Research and Application of Product Design User Requirements Mining Based on Online Comments and Kano Model

Zhiyuan Wang
Mechanical Engineering Institute, Anhui University of Science and Technology,
Huainan City 232001, China
E-mail: xinyuan8610@126.com

Keywords: industrial design, user requirements, online reviews, the kano model, fire rescue drones

Received: June 1, 2023

To solve the problems of strong subjectivity, less data and poor real-time performance of user demand mining in the process of product design, this paper proposes a method of user demand mining based on online reviews and Kano model. Firstly, Octopus is used to crawl the user's online comment data, and the text data is preprocessed by jieba to form a comment dataset, and then the data is classified and labeled by word segmentation. Secondly, the Apriori algorithm is used to extract the product attributes that users pay attention to frequently, and the SO-PMI is used to calculate the product attribute evaluation value. Finally, with the help of KANO model, the user demand classification of product attributes is carried out, and the priority ranking of product attributes is obtained, and then the direction of product optimization and improvement is proposed. A fire rescue drones product is taken as an example to verify the effectiveness of the proposed method. The results show that this method can provide valuable information about user requirements, improve user satisfaction, and promote the successful development of product design.

Povzetek: Članek predlaga metodo rudarjenja uporabniških zahtev z uporabo spletnih mnenj in Kano modela za izboljšanje oblikovanja izdelkov, preverjeno na reševalnih dronih.

1 Introduction

The process of product design quality evaluation that how to better meet the diversified requirements of customers. Besides, it is one of the important means for enterprises to improve their competitiveness to convene the personalized and diversify needs of user and improve their satisfaction in the highly competitive global market. A deep insight and understanding of the needs of the target user groups is helpful to the successful development of products. [1] Traditional user requirements mining methods are hard to meet the personalized and diversified needs of users due to the strong subjectivity, less data, and poor real-time performance. By contrast, online review data has become a new source of mining user requirements because the user data obtained is more reliable. [2-3] Online reviews, as a link between products and users, imply valuable feedback information from users, truly reflect users' feelings after using the products, and are characterized by large amount of data, rich content and strong real-time performance. Thus, the current domestic and foreign scholars focus on how to get useful information. [4] Jin et al. [5] extracted user opinion information from online product reviews based on clustering method. Qi et al. [6] established a filtering model for online reviews and analyzed users' opinions through Cano model. To gather a lot of data, do in-depth surveys, interviews, and market research. To identify both explicit and implicit needs, make use of a variety of data sources, such as reviews, consumer input, and support interactions. Wang Keqin et al. [7] proposed an importance performance competitor analysis method based on online reviews. Yang Cheng [8] proposed a research method of product design improvement based on review big data. Wang Ying [9] divided the software requirement categories contained in APP user comment data, and constructed a comment requirement mining data set, which was subjected to model training and cross-validation. Jia Danping et al. [10] proposed a user demand mining method for semi-automatic construction of perceptual emotion dictionary based on kansei engineering perspective. Yang Huan [11] constructed the big data analysis in the field of Internet design "from creating user portraits to analyzing user experience journey and then combining scenario analysis", and derived the innovative path of user demand insight. Li Shaobo et al. [12] proposed a data-driven kansei engineering method for product online review, which provided a new direction for obtaining kansei words and kansei evaluation. Li Xiang et al. [13] put forward systematic research on user requirements in physical product design by combining SPSS analysis and online review mining method. Ji Xue et al. [14] proposed a review mining and requirement acquisition method considering product attribute hierarchy. Wang Jun et al. [15] suggested a product design method based on user online reviews and combined with fuzzy Kano-TOPSIS method.

It is rare to combine online review with Kano model and apply it to the research of product design field, and focus on the application of data mining technology, text semantic analysis technology, machine learning method and statistical method to the obtained online review data for subject attribute identification, emotion polarity analysis and subject-emotion identification, etc. Moreover, the above study has not clearly illustrated the relationship between the positive and negative tendencies of online reviews and user requirements, nor has it analyzed the categories of user requirements represented by review information from a psychological perspective. The thematic features of positive emotions in online review data reflect the satisfaction degree and demand orientation of users to the products. Some aspects of user-satisfied products will increase the loyalty of users and thus generate word-of-mouth effect on the one hand, and also contain the charm demand and expectation demand of users on the other hand, which is where the products need to be optimized. Negative reviews reflect users' unsatisfied demands for products and contain users' expected demands. Dissatisfied aspects of users will adversely affect the purchase intention of potential consumers, which is also the direction of product improvement. Based on this, this paper proposes a user requirement mining method based on online reviews and Kano model.

Methods	Drawbacks
Intuitionistic Fuzzy Kano Model (IFKM)[29]	The model becomes more complicated due to intuitionistic fuzzy sets. Some users, particularly those unfamiliar with fuzzy logic and associated topics, may find it difficult to comprehend and use intuitionistic fuzzy sets.
Ensemble Neural Network Based Model [30]	Ensemble models increase the system's total complexity. It might be difficult to coordinate the outputs of many models, manage them, and choose the best ensemble approach.
Classification of customer Requirements (CRs)[31]	When particular client demands are categorized into generic categories, it might result in overgeneralization. This may cause significant subtleties and details to be lost. Overgeneralization can end in the creation of goods that fall short of completely satisfying the specific requirements of certain clientele groups.
Fuzzy Analytic Hierarchy Process (FAHP)[32]	Compared to regular AHP, FAHP is less transparent due to its fuzzy character. Users unfamiliar with fuzzy logic ideas may find it difficult to comprehend and analyze the outcomes of a fuzzy decision-making process.

2 Theoretical basis

2.1 Online review

Online reviews, first proposed by Chatteije[16] in 2001, are comments about products shared on a virtual online community platform. With the development of the Internet era, the scope of online reviews is not only about some views and opinions of the product, but also gradually extends to the evaluation of the product design by users and the experience feedback in the whole process. Online reviews contain evaluation information of multiple attributes of goods, and users may have their own preference expression on each attribute. Therefore, it is necessary to mine the needs of the attributes expressed by users, and analyze the emotional tendencies of users on different attributes in order to accurately locate the user requirements. [16] Data mining based on online reviews can fully represent the real user requirements, which has important reference significance for improving and optimizing products and enhancing market competitiveness.

2.2 Kano model

Kano [18] model is a user model proposed by Noriaki Kano, a famous professor at Tokyo University of Technology, in 1984 under the inspiration of Frederick

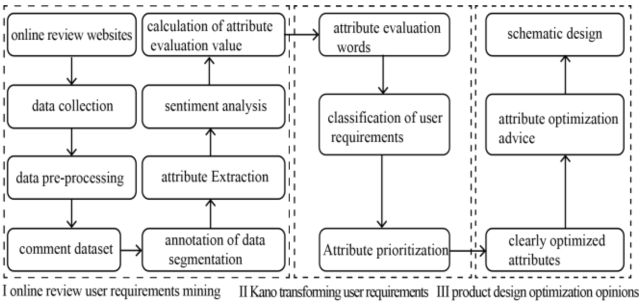
Herzberg's two-factor theory. It can classify and sort customer requirements, find the nonlinear relationship between product performance and customer satisfaction, and then determine the priority of product attributes. According to the nonlinear relationship between product attributes and customer satisfaction, Kano model divides user requirements into five categories: Must-be Quality, One-dimensional Quality, Attractive Quality, Indifferent Quality and Reverse Quality. [16]

3 Methodologies

A method of user requirements mining based on online review and Kano model was proposed, which was divided into three parts: online review user requirements mining, Kano transforming user requirements and product design optimization opinions. The method flow is shown in Fig. 1.

1. Online review user requirements mining. An online review website was selected for data collection and preprocessing to form a review data set, word segmentation in the data was classified and labeled, product attributes were extracted for sentiment analysis, and product attribute evaluation values were calculated.
2. Kano transforming user requirements. Based on the product attribute evaluation value, the user requirements were classified by Kano model to get the attribute priority.
3. Product design optimization opinions. The attributes that need to be optimized can be known according to the above classification of product users' requirements and the prioritization of attributes, which can be used as the goal of product design optimization to design the scheme. Detailed steps are introduced in the following sections.

Figure 1: Method flow chart



3.1 Mining user requirements for online reviews

3.1.1 Data acquisition and preprocessing

Nowadays, e-commerce, Weibo, blogs, BBS, forums and other websites provide users with platforms for online reviews. When selecting online review websites, it is necessary to select a representative website or multiple websites for data collection according to the specific situation of products. For example, a determined keyword is firstly selected through an in-station search function in the website to search for online reviews, then a network information acquisition tool is used to acquire search results, and the acquired content of each record comprises a user name, review content and response number, and finally all data are exported to an Excel file for storage. In order to ensure that valuable data can be collected, some useless data need to be cleaned out, mainly including information unrelated to the subject and repeated reviews of a user, so that the preprocessed data can form a data comment set that matches the search keywords.

3.1.2 Annotation of data segmentation

As the user data of online reviews are colloquial and ruleless unstructured text data, which cannot be directly analyzed by in-depth online mining. Therefore, further Chinese word segmentation processing is needed for the review set data. The jieba library in python is selected to segment the acquired data. After word segmentation processing of review data results, a word frequency table is generated and marked by python according to the segmented data. The word frequency and part-of-speech table are obtained by sorting the occurrence times of all word segmentation data.

3.1.3 Extraction of product attributes

It is necessary to filter out irrelevant information by manually screening and cleaning whether the product itself is relevant or not mainly according to the information conveyed by the word when constructing product attribute dictionary according to the obtained word frequency to improve the accuracy of data mining. The cleaned data is calculated by Apriori [20] algorithm to get the frequent item sets and extract more accurate user high-frequency attention attributes. Apriori algorithm is a commonly used method in data mining to search for related itemsets from a large number of data sets.[21] The support SUPP(X) of event (or attribute) X can be obtained by Formula (1).

$$\begin{cases} SUPP(X) = \dfrac{N(X)}{N} \\ N(X) = \displaystyle\sum_{i=1}^{N} I_i(X) \end{cases} \tag{1}$$

Where,

N= the total length of the data set;

X=the event (or attribute);

SUPP (X) =the support degree of X;

and $I_i(X)$ =the indicative function of the occurrence of X in the i-th item.

If SUPP (X) reaches the preset support threshold, the itemset {X} is a frequent itemset, i.e., the product attribute that the user pays high-frequency attention to.

3.1.4 Sentiment analysis

According to the product attributes extracted in the previous step, the emotion analysis method is applied to get the emotion tendency of each attribute, so as to better identify the views of users. There have been many studies related to sentiment analysis, which are mainly divided into three types according to the calculations used: emotion dictionary-based approach[16], ML-based approach[16] and rule-based approach [16]. Most of the online reviews involved in this paper are related to the product field, so sentiment dictionary-based approach is chosen for sentiment analysis, i.e., calculating the data using the emotion dictionary to judge its emotional tendency and degree. The construction of the emotion dictionary is based on the automatic way of the polar dictionary, which judges the relevance between the polar words and the seed words in the dictionary by a certain algorithm, and calculates the intensity of the polar words. HowNet[16] sentiment word dictionary is a commonly used polarity dictionary, which is selected in this paper to analyze polarity words. When calculating the intensity of polarity words, a polarity dictionary can be established based on the types of polarity words and their respective emotional tendency levels, as shown in Table 1.

Table 1: Classification of polarity dictionary

Emotional tendency	Meaning	Score
Lexicon of commendatory w ords	Expectation	0.5
	Mightiness	1
Lexicon of derogatory terms	Boredom	-0.5
	Badness	-1
Lexicon of neutral words	Believe	0

3.1.5 Calculation of evaluation value of each product attributes

The product attribute evaluation value is calculated by analyzing the sentences in which the attribute words appear, and comprehensively analyzing the polarity of the emotion words in the sentences. In this paper, SO-PMI [16] algorithm was use to calculate, because it will comprehensively consider the interaction of the three parts of speech of polarity words, adverbs and negative words when expressing the meaning of a sentence when calculating the polarity intensity of a sentence, and gives a calculation Formula to quantify the polarity degree of the whole sentence. The degree of polarity refers to the direction and intensity of the reviewers' emotions. Polarity words (adjectives, verbs, and nouns), as the core part of the sentence, contain most of the information of the whole comment sentence. Adverbs and negative words are the components that further modify the polarity of the whole comment sentence on the basis of the polarity words. The flow of SO-PMI algorithm was divided into three steps: calculation of word polarity intensity, calculation of sentence polarity intensity and calculation of attribute evaluation value.

（1）Calculation of word polarity intensity. Pointwise mutual information (PMI) reflected the relationship between the two words, where PMI = 0 indicated that they were not related, while PMI > 0 indicated that the two words were related, and the degree of correlation was increased with the increase of PMI value, and PMI < 0 indicated that they were mutually exclusive, as shown in Formula (2).

$$PMI(word_1, word_2) = \mathrm{I}b \frac{P(word_1 word_2)}{P(word_1)P(word_2)} \quad (2)$$

Where P (word1 & word2) refers to the probability of occurrence of word1 and word2 at the same time; P(word1) refers to the probability of word1, and the same is true for P(word2). All adjectives, verbs and nouns are regarded as polarity words, and the polarity intensity of the new polarity word and seed word can be obtained by calculating the PMI of the word, as shown in Formula (3).

$$Score(word) = \sum_{pword \in pwords} PMI(word, pword) - \sum_{pword \in pwords} PMI(word, nword) \quad (3)$$

Where,

Word= a new word;

pword= a commendatory seed word;

and nword =a derogatory seed word.

When Score = 0, the word is neutral; When Score > 0, the word is commendatory; When Score < 0, the word is derogatory.

Calculation of sentence polarity intensity. Judgment of sentence polarity is not only to sum up the strength of all words expressing emotional polarity in a sentence, but also to further analyze the adverbs of degree in a sentence. According to

the classification of degree adverbs in HowNet, they were divided into 4 grades with values of 0.6, 0.8, 1.2 and 1.5, as shown in Table 2.

Table 2: Table of partial degree adverbs

Examples of negative words
Deny, none, not, off, fail, stop, little, few; neither, no, nor, never, reject, without, nowhere, hardly, barely, seldom, rarely

Moreover, negative words in sentences generally express opposite emotional tendencies[16]]. According to the method proposed by Yao Tianfang et al. [16] to invert and halve the intensity of emotional polarity words with negative words, the parameter value of negative words was set to -0.5. The contents of common negative words are exposed in Table 3.

Table 3: Listing of partial negative words

Degree	Values	Examples
Maximum	1.5	The most, super, extreme, max out, incomparable, ultimate and perfect…
High	1.2	Very, more, awful, special, extra, quite, much ....
Medium	0.8	Rather, worthy of, as, also…
Low	0.6	Slight, a little, somewhat, some, insignificant, more or less…

Considering the effects of degree words and negative words, the computational formula of the sentence polarity intensity is shown in Formula (3), where wp_i represents the intensity value of the i-th positive word; wn_j represents the intensity value of the j-th negative polarity word; pt_i and pt_j are the signs of negative words in the sentence. If no negative word appears, the value is 1; if there is a negative word, the value is -0.5. dt_i and dt_j are the signs of degree adverbs. If there is no degree adverb, its value is 1. If there is a degree adverb.

$$S=\frac{\sum_{i=1}^{m}wp_i pt_i dt_i+\sum_{j=1}^{n}wn_j pt_j dt_j}{m+n} \qquad (4)$$

(4) Calculation of evaluation value. After the polarity intensities of all attribute sentences corresponding to different attributes are calculated, the weighted average is the comprehensive evaluation value of the attribute, as shown in Formula (5), in which W_a is the comprehensive evaluation value of attribute a, and S_i is the polarity intensity of the i-th attribute sentence of attribute a.

$$W_a=\frac{\sum_{i=1}^{n}S_i}{n} \qquad (5)$$

3.2 Kano Transforming user requirements

3.2.1 Hierarchical classification of user requirements

After weighting the initial emotional score of the emotional dictionary and the grade score of degree adverbs, the emotional value is within the range of (-2,2), so the Score∈(-2,2). The optimized requirements are classified according to the value of score: Score > 1 indicates that users are satisfied with this attribute, so it is classified as an attractive quality; Score∈(0,1) indicates that the user's satisfaction is not high, the product attribute does not exceed the user's expectation, and the user wishes to get a better experience, so it is classified as the one-dimensional quality, and Score < 0 indicates that the user is not satisfied, so it is classified as the must-be quality.

3.2.2 Priority ranking of attributes

According to the classification of user requirements, the attention degree of each attribute of the product and the proportion of positive and negative evaluations are analyzed, and the priority ranking of product attribute requirements is obtained through formula (6), and then suggestions are put forward for the optimal design of the product. Where priority represents demand priority, a represents attention to product attributes (the ratio of attribute evaluation times to total evaluation times), and neg% represents the proportion of negative evaluations.

$$Priority=|a*neg\%| \qquad (6)$$

4 Case study

Based on the fire rescue drones products as a case for empirical study, JD.COM e-commerce platform was selected as the online review data source, because it is the largest 3C digital online shopping platform in China B2C market, and its high-speed logistics and product quality are recognized by consumers.

4.1 Data acquisition and preprocessing

Firstly, the octopus collector was used to search the product, and two key words- "fire rescue drones" and "firefighting drones" were set to crawl the first five pages of results under each key word, and 240 pieces of product information were finally obtained after removing duplicates. Then the crawled data was imported into Excel for manual processing, mainly eliminating some commodity data with few comments and product positioning different from that studied in this paper. Finally, six products were extracted from the perspectives of two keywords respectively to serve as an online review data crawling source for subsequent products, as shown in Table 4.

Table 4: Selected products to be crawled

S/N	Search keywords	Product name	Product SKU	Product Url	Price	Number of reviewers	Comment Url	Sellers' name
1	Fire rescue drones	DJI Tello educational programming drones	1000001650 29	https://item.jd.com/100000165029.html	999	50000+	https://item.jd.com/100000165029.html#comment	DJI Innovation JD flagship store
2	Fire rescue drones	JJR/C Professional drones flying camera	1000295900 02	https://item.jd.com/100029590002.html	1499	5000+	https://item.jd.com/100029590002.html#comment	Optimal JD self-operated Zone
3	Fire rescue drones	JJR/C Professional drones flying camera	1000092776 37	https://item.jd.com/100009277637.html	1499	5000+	https://item.jd.com/100009277637.html#comment	Optimal JD Self-operated Zone
4	Fire rescue drones	Autel EVO II Pro foldable drones for aerial photography	1000206879 36	https://item.jd.com/100020687936.html	11899	200+	https://item.jd.com/100020687936.html#comment	JD self-operated category store for droness
5	Fire rescue drones	Xingkong Smart drones	10042241265798	https://item.jd.com/10042241265798.html	2999	2000+	https://item.jd.com/10042241265798.html#comment	Xingkong smart toy flagship store
6	Fire rescue drones	Autel drones	10040386564737	https://item.jd.com/10040386564737.html	10499	100+	https://item.jd.com/10040386564737.html#comment	Autel smart drones flagship store
7	Firefighting drones	DJI Mavic3 flying drones suit	1000172242 87	https://item.jd.com/100017224287.html	21988	5000+	https://item.jd.com/100017224287.html#comment	DJI Innovation JD flagship store
8	Firefighting drones	HUBSAN	10021039960983	https://item.jd.com/10021039960983.html	4399	200+	https://item.jd.com/10021039960983.html#comment	Masha intelligent equipment specialty store
9	Firefighting drones	DJI drones Yu 3	10048067589898	https://item.jd.com/10048067589898.html	12888	1 万+	https://item.jd.com/10048067589898.html#comment	Penglaike digital specialty store

10	Firefighting drones	Xingkong Smart drones	10042241265793	https://item.jd.com/10042241265793.html	2699	2000+	https://item.jd.com/10042241265793.html#comment	Xingkong smart toy flagship store
11	Firefighting drones	Beast 3 drones	10037811932106	https://item.jd.com/10037811932106.html	1598	5000+	https://item.jd.com/10037811932106.html#comment	Xinguyan toy flagship specialty store
12	Firefighting drones	Xingkong Smart drones	10042241265791	https://item.jd.com/10042241265791.html	2399	2000+	https://item.jd.com/10042241265791.html#comment	Xingkong smart toy flagship store

The user reviews of 12 products were crawled by tools continuously. Under the protection of JD.COM's anti-crawler mechanism, 25 pages of data were obtained for each product, and a total of 2,932 online review data were obtained, including 34 repeated review data, 65 irrelevant review data, and 1,003 defaults, too short and invalid data. Finally, a total of 1,830 valid data were obtained.

4.2 Annotation of data segmentation

(1) Data segmentation

The accurate word segmentation mode in jieba library in python was used to perform batch word segmentation on 1,830 pieces of online review data after preprocessing, and the word segmentation outcome are revealed in Fig. 2.

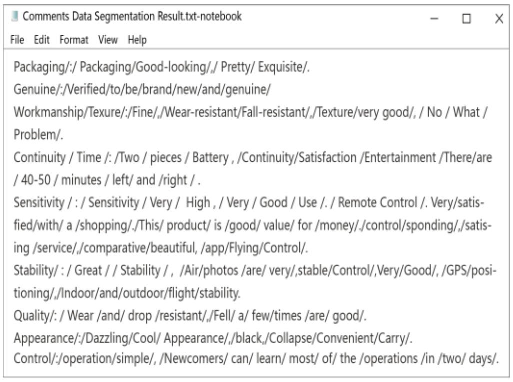
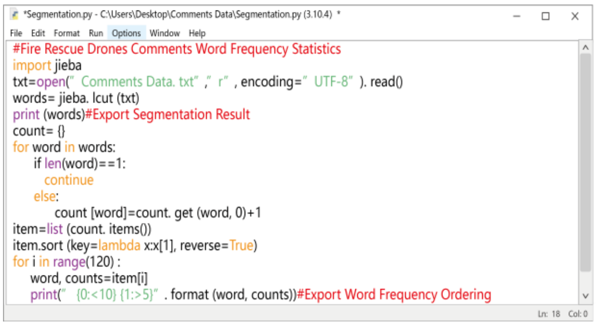

Comments Data Segmentation Result.txt-notebook
File Edit Format View Help
Packaging/:/ Packaging/Good-looking/,/ Pretty/ Exquisite/.
Genuine/:/Verified/to/be/brand/new/and/genuine/
Workmanship/Texure/:/Fine/,/Wear-resistant/Fall-resistant/,/Texture/very good/, / No / What / Problem/.
Continuity / Time /: /Two / pieces / Battery , /Continuity/Satisfaction /Entertainment /There/are / 40-50 / minutes / left/ and /right / .
Sensitivity / : / Sensitivity / Very / High , / Very / Good / Use /. / Remote Control /. Very/satisfied/with/ a /shopping/./This/ product/ is /good/ value/ for /money/./control/sponding/,/satising /service/,/comparative/beautiful, /app/Flying/Control/.
Stability/ : / Great / / Stability / , /Air/photos /are/ very/,stable/Control/,Very/Good/, /GPS/positioning/,/Indoor/and/outdoor/flight/stability.
Quality/: / Wear /and/ drop /resistant/,/Fell/ a/ few/times /are/ good/.
Appearance/:/Dazzling/Cool/ Appearance/,/black,/Collapse/Convenient/Carry/.
Control/:/operation/simple/, /Newcomers/ can/ learn/ most/ of/ the /operations /in /two/ days/.

Figure 2: Word segmentation results

(2) Annotation of word segmentation

After word segmentation of review data, a word frequency table was generated and word frequency was marked using python. Part-of-speech and word frequency lists were obtained by sequencing the occurrence times of all word segmentation data. The main operation process and part of the code are as shown in Fig. 3.


*Segmentation.py - C:\Users\Desktop\Comments Data\Segmentation.py (3.10.4) *
File Edit Format Run Options Window Help

```
#Fire Rescue Drones Comments Word Frequency Statistics
import jieba
txt=open(" Comments Data. txt" ," r" , encoding=" UTF-8" ). read()
words= jieba. lcut (txt)
print (words)#Export Segmentation Result
count= {}
for word in words:
    if len(word)==1:
       continue
    else:
        count [word]=count. get (word, 0)+1
item=list (count. items())
item.sort (key=lambda x:x[1], reverse=True)
for i in range(120) :
   word, counts=item[i]
   print("  {0:<10} {1:>5}" . format (word, counts))#Export Word Frequency Ordering
```

Ln: 18 Col: 0

Figure 3: Acquisition process of word frequency sorting

After that, the synonyms were washed and fused by artificial filtering to obtain a table for top 45 users' high frequency concerned data word frequency (Table 5) and table of the first 42 parts of speech (Table 6).

Table 5: Sorting of word frequency

Attribute	Times	Attribute	Times	Attribute	Times	Attribute	Times
Control	652	Flight	228	Test flight	132	Performance	55
Quality	496	Package	213	Patience	131	Refinement	50
drones	480	Specialty	197	Texture	117	Capacity	49
Endurance	420	Speed	185	Feeling	115	Charging	47
Stability	390	Price	178	Handing	113	Driving	42

						force	
Simplicity	374	Course reversal	173	Design	106	Compactness	41
Photographing	343	Portability	169	Problems	100	Distance	41
Agility	292	Intelligence	159	Effect	92	Foldability	41
Details	286	Logistics	155	Delicacy	79	Remote control	40
Definition	270	Service	151	Manual	70		
Appearance	445	Function	203	Suitability	58		
Battery	220	Video	144	Features	57		

Table 6: Part-of-speech table

S/N	Keywords Times	Keywords Times	Keywords Times	Keywords Times
1	wd 7958	ad 450	cc 164	al 42
2	ude1 2448	wyz 388	new 158	ryv 38
3	vi 2283	vyou 386	vl 152	uyy 38
4	wj 1545	qv 373	ag 105	ws 36
5	vn 1185	rz 356	ww 97	GPS 35
6	ng 1042	qt 351	uzhe 96	nrf 33
7	wm 770	vg 341	wn 93	
8	vshi 700	rzv 315	uls 71	
9	vf 671	mq 314	vd 64	
10	wt 612	an 307	ry 59	
11	rr 603	njtgj 303	uguo 55	
12	ule 595	ns 174	udeng 43	
Remarks	wd comma, ude1 auxiliary word, vi verb (intransitive), wj full stop, vn verb (noun verb), ng noun, wm colon, vshi verb (is), vf verb (directional verb), wt interjection, rr personal pronoun, ule auxiliary word, ad adjective (adjective & adverb), wyz quotation mark (left), vyou verb (have), qv verbal measure words, rz demonstrative pronoun, qt temporary classifiers, vg verbal morpheme, rzv predicate, mq quantifier, njtgi special nound, ns place name, cc coordinate conjunction, new new noun, vl verb (idiom), ag adjective morpheme, ww question mark, uzhe auxiliary word (this), wn unit symbol, uls auxiliary word (speaking), vd auxiliary verb, ry interrogative pronoun, uguo auxiliary word , udeng auxiliary word (etc.), al adjective idiom, ryv predicate interrogative pronoun, uyy auxiliary word (general), ws ellipsis, GPS global positioning system, nrf person name			

4.3 Product attributes extraction

According to the table 7 for top 45 users' high frequency concerned data word frequency, the Apriori algorithm analysis was applied to select the data that had an effective impact on the product, the frequent item set was calculated, and the minimum support=0.1 was set to get more accurate product attributes that the user high-frequency focused on.

Table 7: Product attributes

Attribute	support	confidence	lift	count
Intelligence	0.1028	0.1028	1.0000	159
Portability	0.1114	0.1114	1.0000	169
Price	0.1121	0.1121	1.0000	178
Function	0.1205	0.1205	1.0000	203
Definition	0.1423	0.1423	1.0000	270
Details	0.1563	0.1563	1.0000	286
Agility	0.1691	0.1691	1.0000	292
Stability	0.2131	0.2131	1.0000	390
Appearance	0.2431	0.2431	1.0000	445
Quality	0.2709	0.2709	1.0000	496
Endurance	0.3025	0.3025	1.0000	640
Control	0.3118	0.3118	1.0000	652

4.4 Calculation of attribute evaluation value

After obtaining the product attributes that users paid high attention to, the emotional words in the comment sentences corresponding to the attributes were extracted, and the evaluation values of each attribute of the product were calculated by using the SO-PMI algorithm. All adjectives, verbs and nouns were regarded as polarity words to calculate the correlation degree of the polarity words and the seed words by R software programming using the formulas (2)(3) to obtain the intensity of the polarity words, and then the polarity intensity of the characteristic sentences was calculated according to the formula (4) to obtain the comprehensive evaluation value of each feature by weighted average of the polarity intensities of all the characteristic sentences corresponding to different characteristic words, as shown in Table 8.

Table 8 Comprehensive evaluation values of

Attributes	Evaluation values	Frequency
Intelligence	-0.22	159
Portability	0.16	169
Price	0.13	178
Function	0.21	203
Definition	0.24	270
Details	0.29	286
Agility	0.31	292
Stability	-0.34	390
Appearance	0.56	445
Quality	1.01	496
Endurance	-0.91	640
Control	1.09	652

4.5 Hierarchical classification and prioritization of user requirements

First, according to the evaluation values of each attribute of fire rescue droness obtained in Table 8, hierarchical classification of requirements was performed combined with the Kano model. Control and quality are the attractive quality of the fire rescue drones because their attribute evaluation value is greater than 1. Portable, price, function, definition, details, agility and appearance are the one-dimensional quality of fire rescue drones because their evaluation value is greater than 0 and less than 1. Intelligence, stability and endurance are the must-be quality of fire rescue drones, because their attribute evaluation value is less than 0.

Then, according to the attribute requirement classification of the fire rescue drones, the attention degree of each attribute of the product and the proportion of positive and negative evaluations were analyzed, and the priority order of each attribute requirement of the product was obtained through formula (6), as shown in Table 9.

Table 9: Priority ordering of each attribute requirement

Attribute	Evaluation value	Type of requirement	Attention	Priority
Intelligence	-0.22	Should-be worth	2.28%	3
Portability	0.16	One-dimensional worth	1.56%	5
Price	0.13	One-dimensional worth	2.01%	4
Function	0.21	One-dimensional worth	4.15%	6
Definition	0.24	One-dimensional worth	5.02%	7
Details	0.29	One-dimensional worth	5.42%	8
Agility	0.31	One-dimensional worth	11.07%	9
Stability	-0.34	Must-be quality	13.42%	2
Appearance	0.56	One-dimensional worth	12.01%	10
Quality	1.01	Attractive quality	14.32%	11
Endurance	-0.91	Should-be quality	14.71%	1
Control	1.09	Attractive quality	14.03%	12

4.6 Result analysis

Table 9 shows that the evaluation values of endurance, stability and intelligence of the fire rescue drones are negative, with priority ranking of 1, 2 and 3, indicating that users' evaluation of these two aspects of the product is negative, and their satisfaction with these three attributes is low, which is the attribute that needs to be optimized most. The evaluation values of price, portability, function, definition, details, agility and appearance are positive, indicating that the users' evaluation of these 7 aspects of the product is positive, and the users have high satisfaction with them. The evaluation values of quality and control are positive, with priority ranking of 11 and 12, indicating that the users' evaluation of these two aspects of the product is positive, and the users have the highest satisfaction with them, which are the advantages of the existing product. In view of the principle that product design should be centered on meeting the requirements of users, endurance, stability and intelligence are three attributes that need to be improved urgently in product design.

4.7 Schematic design

The scheme of fire rescue drones is designed according to the above analysis. The user's satisfaction with the endurance, stability and intelligence of the fire rescue drones is low, which are the attributes that need to be optimized most. Therefore, in terms of battery life, a new type of lithium battery should be selected, with high-efficiency battery conversion to ensure long-term continuous battery life. Moreover, the battery module should adopt detachable structural design and be installed in the center of the fuselage to be safe and stable, for convenient replacement. In terms of stability, the flight module with multi-rotor and three-blade propellers should be adopted, and the vortex air-conditioning control system should be installed side by side under the flight propellers to ensure the flight stability and precise control ability of the fire scene, so as to make the flight movement change more stable and precise. In order to ensure the stability during take-off and landing, the necessary buffer support device should be equipped under the fuselage, and the curved flexible material should be used to avoid slight vibration and provide the stability coefficient to meet the needs of use. In terms of intelligence, automatic algorithm control system should be adopted to integrate the information collaboration among various devices and instruments, and the integrated algorithm should be used to control the drones to realize intelligent control. The space detection and rendering instrument and multiple cameras mounted on it are connected in series to render the environmental VR model, and the data are transmitted to the ground controllers, and the simulation flight control is realized through the VR control equipment. At the same time, the intelligent integration algorithm ensures that it can realize independent analysis and intelligent obstacle avoidance flight when encountering sudden extreme environment. While ensuring the effectiveness of data communication, a voice transmission structure shall be arranged on the unmanned aerial vehicle to realize the transmission of the fire scene propaganda considering the communication with the trapped personnel at the fire scene, and at the same time, the striking part of the machine body should be provided with light information to remind the trapped personnel. In the aspect of modeling, the design of flowing lines enhances the visual tension and demonstrates its control and agility attributes .The top of the unmanned aerial vehicle is provided with a binocular vision obstacle avoidance module and a battery pack, the side is provided with an annular prompt and a searchlight, and the bottom is provided with a load platform of a panoramic scanning camera group and a precision instrument, thereby meeting the functional attributes of rescue, patrol and monitoring. The final design effect diagram is shown in Figs. 4, 5, 6, 7, 8 and 9.

Figure 4: Main visual diagram of the product

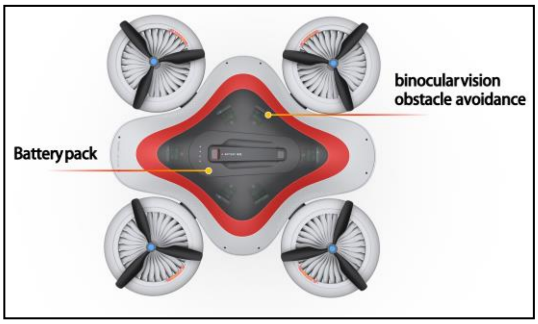


Figure 5: Battery pack and binocular vision obstacle avoidance system

Figure 6: Sensor and gas monitoring device

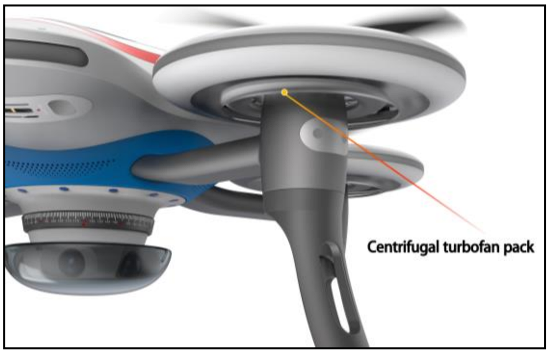


Figure 7: Centrifugal turbofan pack

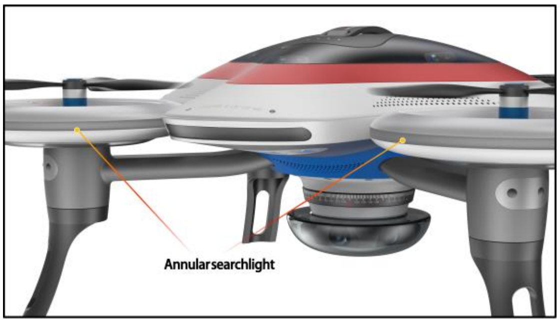


Figure 8: Panoramic scanning camera group

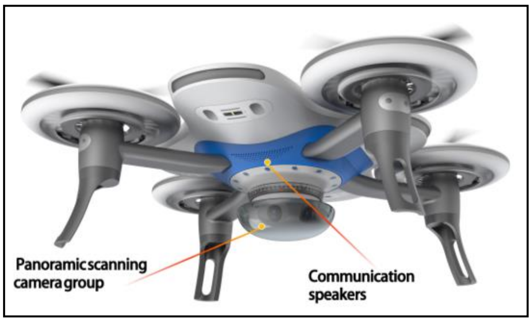


Figure 9: Annular searchlight

5 Discussion

The Fuzzy kano Topsis [15] of the method relies on the assumption that the criteria lack dimensions and are comparable. It is essential to normalize the criteria, and the selected normalization method can impact the results. The method's susceptibility to the chosen normalization approach may pose a disadvantage. SOTA models often perform well on certain benchmark tasks, but they may have trouble generalizing to novel or untested contexts. While they may work very well on carefully chosen datasets, they might not work well in practical settings. SOTA models may give rise to ethical questions, particularly when they deal with sensitive data or choices that have important practical ramifications. It is a never-ending task to make sure these models are applied appropriately and cause no damage to people or communities. Our proposed approach SO-PMI to resolve these problems.

6 Conclusions

Since online reviews reflect users' satisfaction and requirement orientation of products and contain users' attractive and expected requirements, mining based on online review data can help enterprises to obtain valuable product design improvement information. Therefore, a method of user requirement mining based on online reviews and Kano model is proposed. Firstly, users' requirements are mined based on online reviews, and then transformed with the help of KANO model, and finally, the direction of product optimization and improvement is proposed. The method improves the authenticity, timeliness and effectiveness of the user requirement acquisition, promotes the product optimization and improvement target to be more targeted, and improves the satisfaction degree of the user on the product. Furthermore, The proposed method's efficacy is demonstrated through the utilization of a fire rescue drones product as an illustrative example. Nevertheless, certain limitations persist in this paper, particularly in the phases of data collection and preprocessing. Manual screening is used to eliminate useless information to obtain review data sets, and further research on intelligent identification is needed in the future. In addition, it is one-sided to obtain user requirement information from text data. In the follow-up work, the image and text data should be combined to achieve a comprehensive mining of user requirement information.

7 Limitations and future work

The Kano Model, a framework for identifying and ranking consumer wants and preferences in product creation. It divides elements into three major categories: performance needs, passion needs, and fundamental necessities. It's possible that not all users have comments on the internet. A partial knowledge of the requirements of all customers may result from feedback that is biased towards a particular user group or demography. Create and improve automated systems for sentiment analysis to classify internet comments into Kano Model categories. This will help companies in promptly identifying areas for development and gauging client happiness.

Data availability

The data used to support the findings of this study are included within the article.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Funding statement

This research was supported by the project of Humanities and Social Sciences in Colleges and universities of Anhui Province (Funder ID: SK2021A0212), which are greatly acknowledged.

Acknowledgement

Humanities and Social Sciences in Colleges and universities of Anhui Province (Funder ID: SK2021A0212)

References

[1] Schaffhausen CR, Kowalewski TM. Assessing quality of unmet user needs: Effects of need statement characteristics. Design Studies. 2016 May 1; 44:1-27.

[2] Salehan, M. and Kim, D.J., 2016. Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems*, *81*, pp.30-40.

[3] Jin, J., Liu, Y., Ji, P. and Kwong, C.K., 2019. Review on recent advances in information mining from big consumer opinion data for product design. *Journal of Computing and Information Science in Engineering*, *19*(1), p.010801.

[4] Zhang L, wu L, mattila A S. Online reviews. Journal of Travel Research.2015,7(4): e3-e5.

[5] Jin, J., Ji, P. and Gu, R., 2016. Identifying comparative customer requirements from product online reviews for competitor analysis. *Engineering Applications of Artificial Intelligence*, *49*, pp.61-73. .

[6] Qi, J., Zhang, Z., Jeon, S. and Zhou, Y., 2016. Mining customer requirements from online reviews: A product improvement perspective. *Information & Management*, *53*(8), pp.951-963. .

[7] WANG Keqin, LIU Zhaoming. Product design improvement based on importance performance Competitor analysis of online reviews. Computer Integrated Manufacturing Systems, （05):1496-1506.

[8] Chang, Y., Kun, T. and Chunyang, Y., 2020. Mobile phone product improvement based on review bigdata. *Comp. Integra. Manuf. Sys*, *26*(11), pp.3074-3083.

[9] Wang, Y., Zheng, L.W., Zhang, Y.Y. and Zhang, X.Y., 2020. Software requirement mining method for Chinese APP user review data. Comp Sci, 47(12), pp.56-64.

[10] JIA Danpin, JIN Jian, GENG QIAN, DENG Siyu.A Kansei Engineering Integrated Approach for Customer-needs Mining from Online Product Reviews. Journal of the China Society for Scientific ad Technical Information,2020,39(03):308-316.

[11] Yang, H., 2019. Fusion of data and design: research on the innovation path of big data analysis to derive insight into user needs. *Decoration*, *4*(05), pp.100-103.

[12] Li, S.B., Quan, H.F., Hu, J.J., WU, Y. and Zhang, A., 2018. Perceptual evaluation method of products based on online reviews data driven. *Computer Integrated Manufacturing Systems*, *24*(3), pp.752-762.

[13] Grădinaru, C., Obadă, D.R., Grădinaru, I.A. and Dabija, D.C., 2022. Enhancing Sustainable Cosmetics Brand Purchase: A Comprehensive Approach Based on the SOR Model and the Triple Bottom Line. *Sustainability*, *14*(21), p.14118.

[14] Ji, X., Gao, Q., Li, X.F. and Gao, F., 2020. Reviews mining and requirements elicitation methodology considering product attributes' hierarchy. Computer Integrated Manufacturing Systems, 26(3), pp.747-759.

[15] WANG Jun, LIU Siyi.Research on Product Design Method by Online Review Text Driven.Machine Design and Research,2022,38(04):1-5+11.

[16] Chatterjee P. Online Reviews: Do Consumers Use Them.Advances in Consumer Research,2001,28.

[17] Zhe, C., Huilan, G., Jiang, W. and Zhongyi, H., 2023. The Ideal and Reality of Metaverse: User Perception of VR Products Based on Review Mining. *Data Analysis and Knowledge Discovery*, *7*(1), pp.49-62.

[18] Kano, N., 1984. Attractive quality and must-be quality. *Journal of the Japanese society for quality control*, *31*(4), pp.147-156. .

[19] Kim, H., Kim, D., Yang, H., Lee, K., Seo, K., Chang, D. and Kim, J., 2008. Development of a wall-climbing robot using a tracked wheel mechanism. *Journal of mechanical science and technology*, *22*, pp.1490-1498. .

[20] LUO, L., GE, Q.D., ZHOU, Y.X. and XIA, B., 2021. Association Rule Mining of Equipment Data Based on Apriori Algorithm. *Command Control and Simulation*, *43*(6), pp.29-33.

[21] LI Jiangyong, Zhang Wei.Product Improvement Design from the Perspective of Online Comment Data Mining.Packaging Engineering,2021,42(06):135-141.

[22] Wang, K. and Xia, R., 2016. A survey on automatical construction methods of sentiment lexicons. Acta Automatica Sinica, 42(4), pp.495-511. .

[23] Jakob, N. and Gurevych, I., 2010, October. Extracting opinion targets in a single and cross-domain setting with conditional random fields. In *Proceedings of the 2010 conference on empirical methods in natural language processing* (pp. 1035-1045).

[24] Zhao, Y., Qin, B., Che, W. and Liu, T., 2011. Appraisal expression recognition with syntactic path for sentence sentiment classification. *International Journal of Computer Processing of Languages*, *23*(01), pp.21-37.

[25] WANG Keqin, PEI Fengjun.Online Comment Mining for Product Design.Improvement [J]. Computer Engineering and Applications,2019,55(19):235-245.

[26] Dewaele, P.B.P.A.G., 2008. Investigating self-similarity and heavy tailed distributions on a large-scale experimental facility.

[27] Stegeman, A., 2000. Heavy tails versus long-range dependence in self-similar network traffic. *Statistica Neerlandica*, *54*(3), pp.293-314.

[28] Yao, T.F. and Lou, D.C., 2007. Research on semantic orientation analysis for topics in Chinese sentences. *Journal of Chinese information processing*, *21*(5), pp.73-79.

[29] Xi, L., Zhang, H., Li, S. and Cheng, J., 2020. Integrating fuzzy Kano model and fuzzy importance–performance analysis to analyse the attractive factors of new products. *International Journal of Distributed Sensor Networks*, *16*(5), p.1550147720920222.

[30] Kumari, P. and Toshniwal, D., 2021. Extreme gradient boosting and deep neural network-based ensemble learning approach to forecast hourly solar irradiance. *Journal of Cleaner Production*, *279*, p.123285.

[31] Shi, Y. and Peng, Q., 2021. Enhanced customer requirement classification for product design using big data and improved Kano model. *Advanced Engineering Informatics*, *49*, p.101340.

[32] Chan, H.K., Sun, X. and Chung, S.H., 2019. When should fuzzy analytic hierarchy process be used instead of analytic hierarchy process? *Decision Support Systems*, *125*, p.113114.