_j)}{\log M} \right) * \frac{1}{\sqrt{\sum_i (G_i * L_{ij})^2}}. \quad (17)$$

3.3. The Product Feature Selection Model for the Redesign.

The potential features extracted from the structured model contain customer preferences to help designers understand the customers' needs. However, the structured model only describes the opinions and attitudes of customers. During the product redesign process, the company's possibilities and manufacturing limitations to implementing the customers' needs must-be evaluated.

In this study, we constructed an optimization model to select potential features to be improved in product redesign. The objective of this model is to satisfy the customers' needs considering the production constraints. We considered the product redesign cost, redesign lead time, and carbon emissions as the primary constraints.

Product redesign cost is the first constraint that needs to be considered. We considered traditional engineering costs, such as materials, labor, and product transportation, as well as environmental protection costs throughout the lifecycle of products [67].

$$RC_i = TC_i + EC_i, \quad (18)$$

where RC_i is the cost of product redesign. TC_i and EC_i are denoted as the engineering costs and environmental protection costs for the potential feature i , respectively.

In this study, we assumed that all redesign tasks were conducted at the same time. Thus, the redesign lead time was determined by the task requiring the longest lead time among all the redesign tasks.

$$RT = \text{Max}(t_i * x_i), \quad (19)$$

where t_i is the time of redesign task for potential feature i . x_i is the 0–1 decision variable. If the potential feature i is selected, $x_i = 1$, otherwise $x_i = 0$.

In this study, we set a carbon emission threshold for product redesign. We assumed that all product redesign tasks could not exceed this threshold.

$$\sum_1^n e_i x_i \leq E, \quad (20)$$

where e_i is denoted as the carbon emission of the redesign task for feature i , and E is the carbon emission threshold of all the redesign tasks.

The objective function of the target product feature selection model in product redesign is defined in equation (21). As we considered the importance of product topics (w_p), x_{pi} is denoted as the decision variable when the candidate feature i in topic p is selected, $x_{pi} = 1$, otherwise, $x_{pi} = 0$. l_p is denoted as the number of features in topic p .

$$\max U = \sum_{p=1}^k \sum_{i=1}^{l_p} w_p * FA_{pi} * x_{pi}. \quad (21)$$

The constraints of the target product feature selection model can be denoted as

$$\left\{ \sum_{p=1}^k \sum_{i=1}^{l_p} (TC_{pi} + EC_{pi}) * x_{pi} \leq C, \text{Max}(t_{pi} * x_{pi}) \leq T, \sum_{p=1}^k \sum_{i=1}^{l_p} e_{pi} * x_{pi} \leq E. \right. \quad (22)$$

4. Evaluation

4.1. Data. In recent years, with the rapid development of social media platforms, new media, mainly in the form of short videos, have become popular. These short-video platforms provide rich real-time data and are popular channels for customers to exchange their opinions about the

product. Bilibili is one of the largest short-video platforms in China, and there are many online reviews of products provided by customers on the platform. We selected Matepad 11, a tablet computer manufactured by Huawei, as the product to test our model due to the availability of a large number of online reviews about Matepad 11 on Bilibili. Based on the comprehensive ranking of uploaded videos by

Bilibili, 10 videos about Matepad 11 were selected to collect online reviews. *Python* 3.7 was used to clean the online reviews. To clean the data, we mainly removed duplicated and abnormal data, normalized the text, segmented the words, and deleted stop words. This resulted in a collection of 21112 customer reviews about Huawei Matepad 11.

4.2. Identification of Helpful Reviews. A model to identify helpful reviews was created based on the explanations in Section 3.1. We divided the training and test datasets according to a ratio of 7:3. The average 10-fold cross-validation was used as the performance of each model. In our experiments, we have set enough epoch for model training, and the parameters of each were iteratively optimized until the model converged on our training set. The performance of each model on the testing set is shown in Table 3. Precision, recall, and the *F*-measure were used to evaluate the performance of the classification algorithms:

$$\begin{aligned} \text{precision} = P &= \frac{TP}{(TP + FP)}, \\ \text{recall} = R &= \frac{TP}{(TP + FN)}, \\ F_measure &= \frac{2 * P * R}{(P + R)}, \end{aligned} \quad (23)$$

where *F*-measure denotes the harmonic mean of precision and recall. True positive (TP) refers to positive samples that are correctly judged as positive, false positive (FP) refers to negative samples that are incorrectly judged as positive, and false negative (FN) refers to positive samples that are incorrectly judged as negative. Table 3 shows the results obtained using different models of helpful review identification.

Based on the numbers in Table 3, the SVM as the kernel algorithm performed better than NB, KNN, and RF algorithms in reviews identification. The precision of the model reached its best when SVM was used. The performance of SVM and NB on recall was close, but the *F*-measure was best when SVM was used. Therefore, we chose SVM to identify helpful reviews, resulting in the identification of 4111 helpful reviews.

TABLE 3: Results of identifying helpful reviews.

Models	Precision	Recall	F-measure
NB	0.857	0.862	0.859
KNN	0.823	0.836	0.829
RF	0.845	0.835	0.840
SVM	0.873	0.865	0.868

4.3. Construction of the Structured Model

4.3.1. Topic Modeling. As shown in Figure 4(a), the optimal number of topics selected by the PSV was 4. The optimal number of topics selected by perplexity was 6. The distances between each latent topic obtained using PSV and perplexity are shown in Figures 5(a) and 5(b), respectively. As shown in Figure 5(b), three topics with high similarity were obtained using perplexity. Therefore, PSV was selected to calculate the number of topics.

We named four topics selected from the PSV, as shown in Table 4. The real name of each topic was manually defined by its major features. The features such as “notebook” or “design” in “Topic 1” relate to office scenes. Thus, the label of “Topic 1” can be described as “Office Scene.” Similarly, the label of the other topics can be described separately as: “Operating System,” “Audio-Visual entertainment,” “Relevant Accessory and Hardware.” The degree of importance for each topic was calculated using equations (1) and (2), $W_p = (W_1, W_2, W_3, W_4) = (0.24, 0.35, 0.16, 0.25)$.

4.3.2. The Calculation of FA. A set of 38 features from the four topics were extracted, as shown in Table 4. The local and global preferences were calculated using equations (7)–(16). The *FA* of each feature was calculated using equation (17). The results are shown in Table 5.

4.4. Identifying the Potential Product Features for Product Redesign. In this empirical study, the total budget for product redesign was calculated as 3000000 dollars, the longest lead time for product redesign was calculated as 2400 hours, and carbon emission was calculated as 10000 kg. The to-be-improved product feature selection model was defined by

$$\max U = \sum_{p=1}^4 \sum_{i=1}^{I_p} w_p * FA_{pi} * x_{pi}. \quad (24)$$

Subject to

$$\left\{ \begin{aligned} \sum_{p=1}^4 \sum_{i=1}^{I_p} (TC_{pi} + EC_{pi}) * x_{pi} &\leq 3000000 \text{Max}(t_{pi} * x_{pi}) \leq 2400, \sum_{p=1}^4 \sum_{i=1}^{I_p} e_{pi} * x_{pi} \leq 10000. \end{aligned} \right. \quad (25)$$

As shown in Figure 6, after solving the target feature selection model for product redesign, the top 10 features were identified from the potential features. The manufacturers can choose the number of features for redesign

according to their actual situation of product redesign. In this study, the top 10 features were selected as the target features for redesign. These 10 features are *software usage ecology*, *system fluency*, *CPU*, *multi-screen collaboration*,

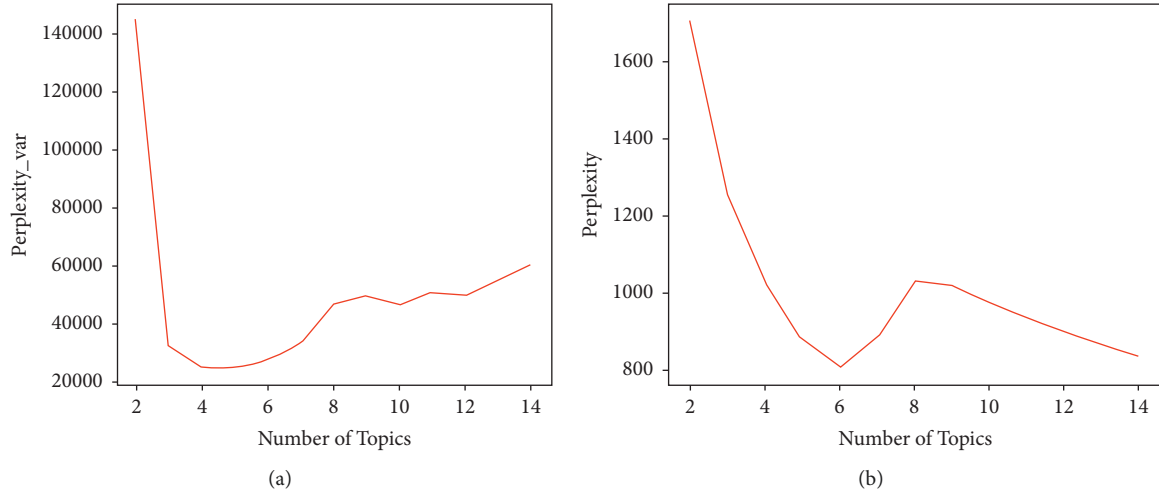


FIGURE 4: The selected optimal number of topics. (a) Select the optimal number of topics by PSV. (b) Select the optimal number of topics by perplexity.

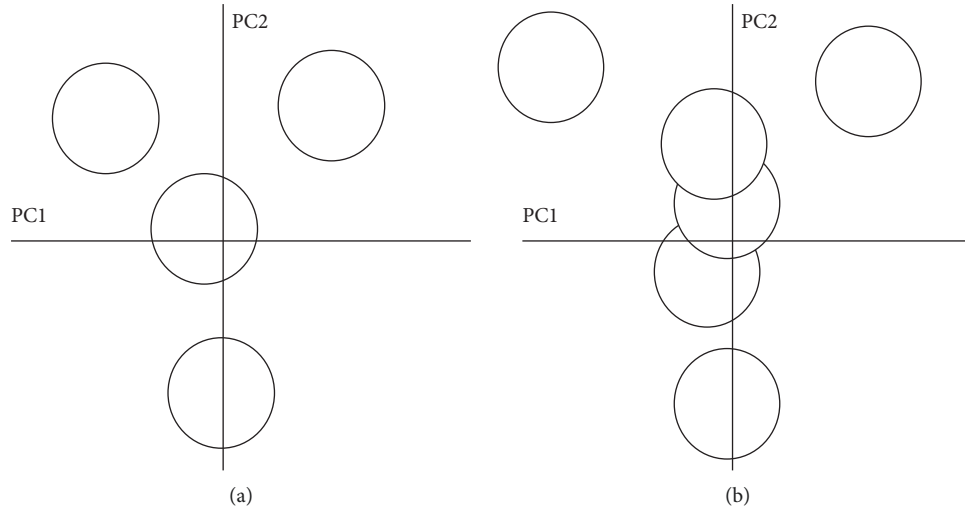


FIGURE 5: Distance between latent topics. (a) Topic distance calculated by the PSV. (b) Topic distance calculated by perplexity.

TABLE 4: Results of topic modeling.

Topics	Features
Topic 1, office scenes	Notebook, write, multi-screen collaboration, Huawei share, cloud, design, document, cast screen, MeeTime
Topic 2, operating system	Software usage ecology, bring two into view, app snippets, screen refresh rate, supper device, control panel, software optimization, backend system, search, system fluency
Topic 3, audio-visual entertainment	Film, painting, wallpapers, music, download, advertisement, network, photos, game
Topic 4, relevant accessory and hardware	Screen, camera, keyboard, color, speaker, memory, CPU, fingerprint, battery, stylus, bluetooth

supper device, bring two into view, notebook, Huawei share, screen refresh rate, and document. According to the results, the features *software usage ecology* and *system fluency* have significant potential for the product redesign. The features *multi-screen collaboration, notebook, Huawei share, and document* belong to Topic 1. The features *software usage*

ecology, system fluency, upper device, bring two into view, and screen refresh rate belong to Topic 2. The feature *CPU* belongs to Topic 4. Based on the knowledge in Section 4.2, Topics 1, 2, 3, and 4 are denoted as “Office Scenes, Operating System, Audio-Visual Entertainment, and Related Accessories and Hardware. As shown in Figure 7, the number of

TABLE 5: Results of the structured model.

Topic p	Candidate features i	Local preference LP_i		Global preference GP_i		Feature attention FA_i
		$n_{ij} (j = 1, 2, \dots, 4111)$	$p_{ij} (j = 1, 2, \dots, 4111)$	N_{ij}	P_{ij}	
Topic1	Notebook	$n_{1,1} = 1$	$p_{1,1} = 0$	187	151	156.52
		$n_{1,2} = 0$	\dots			
		\dots	$p_{1,12} = 1$			
	...	$n_{1,4110} = 0$	\dots	131	86	83.72
		$n_{1,4111} = 0$	$p_{1,4111} = 0$			
		$n_{1,1} = 0$	$p_{1,1} = 0$			
Topic2	Mee time	\dots	\dots	217	89	137.47
		$n_{1,14} = 1$	$p_{1,22} = 1$			
		\dots	\dots			
	Software usage ecology	$n_{1,4111} = 0$	$p_{1,4111} = 0$	164	155	134.64
		$n_{1,1} = 1$	$p_{1,1} = 0$			
		$n_{1,2} = 0$	\dots			
Topic3	...	\dots	$p_{1,17} = 1$	46	125	62.88
		$n_{1,4110} = 0$	\dots			
		$n_{1,4111} = 0$	$p_{1,4111} = 0$			
	System fluency	$n_{1,1} = 0$	$p_{1,1} = 0$	56	47	40.51
		\dots	\dots			
		$n_{1,21} = 2$	$p_{1,16} = 1$			
Topic4	Film	\dots	\dots	54	179	80.79
		$n_{1,4111} = 0$	$p_{1,4111} = 0$			
		$n_{1,89} = 0$	$p_{1,12} = 45$			
	...	\dots	\dots	34	87	56.06
		$n_{1,1} = 0$	$p_{1,1} = 0$			
		\dots	\dots			
Topic4	Game	$n_{1,121} = 1$	$p_{1,78} = 1$	34	87	56.06
		\dots	\dots			
		$n_{1,4111} = 0$	$p_{1,4111} = 0$			
	Screen	$n_{1,1} = 0$	$p_{1,1} = 0$	34	87	56.06
		$n_{1,1} = 0$	$p_{1,1} = 0$			
		\dots	\dots			
Topic4	Bluetooth	$n_{1,48} = 2$	$p_{1,14} = 1$	34	87	56.06
		\dots	\dots			
		$n_{1,4111} = 0$	$p_{1,4111} = 0$			
	...	$n_{1,1} = 0$	$p_{1,1} = 0$	34	87	56.06
		$n_{1,2} = 1$	\dots			
		\dots	$p_{1,41} = 1$			
Topic4	Bluetooth	$n_{1,4110} = 0$	\dots	34	87	56.06
		$n_{1,4111} = 0$	$p_{1,4111} = 0$			

target features accounted for 50% in Topic 2, 40% in Topic 1, and 10% in Topic 4. Based on the results from Figure 8, the topic “Operating System” had the largest proportion, and no feature was selected in Topic 3 as the potential features for the product redesign. However, this might be related to the strategic positioning of the Huawei Matepad series because the main selling points of the Huawei Matepad series are the operating system and smart office.

4.5. The Topic Weight Analysis for the Product Redesign. This study changed the weight of the topic importance to better describe the impact of the topic importance (w_1, w_2, w_3, w_4) on the selection of product design attributes. We set a transformation factor (h), which represents the increase or decrease of the weight of the topic importance each time, and h was set as 0.05, 0.1, and 0.15, respectively. After setting h , six experiments were conducted. For

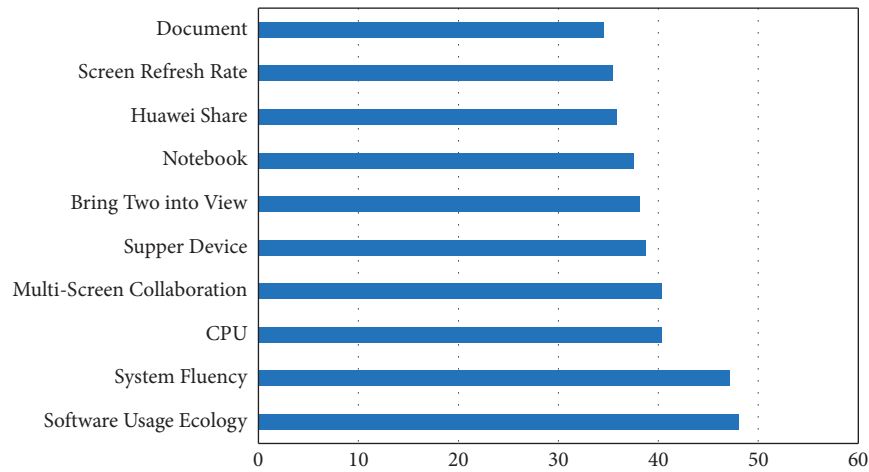


FIGURE 6: The target product features for the product redesign.

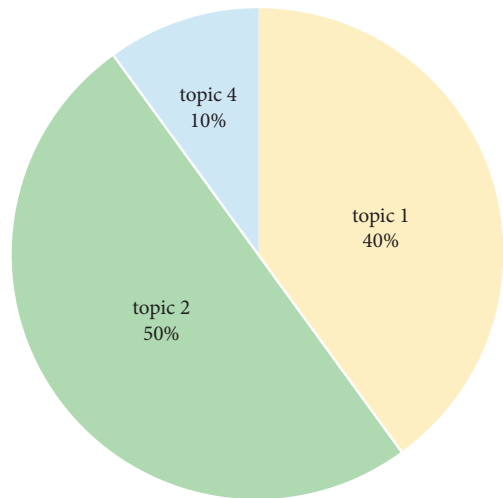


FIGURE 7: Proportion of the target product features for redesign.

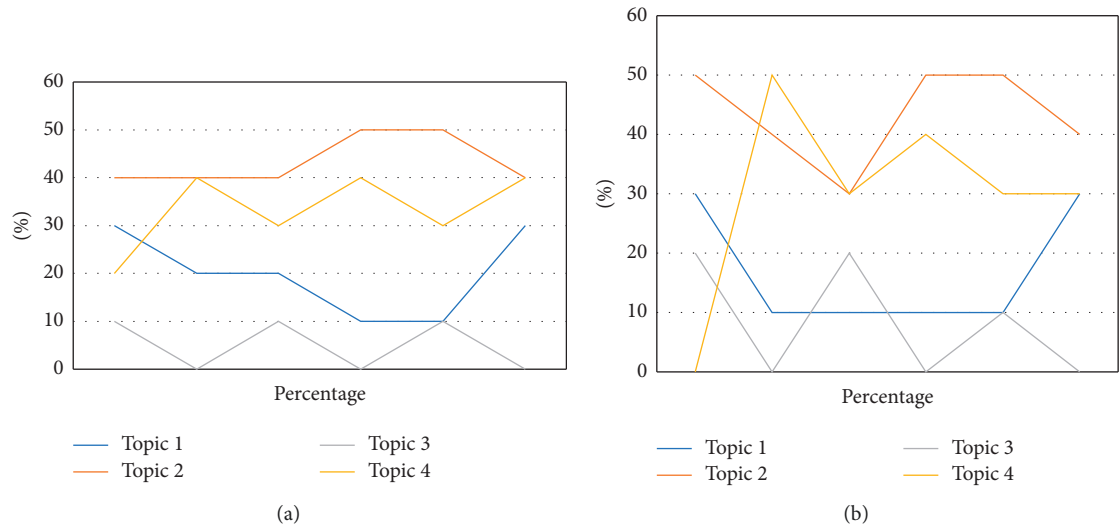


FIGURE 8: Continued.

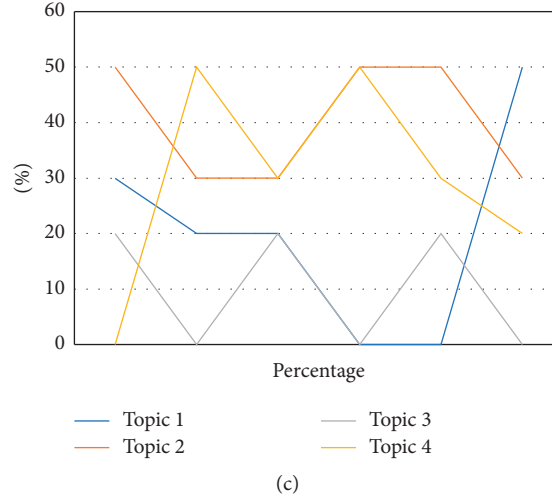


FIGURE 8: Line chart of percentages for each topic in different transformation factors. (a) $h = 0.05$. (b) $h = 0.1$. (c) $h = 0.15$.

example, when $h = 0.05$, in Experiment 1, w_1 and w_2 remained unchanged, w_3 increased by 0.05, and w_4 decreased by 0.05. Similar to Experiment 1, other experiments also had two control variables, with the weights of the other two topics directionally changed.

Figure 8 shows that with the change in topic weight, the percentage of each topic was constantly changing. Hence, the experiments illustrate that the changes in w_1 , w_2 , w_3 , and w_4 had a significant impact on the selection of product redesign features. With the increasing of transformation factor, the magnitude of changing percentage of each topic was increasing, as shown in Figure 9. Moreover, as shown in Figure 10, compared to w_1 , w_2 and w_4 , w_3 was relatively weak, and w_2 had the greatest influence. In this study, the feature selection model was effective in terms of topic importance. Topic importance is an indispensable key factor in product redesign. Customers and manufacturers paid significant attention to Topic 1—Office Scene, Topic 2—Operation System, and Topic 4—Relevant Accessory and Hardware, whereas Topic 3—Audio-visual Entertainment was of less concern. Among the four topics, Topic 2 was the most important future product redesign direction. Although the change in topic weight affected its proportion, Topic 2 accounted for a large proportion in each experiment.

4.6. Comparison with Other Models. To verify the validity of the proposed model (DPFR), we compared it with other models. Bi et al. [12] proposed a method from the customers' needs perspective. In their method, LDA is used to extract different customer satisfaction dimensions, and then the sentiment orientations of the customer satisfaction dimensions are measured based on the sentiment analysis. Therefore, the category of each customer satisfaction dimension is identified by the KANO model. According to Bi et al. [12], for product improvement process, manufacturers mainly focus on one-dimensional and attractive categories,

and the priority order of one-dimensional is higher than the attractive category. Therefore, product features that belong to one-dimensional and attractive categories are selected for improvement. By contrast, in the DPFR model, the customers' needs and production constraints are both considered. To further verify the validity of the DPFR, we need to have equal functionalities across the different models being compared. Hence, the production constraints proposed in Section 3.3 are removed from the DPFR model, and the selection model is constructed as equation (22). The results of the three methods are shown in Table 6.

$$\max U = \sum_{p=1}^k \sum_{i=1}^{l_p} w_p * FA_{pi} * x_{pi}. \quad (26)$$

As shown in Table 6, ignoring the product manufacturing constraints, 80% of the target features extracted by DPFR are the same as the features extracted using the KANO model by Bi et al. [12]. Hence, we can validate the effectiveness of DPFR model when focusing only on customers' needs. Including the constraints of product redesign, the results of DPFR highlight 3 target features different from the DPFR without considering the product constraints. The 3 target features are "huaweiwei share, screen refresh rate, and document" when manufacturing constraints are included in DPFR. While the 3 target features are "camera, stylus, and keyboard" when using DPFR without considering the product constraints. The reason for the mentioned difference is that redesign cost, redesign time, and carbon emissions are considered as constraints for product redesign in the DPFR. In the model of DPFR excluding product constraints, "camera, stylus, and keyboard" are highlighted, indicating that they attract attention from the customers, while these features are demanding regarding cost, time, and carbon emissions. In the actual production process, in order to select product features for the effective redesign, manufacturers should consider not only the

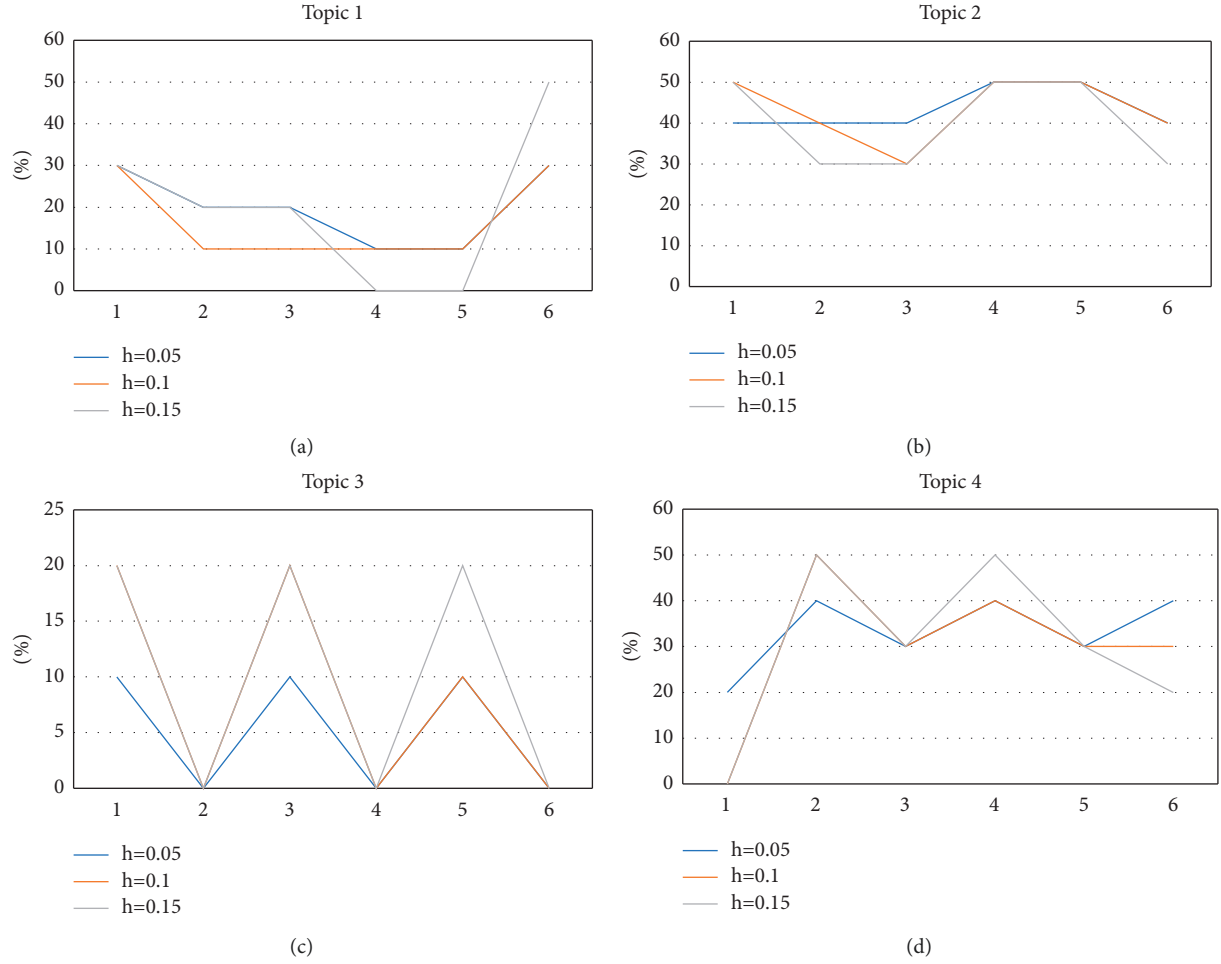


FIGURE 9: Line chart of changing trend of each topic in different transformation factors. (a) The changing trend of topic 1 in different transformation factors. (b) The changing trend of topic 2 in different transformation factors. (c) The changing trend of topic 3 in different transformation factors. (d) The changing trend of topic 4 in different transformation factors.

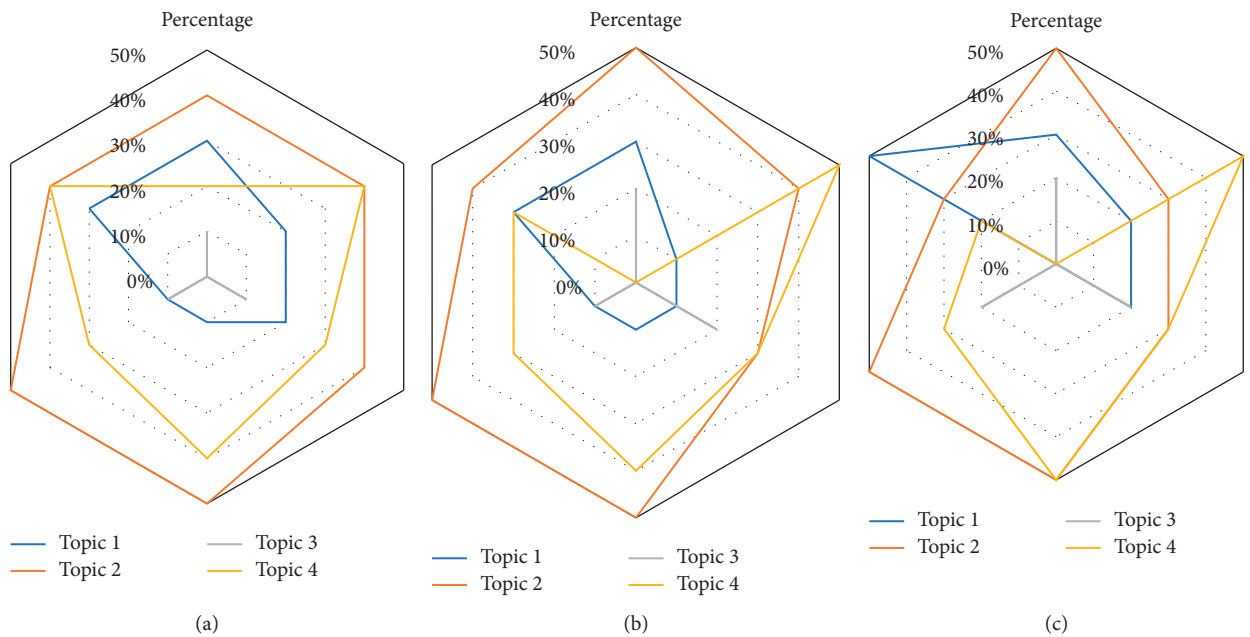


FIGURE 10: Radar map of percentages for each topic in different transformation factors. (a) $h = 0.05$. (b) $h = 0.1$. (c) $h = 0.15$.

TABLE 6: Results of the three methods.

Methods	Target features
Bi et al. [12]	Software usage ecology, speaker, bring two into view, supper device, fingerprint
DPFR	System fluency, system fluency, battery, color, write, camera, stylus, notebook, keyboard, screen
DPFR without considering the production constraints	Software usage ecology, system fluency, CPU, multi-screen collaboration, supper device, bring two into view, notebook, huawei share, screen refresh rate, document
	Software usage ecology, system fluency, CPU, multi-screen collaboration, supper device, bring two into view, notebook, camera, stylus, keyboard

preference of consumers but also the product constraints for the redesign.

5. Discussion and Conclusions

In this study, we propose a novel model (DPFR) to identify the potential features for product redesign from both the customers' and the manufacturers' perspectives. First, we used semantic analysis to build feature engineering and to create machine learning models to identify helpful reviews. The results showed that many useless reviews were removed indicating that identifying useful reviews is an important step, which can reduce the redundancy of useless reviews and improve the computational efficiency to discover product features for the redesign. Second, a structured model was constructed based on helpful reviews. In this step, LDA was used to identify latent *product topics* that customers discussed in helpful reviews and to quantify the importance of each topic. Moreover, we proposed a new measure (PSA) to select the optimal number of topics. The results showed that PSA performed better than the perplexity algorithm. We introduced an index called *feature attention* to measure the priority of all potential product features for the redesign. The local and global preferences of each feature are considered in *feature attention*. Lastly, a target features selection model was constructed by combining the various perspectives of both the customer and the manufacturer to select potential product features from all the features, considering engineering cost, environmental protection cost, redesign lead time, and carbon emission. The results from this study validate the effectiveness of DPFR when the focus is only on customer needs. Moreover, including the production constraints, DPFR can balance customer needs and the manufacturing constraints (such as redesign cost, redesign time, and carbon emissions) and can provide guidelines for practitioners to discover product features for the redesign.

5.1. Research Contributions. This study contributes to academia from three different perspectives.

First, this study verifies that identifying helpful reviews can help manufacturers make strategic decisions regarding product improvement. Although most previous studies have also focused on identifying helpful reviews [17, 29, 40], they have not provided a fine-grained analysis of product improvement and redesign.

Second, in this study, *product topics* for customers were generated from helpful reviews, and the influence of each topic was quantified. Previous studies have mainly focused

on the polarity detection of product features [12, 14] and neglected the effect of different topics. In our study, a structured model was constructed to identify the topics and calculate the importance of the product features of each topic for product improvement.

Third, this study combines customer satisfaction with the manufacturing constraints of product design. Previous studies mainly focused only on customer satisfaction [12, 19] and did not investigate a method for aligning the customers' preferences, with manufacturing design. Moreover, our study proposes a feature selection model that considers economic and environmental factors to select to-be-improved features for product redesign.

5.2. Practical Implications. This study makes several practical and industrial contributions.

First, we used a real dataset of a well-known short-video platform in China to test and validate the proposed method. The results show that our method can effectively identify potential features for product redesign.

Second, our method is feasible for product redesign because it considers both customer satisfaction and the manufacturing constraints of product redesign. The selected potential features are usually product attributes that both customers and manufacturers are most concerned about. Therefore, decision makers can further understand the strengths and weaknesses of the product and make strategic decisions about product improvements.

Third, manufacturers can use our method to develop an automatic software system connected to online platforms to constantly analyze real-time data from customers. Given the emergence of new media and globalization, it is vital for companies to rapidly track and deal with customer needs at an early stage to retain their competitive position in the market. Under these circumstances, our method will be an efficient tool.

5.3. Limitations and Future Work. This study has several limitations. First, the model for identifying helpful reviews requires a significant amount of data to be manually labeled. Future work needs to consider unsupervised learning techniques to reduce the manual labor required. Second, this study considers only economic and environmental constraints in the simplest form to construct an optimal configuration model for product features. However, in the actual production process of a company, the constraints are more complicated and interrelated. Therefore, future work should consider more complex constraints, such as combining

economic constraints, environmental constraints, and corporate social responsibility constraints. Third, in the whole life cycle of product redesign, continuous learning techniques can be applied to learn and update knowledge constantly during product improvement [78, 79]. Furthermore, although this study was applied to one example product, it has the potential to be applied in many domains in future work.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Xiaolong Wang performed conceptualization, data curation, formal analysis, methodology, writing of the original draft, reviewing, and editing. Yu Wang was responsible for project administration, supervision, resources, and funding acquisition. Jiacong Wu and Sara Shafiee carried out review and editing.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 7217020185 and 71772075) and National Foreign Expert Program of the Ministry of Science and Technology of China (Grant No. G2021199032L).

References

- [1] Y. Dong, R. Tan, P. Zhang, Q. Peng, and P. Shao, "Product redesign using functional backtrack with digital twin," *Advanced Engineering Informatics*, vol. 49, 2021.
- [2] B. Ma, H. Ma, Y. Li, X. Chu, and Q. Liu, "Design improvement supporting method based on product performance analysis," *International Journal of Computer Integrated Manufacturing*, vol. 32, no. 10, pp. 1009–1022, 2019.
- [3] M. C. Cohen, "Big data and service operations," *Production and Operations Management*, vol. 27, no. 9, pp. 1709–1723, 2018.
- [4] B. Jeong, J. Yoon, and J. M. Lee, "Social media mining for product planning: a product opportunity mining approach based on topic modeling and sentiment analysis," *International Journal of Information Management*, vol. 48, pp. 280–290, 2019.
- [5] J. Qi, Z. Zhang, S. Jeon, and Y. Zhou, "Mining customer requirements from online reviews: a product improvement perspective," *Information & Management*, vol. 53, no. 8, pp. 951–963, 2016.
- [6] Y. Guo, S. J. Barnes, and Q. Jia, "Mining meaning from online ratings and reviews: tourist satisfaction analysis using latent dirichlet allocation," *Tourism Management*, vol. 59, pp. 467–483, 2017.
- [7] B. Yang, Y. Liu, Y. Liang, and M. Tang, "Exploiting user experience from online customer reviews for product design," *International Journal of Information Management*, vol. 46, pp. 173–186, 2019.
- [8] T. Hou, B. Yannou, Y. Leroy, and E. Poirson, "Mining changes in user expectation over time from online reviews," *Journal of Mechanical Design*, vol. 141, no. 9, 2019.
- [9] X. Liang, J. Guo, Y. Sun, and X. X. Liu, "A Method of Product Selection Based on Online Reviews," *Mobile Information Systems*, vol. 2021, Article ID 9656315, 16 pages, 2021.
- [10] E. Maslowska, C. M. Segijn, K. A. Vakeel, and V. Viswanathan, "How consumers attend to online reviews: an eye-tracking and network analysis approach," *International Journal of Advertising*, vol. 39, no. 2, pp. 282–306, 2019.
- [11] D. Yuan, "Intelligent innovative knowledge management integration method based on user generated content," *Cluster Computing*, vol. 22, no. S2, pp. 4793–4803, 2018.
- [12] J.-W. Bi, Y. Liu, Z.-P. Fan, and E. Cambria, "Modelling customer satisfaction from online reviews using ensemble neural network and effect-based Kano model," *International Journal of Production Research*, vol. 57, no. 22, pp. 7068–7088, 2019.
- [13] N. F. Ibrahim and X. J. Wang, "A text analytics approach for online retailing service improvement: evidence from Twitter," *Decision Support Systems*, vol. 121, pp. 37–50, 2019.
- [14] J. Joungh and H. M. Kim, "Explainable neural network-based approach to Kano categorisation of product features from online reviews," *International Journal of Production Research*, pp. 1–21, 2021.
- [15] S. Lee, S. Park, and M. Kwak, "Revealing the dual importance and Kano type of attributes through customer review analytics," *Advanced Engineering Informatics*, vol. 51, 2022.
- [16] J. Wu, Y. Wang, R. Zhang, and J. Cai, "An approach to discovering product/service improvement priorities: using dynamic importance-performance analysis," *Sustainability*, vol. 10, no. 10, 2018.
- [17] M. Akbarabadi and M. Hosseini, "Predicting the helpfulness of online customer reviews: the role of title features," *International Journal of Market Research*, vol. 62, no. 3, pp. 272–287, 2018.
- [18] C. Jiang, Y. Liu, Y. Ding, K. Liang, and R. Duan, "Capturing helpful reviews from social media for product quality improvement: a multi-class classification approach," *International Journal of Production Research*, vol. 55, no. 12, pp. 3528–3541, 2017.
- [19] M. Tang, J. Jin, Y. Liu, C. Li, and W. Zhang, "Integrating topic, sentiment, and syntax for modeling online reviews: a topic model approach," *Journal of Computing and Information Science in Engineering*, vol. 19, no. 1, 2019.
- [20] S. Ansari and S. Gupta, "Customer perception of the deceptiveness of online product reviews: a speech act theory perspective," *International Journal of Information Management*, vol. 57, 2021.
- [21] K. Y. Chan, C. K. Kwong, and G. E. Kremer, "Predicting Customer Satisfaction Based on Online Reviews and Hybrid Ensemble Genetic Programming Algorithms," *Engineering Applications of Artificial Intelligence*, vol. 95, 2020.
- [22] A. Dash, D. S. Zhang, and L. N. Zhou, "Personalized ranking of online reviews based on consumer preferences in product features," *International Journal of Electronic Commerce*, vol. 25, no. 1, pp. 29–50, 2021.
- [23] Y. Meng, N. H. Yang, Z. L. Qian, and G. Y. Zhang, "What makes an online review more helpful: an interpretation framework using XGBoost and SHAP values," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 16, no. 3, pp. 466–490, 2021.

- [24] L. Martin and P. Pu, "Prediction of helpful reviews using emotions extraction," in *Proceedings of the Twenty-Eighth Aaai Conference on Artificial Intelligence*, pp. 1551–1557, Québec, Canada, July 2014.
- [25] S. P. Eslami, M. Ghasemaghaei, and K. Hassanein, "Which online reviews do consumers find most helpful? A multi-method investigation," *Decision Support Systems*, vol. 113, pp. 32–42, 2018.
- [26] L. Luo, S. Duan, S. Shang, and Y. Pan, "What makes a helpful online review? Empirical evidence on the effects of review and reviewer characteristics," *Online Information Review*, vol. 45, no. 3, pp. 614–632, 2021.
- [27] S. M. Kim, P. Pantel, T. Chklovski, and M. Pennacchiotti, "Automatically Assessing Review Helpfulness," in *Proceedings of the 2006 Conference On Empirical Methods In Natural Language Processing(EMNLP)*, pp. 423–430, Sydney, Australia, January 2006.
- [28] S. R. Zeenia Singla and J. Sushma, "Statistical and sentiment analysis of consumer product reviews. 2017 8th International Conference on Computing," in *Proceedings of the 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Delhi, India, July 2017.
- [29] B. Lutz, N. Pröllochs, and D. Neumann, "Are longer reviews always more helpful? Disentangling the interplay between review length and line of argumentation," *Journal of Business Research*, vol. 144, pp. 888–901, 2022.
- [30] Q. Cao, W. J. Duan, and Q. W. Gan, "Exploring determinants of voting for the "helpfulness" of online user reviews: a text mining approach," *Decision Support Systems*, vol. 50, no. 2, pp. 511–521, 2011.
- [31] A. Y. K. Chua and S. Banerjee, "Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality," *Computers in Human Behavior*, vol. 54, pp. 547–554, 2016.
- [32] Y. Liu, J. Jin, P. Ji, J. A. Harding, and R. Y. K. Fung, "Identifying helpful online reviews: A product designer's perspective," *Computer-Aided Design*, vol. 45, no. 2, pp. 180–194, 2013.
- [33] A. I. S. K. Oumayma Oueslati and O. Habib, "Sentiment analysis for helpful reviews prediction," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 7, no. 3, pp. 34–40, 2018.
- [34] Z. Zeng, Z. Zhou, and X. Mu, "User review helpfulness assessment based on sentiment analysis," *The Electronic Library*, vol. 38, no. 2, pp. 337–351, 2020.
- [35] A. Y. K. Chua and X. Y. Chen, "Online "helpful" lies: an empirical study of helpfulness in fake and authentic online reviews," *Information for a Better World: Shaping the Global Future*, vol. Pt I 13192, pp. 91–99, 2022.
- [36] K. Kaushik, R. Mishra, N. P. Rana, and Y. K. Dwivedi, "Exploring reviews and review sequences on e-commerce platform: A study of helpful reviews on Amazon.in," *Journal of Retailing and Consumer Services*, vol. 45, pp. 21–32, 2018.
- [37] S.-T. Li, T.-T. Pham, and H.-C. Chuang, "Do reviewers' words affect predicting their helpfulness ratings? Locating helpful reviewers by linguistics styles," *Information & Management*, vol. 56, no. 1, pp. 28–38, 2019.
- [38] J. Kikuchi and V. Klyuev, "Gathering User Reviews for an Opinion Dictionary," in *Proceedings of the 2016 18th International Conference On Advanced Communications Technology (Icact) - Information And Communications For Safe And Secure Life*, pp. 566–569, PyeongChang, Korea (South), January 2016.
- [39] S. Krishnamoorthy, "Linguistic features for review helpfulness prediction," *Expert Systems with Applications*, vol. 42, no. 7, pp. 3751–3759, 2015.
- [40] J. Du, J. Rong, S. Michalska, H. Wang, and Y. Zhang, "Feature selection for helpfulness prediction of online product reviews: an empirical study," *PLoS One*, vol. 14, no. 12, Article ID e0226902, 2019.
- [41] G. Shan, L. Zhou, and D. Zhang, "From conflicts and confusion to doubts: examining review inconsistency for fake review detection," *Decision Support Systems*, vol. 144, 2021.
- [42] S. Chatterjee, "Drivers of helpfulness of online hotel reviews: a sentiment and emotion mining approach," *International Journal of Hospitality Management*, vol. 85, 2020.
- [43] J. P. Singh, S. Irani, N. P. Rana, Y. K. Dwivedi, S. Saumya, and P. K. Kumar Roy, "Predicting the "helpfulness" of online consumer reviews," *Journal of Business Research*, vol. 70, pp. 346–355, 2017.
- [44] S. S. Htay and K. T. Lynn, "Extracting product features and opinion words using pattern knowledge in customer reviews," *The Scientific World Journal*, vol. 2013, Article ID 394758, 5 pages, 2013.
- [45] S.-G. Lee, S. Trimi, and C.-G. Yang, "Perceived usefulness factors of online reviews: a study of Amazon.com," *Journal of Computer Information Systems*, vol. 58, no. 4, pp. 344–352, 2017.
- [46] G. R. Semin and K. Fiedler, "The linguistic category model, its bases, applications and range," *European Review of Social Psychology*, vol. 2, no. 1, pp. 1–30, 1991.
- [47] H. Jelodar, Y. Wang, C. Yuan et al., "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey," *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15169–15211, 2018.
- [48] Y. L. Zhang and L. L. Zhang, "Movie Recommendation Algorithm Based on Sentiment Analysis and LDA," *Procedia Computer Science*, vol. 199, pp. 871–878, 2022.
- [49] N. Korfiatis, P. Stamolampros, P. Kourouthanassis, and V. Sagiadinos, "Measuring service quality from unstructured data: A topic modeling application on airline passengers' online reviews," *Expert Systems with Applications*, vol. 116, pp. 472–486, 2019.
- [50] A. Poushneh and R. Rajabi, "Can reviews predict reviewers' numerical ratings? The underlying mechanisms of customers' decisions to rate products using Latent Dirichlet Allocation (LDA)," *Journal of Consumer Marketing*, vol. 39, no. 2, pp. 230–241, 2022.
- [51] T. Schmiedel, O. Müller, and J. vom Brocke, "Topic modeling as a strategy of inquiry in organizational research: a tutorial with an application example on organizational culture," *Organizational Research Methods*, vol. 22, no. 4, pp. 941–968, 2019.
- [52] M. Nilashi, A. Mardani, H. C. Liao, H. Ahmadi, A. A. Manaf, and W. Almukadi, "A hybrid method with TOPSIS and machine learning techniques for sustainable development of green hotels considering online reviews," *Sustainability*, vol. 11, no. 21, 2019.
- [53] N. Ko, B. Jeong, S. Choi, and J. Yoon, "Identifying product opportunities using social media mining: application of topic modeling and chance discovery theory," *IEEE Access*, vol. 6, pp. 1680–1693, 2018.
- [54] Y. Yu, B.-H. Nguyen, F. Yu, and V.-N. Huynh, "Discovering topics of interest on Steam community using an LDA approach," in *Advances in the Human Side of Service Engineering*, pp. 510–517, Cham, Switzerland, Article ID Springer, Cham, 2021.

- [55] T. L. James, E. D. V. Villacis Calderon, and D. F. Cook, "Exploring patient perceptions of healthcare service quality through analysis of unstructured feedback," *Expert Systems with Applications*, vol. 71, pp. 479–492, 2017.
- [56] B. E. Bürdek, *Design: History, Theory and Practice of Product Design*, Walter de Gruyter, Berlin, Germany, 2005.
- [57] L. Zhang, X. Chu, and D. Xue, "Identification of the to-be-improved product features based on online reviews for product redesign," *International Journal of Production Research*, vol. 57, no. 8, pp. 2464–2479, 2018.
- [58] Z. Kosmadoudi, T. Lim, J. Ritchie, S. Louchart, Y. Liu, and R. Sung, "Engineering design using game-enhanced CAD: the potential to augment the user experience with game elements," *Computer-Aided Design*, vol. 45, no. 3, pp. 777–795, 2013.
- [59] E. L.-C. Law, P. van Schaik, and V. Roto, "Attitudes towards user experience (UX) measurement," *International Journal of Human-Computer Studies*, vol. 72, no. 6, pp. 526–541, 2014.
- [60] J. Park, S. H. Han, H. K. Kim, S. Oh, and H. Moon, "Modeling user experience: a case study on a mobile device," *International Journal of Industrial Ergonomics*, vol. 43, no. 2, pp. 187–196, 2013.
- [61] B. Sun, C. Ao, J. Wang, B. Mao, and L. Xu, "Listen to the voices from tourists: evaluation of wetland ecotourism satisfaction using an online reviews mining approach," *Wetlands*, vol. 40, no. 5, pp. 1379–1393, 2020.
- [62] H. K. Chan, E. Lacka, R. W. Y. Yee, and M. K. Lim, "The role of social media data in operations and production management," *International Journal of Production Research*, vol. 55, no. 17, pp. 5027–5036, 2015.
- [63] F. Zhou, J. Ayoub, Q. Xu, and X. Jessie Yang, "A machine learning approach to customer needs analysis for product ecosystems," *Journal of Mechanical Design*, vol. 142, no. 1, 2020.
- [64] J. Jin, D. Jia, and K. Chen, "Mining online reviews with a Kansei-integrated Kano model for innovative product design," *International Journal of Production Research*, pp. 1–20, 2021.
- [65] J. Mansuy, S. Verlinde, and C. Macharis, "Understanding preferences for EEE collection services: a choice-based conjoint analysis," *Resources, Conservation and Recycling*, vol. 161, 2020.
- [66] M. R. Martínez-Torres, "Content analysis of open innovation communities using latent semantic indexing," *Technology Analysis & Strategic Management*, vol. 27, no. 7, pp. 859–875, 2015.
- [67] M. D. Bovea and B. Wang, "Redesign methodology for developing environmentally conscious products," *International Journal of Production Research*, vol. 45, no. 18–19, pp. 4057–4072, 2007.
- [68] R. Hammami, I. Nouira, and Y. Frein, "Carbon emissions in a multi-echelon production-inventory model with lead time constraints," *International Journal of Production Economics*, vol. 164, pp. 292–307, 2015.
- [69] K. Meng, P. Lou, X. Peng, and V. Prybutok, "Multi-objective optimization decision-making of quality dependent product recovery for sustainability," *International Journal of Production Economics*, vol. 188, pp. 72–85, 2017.
- [70] H. Tavakoli and B. D. Barkdoll, "Sustainability-based optimization algorithm," *International journal of Environmental Science and Technology*, vol. 17, no. 3, pp. 1537–1550, 2019.
- [71] Q. Ding and H. Zhao, "Study on e-commerce logistics cost control methods in the context of COVID-19 prevention and control," *Soft Computing*, vol. 25, no. 18, pp. 11955–11963, 2021.
- [72] D. Huisingh, Z. Zhang, J. C. Moore, Q. Qiao, and Q. Li, "Recent advances in carbon emissions reduction: policies, technologies, monitoring, assessment and modeling," *Journal of Cleaner Production*, vol. 103, pp. 1–12, 2015.
- [73] J. Ma, J. D. Harstvedt, D. Dunaway, L. Bian, and R. Jaradat, "An exploratory investigation of additively manufactured product life cycle sustainability assessment," *Journal of Cleaner Production*, vol. 192, pp. 55–70, 2018.
- [74] X. L. Zheng, S. Zhu, and Z. X. Lin, "Capturing the essence of word-of-mouth for social commerce: assessing the quality of online e-commerce reviews by a semi-supervised approach," *Decision Support Systems*, vol. 56, pp. 211–222, 2013.
- [75] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *The Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.
- [76] J. Gan and Y. Qi, "Selection of the optimal number of topics for LDA topic model-taking patent policy analysis as an example," *Entropy (Basel, Switzerland)*, vol. 23, no. 10, 2021.
- [77] R. Rai, "Identifying key product attributes and their importance levels from online customer reviews," in *Proceedings of the 2012 ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE2012)*, Chicago, IL, USA, August 2012.
- [78] F. S. Naqushbandi and A. John, "Sequence of actions recognition using continual learning," in *Proceedings of the 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, pp. 858–863, Coimbatore, India, February 2022.
- [79] S. Y. Zhu, J. Qi, J. Hu, and H. Q. Huang, "Intelligent product redesign strategy with ontology-based fine-grained sentiment analysis," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 35, no. 3, pp. 295–315, 2021.