

Mining customer requirements from online reviews: A product improvement perspective



Jiayin Qi^{a,*}, Zhenping Zhang^b, Seongmin Jeon^c, Yanquan Zhou^d

^a Management School, Shanghai University of International Business and Economics, No. 1900 Wenxiang Road, Songjiang District, Shanghai 201620, China

^b School of Economic and Management, Beijing University of Posts and Telecommunications, No. 10, Xitucheng Road, Haidian District, Beijing, 10086, China

^c Business School, Gachon University, 1342 Seongnamdaero, Seongnam, 461-701, South Korea

^d School of Computer Science, Beijing University of Posts and Telecommunications, No. 10 Xitucheng Road, Haidian District, Beijing, 10086, China

ARTICLE INFO

Article history:

Received 15 July 2015

Received in revised form 10 May 2016

Accepted 5 June 2016

Available online 29 June 2016

Keywords:

Online review

Big data commerce

Product design

KANO

Conjoint analysis

ABSTRACT

Big data commerce has become an e-commerce trend. Learning how to extract valuable and real time insights from big data to drive smarter and more profitable business decisions is a main task of big data commerce. Using online reviews as an example, manufacturers have come to value how to select helpful online reviews and what can be learned from online reviews for new product development. In this research, we first proposed an automatic filtering model to predict the helpfulness of online reviews from the perspective of the product designer. The KANO method, which is based on the classical conjoint analysis model, is then innovatively applied to analyze online reviews to develop appropriate product improvement strategies. Moreover, an empirical case study using the new method is conducted with the data we acquired from JD.com, one of the largest electronic marketplaces in China. The case study indicates the effectiveness and robustness of the proposed approach. Our research suggests that the combination of big data and classical management models can bring success for big data commerce.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Online reviews have become an important source of information for consumers that significantly influence consumer choices and product sales [38–47]. Research on online reviews has become an established field due to the rapid growth of electronic commerce [36,37]. From the perspective of manufacturers, the selection of helpful online reviews and learning from online reviews for new product development have both become important. Big Data, i.e., the volume, velocity and variety of primary data, provides manufacturers with the opportunity to make use of online reviews for product designs.

Online reviews could be the source of innovative ideas, providing input for new product designs and enhancements. Co-creation, the active involvement of customers in the process of new product and service development, has been identified as a reliable source of competitive advantage [1]. From the viewpoint of manufacturers, online reviews are appealing sources of customer requirements, especially for those manufacturers who must

continually renovate their products in the competitive market [13]. Through online reviews, manufacturers can listen to the voices of customers in the target market [14]. In addition, manufacturers are able to draw knowledge about the market structure and competitive landscape to support their marketing decisions [48].

A vast number of online reviews may, in theory, create problems such as information overload [61]. Previous research has found that many online marketplaces provide mechanisms for consumers to identify helpful reviews [19,37,52,53]. Much less is known, however, about how online reviews are helpful to decisions surrounding the improvement of product designs. Hence, it is important to better understand helpful online reviews from the product designer's perspective when they are developing new products.

Traditionally, companies have been dependent on market surveys to better understand consumers' requirements. Compared with offline or paper-and-pencil surveys, online reviews provide richer information in less time and at a lower cost, as the respondents are willing to participate independently [8,9]. This study proposes a new method to improve product designs using online reviews. The primary research question for this study is how to utilize online reviews for the purpose of improving product

* Corresponding author.

E-mail addresses: qijiayin@139.com (J. Qi), xiaopingzzp@bupt.edu.cn (Z. Zhang), smjeon@gachon.ac.kr (S. Jeon), zhouyanquan@bupt.edu.cn (Y. Zhou).

designs. In this context, we develop the following research questions:

- How do manufacturers identify helpful reviews?
- What is the most effective method in developing the strategies for product design?

To answer these questions, we construct a method to filter out the online reviews that are helpful to product designers in the context of new product development. We then propose a method based on conjoint analysis through attribute identification and sentiment analysis to determine the weights of the product attributes. We validate the robustness and efficiency of our model by comparing it with the existing Support Vector Machine (SVM)-based method. We then develop the product improvement strategies by combining both the conjoint analysis-based model and the KANO model. We test our method using data from JD.com, one of the largest online marketplaces in China.

Our contributions to the literature are two-fold. First, we propose an automatic filtering model to predict the helpfulness of online reviews from a product design perspective. Previous studies have commonly addressed the matter from the perspective of consumers, and emphasize filtering out the most helpful online reviews to help consumers make purchasing decisions.

Second, we provide an innovative way to use the KANO model based on massive online reviews. We bring the traditional KANO model new vitality in the Web 2.0 era. Traditionally, the KANO method has been used for standard questionnaires. In this paper, we combine conjoint analysis with the traditional KANO. Using online reviews, we find a novel way to use the KANO method. We then obtain strategies to improve products intelligently. Taking the KANO model as an example, we seek a way to make traditional management models work effectively in the era of big data commerce.

Our exploration in this study puts forward an efficient and practical way to mine value and separate noise from big data. Big data is just raw material; it is not a solution [62]. One of the main challenges for big data commerce is finding a way to convert data into insights. We must find a way to efficiently use big data, as it is a wealth of important information. In our study, we find that combining the classical management model with big data is a good way to gain insight from big data. Our work clarifies that a knowledge-based view is an important methodology for conducting big data analytics; however, the ignorance-based view of big data analytics attracted greater attention after Google's successful prediction of the flu outbreak in 2008.

Another challenge for big data is to identify what type of data are relevant to the problem and then rapidly extract that data for a timely analysis [71]. In our study, we provide two facets to solve this problem. First, we apply our improved management model to screen the data necessary for further analysis. Second, we use the filtering model to delete most of the noise from the necessary data. The methodologies we use can be extended to other industries' big data commerce practices.

The remainder of this paper is organized as follows. Section 2 briefly describes the related work. Section 3 describes the research framework of our work. Section 4 presents the data pre-processing, which involves online reviews culled from JD.com. In Section 5, the helpfulness model, from the product designer's perspective, is introduced. In Section 6, the product improvement strategies are appropriately generated by combining a conjoint analysis-based model and the traditional KANO model. Finally, this research concludes with a discussion of the managerial implications and future research directions.

2. Literature review

2.1. Big data and decision making

Data provides insights about consumer behavior. Businesses make use of those insights for market intelligence. Data analytics enables decision makers to find hidden patterns in data [63,69]. The emergence of big data, the unprecedented volume, velocity, and variety of data available from individual consumers, has changed the ways that businesses understand consumer behavior, including how businesses make decisions [67].

The three Vs (i.e., volume, velocity, and variety) have been commonly used to define big data [64,70,71]. It is estimated that the amount of big data from the global market doubles every two years. This is the result of businesses' efforts to manage the rapidly increasing volume of big data [72]. Velocity, or the persistent rapidity of data creation, is another dimension of big data [64]. Businesses utilizing rich, insightful, current data can make better decisions grounded in evidence rather than intuition. A variety of sources of big data supply a diverse richness that surpasses traditional data.

One of the major differences between big data and traditional data may be the shift from structured transactional data to unstructured user generated content or UGC data [73]. Structured data have been collected through scanner or sensor data, files, and databases. Unstructured data were mostly captured through social media. It not only contains textual data extracted from blogs and text messages but also non-textual data (e.g., videos, audio and images).

In addition to the three Vs, two more Vs play key roles in explaining big data: veracity and value [64,65]. The veracity of big data can be a major issue when the quality of the data is in question, although the volume, velocity, and variety of data are continuously increasing [70,71]. In addition, the question of value is raised with the ever-increasing size of big data. By eliminating unimportant and irrelevant data, the remaining data can be useful and valuable in providing business insights [3–6,64].

2.2. Helpfulness of online reviews

Previous studies have found that online reviews are much more reliable than the information extracted from other sources [15,16]. The literature has used a variety of measures to measure the performance of online reviews: purchase intentions [49], readership [50], and sales revenue [10,44,51]. Helpfulness is the most widely used measure to predict the performance of online reviews [11,19,52]. It is useful to understand the perceived helpfulness of online reviews, as helpful online reviews play an important role in purchasing decisions [53,54].

Ample research in information systems and marketing has highlighted the helpfulness of online reviews from the perspectives of consumers. Forman et al. evaluated online review helpfulness in the context of e-commerce [55]. They presented reviewer expertise and attractiveness as two dimensions of source credibility that are related to perceived helpfulness. Online review helpfulness could be influenced positively by its length or readability [56]. Zhu et al. found that perceived helpfulness on Yelp is mainly related to the central and peripheral cues of the argument [57]. The valence of online reviews is another important factor that influences the perceived helpfulness. Review valence can be defined as negativity and positivity connotations or the orientation of a review [58]. Positive reviews are likely to have a higher impact on purchasing decisions than negative reviews [47]. Hu et al., however, found that negative reviews contain clearer information than positive reviews [59]. That being said, the effect of negative reviews diminishes over time [58]. In addition to the

review content, the reviewer should be considered. Ngo-Ye et al. found that reviewer engagement is an important factor in review helpfulness [60]. Zhu et al. illustrated that a reviewer's expertise and online attractiveness positively influences perceived review helpfulness [57]. The volume of online reviews, however, is often too large to examine. Furthermore, online reviews vary greatly in quality, with many unsolicited multiple postings [17].

More relevant to our research, Ku et al. (2012) found that four variables successfully discriminate reputable reviewers from others: 1) trust intensity, 2) average trust intensity of trustors, 3) degree of review focus in the target category, and 4) average product rating in the target category. Product type is a significant control variable to use in analyzing the four product categories, and has been used on Epinions.com [18].

An analysis of online reviews across six products from Amazon.com indicated that review extremity, review depth, and product type affect the perceived helpfulness of the online reviews [19]. Similarly, Korfiatis et al. (2008) found that the usefulness score of a particular review is affected by the qualitative characteristics of the review, as measured by readability tests [15].

Based on the analysis of three factors (i.e., the reviewer's expertise, the writing style of the review, and the timeliness of the review), Liu et al. presented a non-linear regression model for to predict helpfulness and demonstrate that the proposed approach is highly effective [17]. Kim et al. used the SVM regression on a variety of features over Amazon.com product reviews to assess the helpfulness of a review and obtained promising results. They also found that the most useful features are the length of the review, its unigrams, and its product rating [20].

Most research takes the helpfulness voting, helpfulness voting ratio, or indicators constructed based on the helpfulness voting, as the golden standard. This assesses review helpfulness from the perspective of helping consumers make a purchasing decision. In contrast, there is little research focusing on the helpfulness of reviews from the perspective of product designers. To the best of our knowledge, only one paper is closely related [21]. Generally, consumers search online reviews to learn more about a product and see if the product meets their expectations. Product designers, on the other hand, must focus on the discovery of product defects and customer requirements implied by the online reviews and attempt to optimize the quality of a product at an appropriate cost. These two types of helpfulness are clearly quite different.

Liu et al. (2013) found a weak correlation when comparing the helpful voting ratio from Amazon.com with the designers' rating [21]. By interviewing designers, they proposed four categories of features: 1) linguistic features, 2) product features, 3) features based on information quality, and 4) features based on information theory. These features can be used to predict the helpfulness of a review from the perspective of a designer. Note that product features were deleted when the feature selection scheme was applied. Furthermore, both the reviewer data and the metadata are ignored, as both focused on extracting variables completely from the review text.

2.3. Measurement of customer requirements

Requirement measurements are fundamental to product positioning and strategic marketing development [27]. Conjoint analysis and the KANO model are two widely known methods used to measure customer requirements [26]. Conjoint analysis is a method that estimates the structure of a consumer's requirements (e.g., part worth, importance weights, ideal points) given his/her overall evaluations of a set of alternatives that are pre-specified in terms of the levels of different attributes [28].

The other method used to evaluate the impact of product attributes on customer satisfaction is the two-dimensional KANO

model [2]. The five main product attribute categories in a KANO model are: 1) must-be quality, whose attributes consist of the basic product criteria. Customers will be extremely dissatisfied if these basic criteria are not fulfilled, although fulfillment will not increase satisfaction level because customers take their presence for granted; 2) one-dimensional quality, the presence of which increases satisfaction levels while its absence proportionally decreases satisfaction levels. This type of attribute provides customer loyalty for firms. Next is 3) attractive quality, whose attributes usually act as a weapon to differentiate companies from their competitors because its functional presence generates absolutely positive satisfaction while customers will not be dissatisfied at all when it is not fulfilled; 4) indifferent quality, whose attributes make little contribution to customer satisfaction regardless of whether they are present or absent in a product; and 5) reverse quality, whose attributes should be removed from a product because their functional presence is harmful to customer satisfaction. Customer satisfaction is likely to be influenced by the various attributes in different ways. Therefore, a strategy should be developed that includes the different types of attributes.

Conjoint analysis and the KANO model are normally implemented by offline investigations or experiments [29] that are easy to implement, although they are also costly. In addition, it is difficult to capture dynamic changes over time with offline methods, as all of the analyses are based on questionnaires. Thus, how to respond to customer requirements immediately becomes an important issue. Some researchers have tried to learn about customer requirements from other data sources (e.g., tweets, customer complaints).

Abrahams et al. defined and validated a method and a system for automated defect detection and prioritization in the automotive industry by employing text mining on a popular social media location used by vehicle enthusiasts: online discussion forums, which had shown that vehicle quality management can be supported by appropriate analysis of social media posts [30]. Lee and Bradlow presented a method to support conjoint study designs by automatically eliciting an initial set of attributes and levels from online customer reviews using the language of the consumer [12]. Taking the pros and cons expressed as the individually perceived strengths and weakness of the respective products, Decker and Tursov proposed a negative binomial regression approach to estimate aggregate consumer requirements from online product reviews [7].

As the research described previously did not consider the influence of opinion spam, that is, untruthful or low-quality reviews, the customer requirement extracting from the online reviews may be biased. The basic potential of detecting opinion spam before a requirement measurement is also emphasized in the outlook on further applications [7]. Our research differentiates from the previous literature in both aspects in that the sentiment orientation of the review text, instead of the overall rating of the review, is used as customer satisfaction and that the attributes are not only prioritized but also classified, based on the KANO model.

3. Research framework

The explosion of online reviews brings both opportunities and challenges. Big data analytics are closely related with business applications. First, we take cell phones as our research object and use online reviews in Chinese as our source of big data. Thus, we must analyze unstructured data to translate the short Chinese text into structured data for further analysis. This is the basic portion of our research, which includes Data investigation, Data crawling, Lexicon building (Attribute lexicon and Sentiment Lexicon), and Sentiment Analysis, see Part I in Fig. 1. Second, the reviews are posted independently by consumers immediately after purchase,

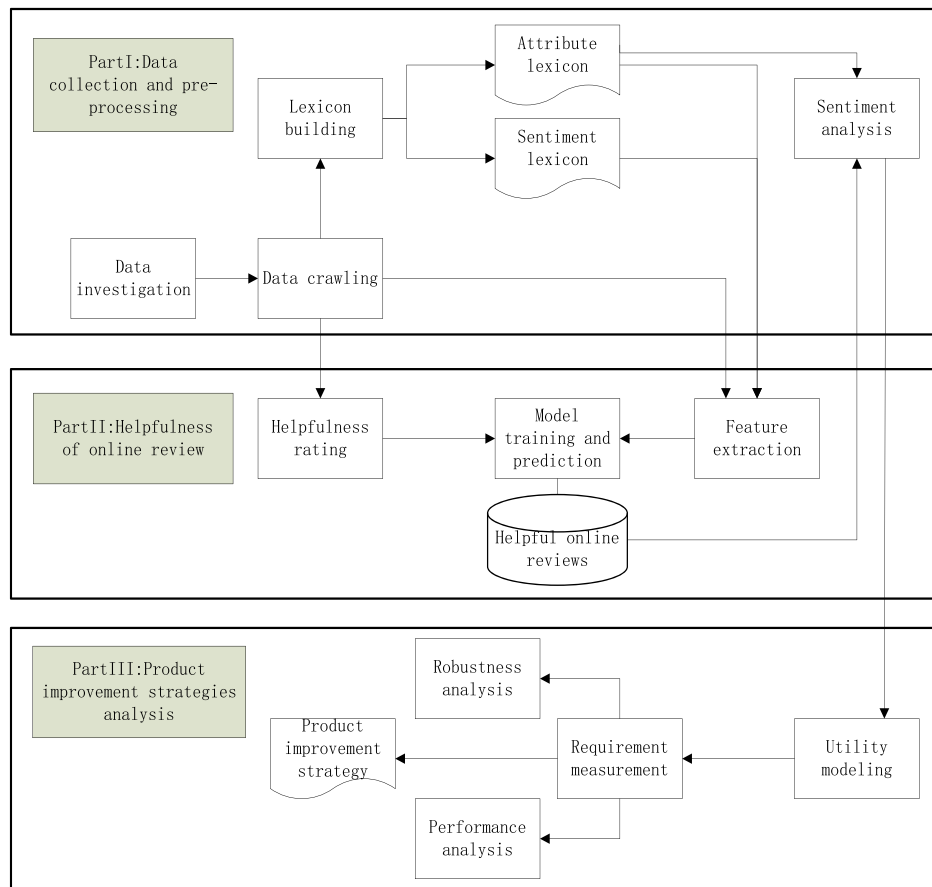


Fig. 1. Research framework.

providing firms with quick feedback; however, there is a large volume of online reviews of varying quality. Useful reviews are like treasures buried in the soil. To catch the essence of the online reviews, we must measure the helpfulness. This leads to the second part regarding the helpfulness of online reviews from an enterprise perspective, including Feature extraction, Helpfulness rating, Model training and prediction, and Helpful online reviews; see Part II in Fig. 1. Third, the amount of value big data can add depends on how important a role big data can play in decision making. In this research, we work to mine the values of big data to help manufacturers make product improvement decisions. This is the last part of our work involving Utility modeling, Requirement measurement (Robustness analysis, Performance analysis) and Product improvement strategy; see Part III in Fig. 1. The logic of three research parts is as follows. The data collection and pre-processing form the base of the research. The usefulness of online reviews is a filtering section based on the original data collection and pre-processing. The product improvement strategies analysis is the decision making section based on useful online reviews. Our research framework is illustrated in Fig. 1.

4. Data collection and pre-processing

In this section, we choose the data resource and research object and then apply natural language processing (NLP) techniques to infer the sentiment orientation towards the attributes and products for further analysis.

4.1. Online review resource selection

JD.com and Taobao.com are two of the most widely known e-commerce websites in China. JD.com primarily focuses on digital products, while Taobao.com is more of an online supermarket. Increasing numbers of Chinese people are purchasing cell phones on JD.com. As such, online reviews of cell phones are more abundant on JD.com than Taobao.com. Therefore, we chose JD.com as the online review resource.

4.2. Research objects of cell phones

Reviews of cell phones are our research priority, and we must determine which cell phones to include in this investigation. We use four principles to select the cell phones. First, we expect that cell phones with the highest market shares should be on the list. Second, we include both high end and low end cell phones. Third, we assume that international brands and domestic brands are both important. The final principle is that all of the selected cell phones should have enough online reviews on our target e-commerce website of JD.com. Using these four principles, we selected 1200 of the best-selling cell phones.

4.3. Research samples

We developed a data crawler to collect all of the online reviews of these 1200 cell phones on JD.com. Because of the limitations of

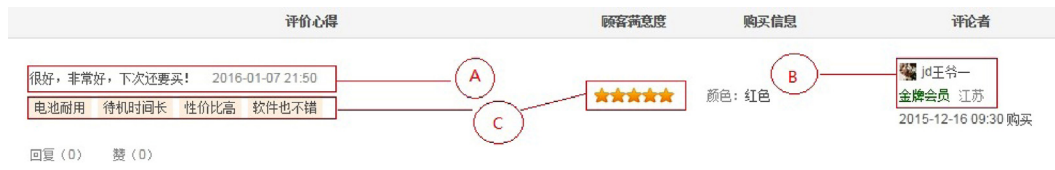


Fig. 2. An example of a Chinese online review.

anti-crawling, only 757 of the 1200 cell phones were included; from these, 679,422 online reviews were crawled from JD.com.

4.4. Data pre-processing

Fig. 2 illustrates a typical online review posted in Chinese. Three pieces of information are included. Part A is the text, which is the actual review posted by the consumer. Part B is the reviewer's data, which indicates the name and grade of the reviewer. Part C is the metadata, which includes the data about the review. By metadata, we refer to the consumer's conclusion of this review (e.g., the labels, ratings, pros, cons, and interactive information

that concerns other consumers, such as the number of helpful votes and the number of replies).

4.4.1. Attribute identification and sentiment analysis

The first step is to identify the product attributes. We apply POS (Part-of-Speech) tagging. A Latent Dirichlet Allocation and Page Rank are then used to rank the terms based on the frequency and the semantic relationship of the terms. A total of 4105 terms were extracted as candidate attributes. We then classified the filtered attribute terms into 15 categories, consulting professional cell phone designers. A lexicon of 1123 words was constructed to identify the attribute terms in the online reviews.¹ The classified attributes and relative number of terms are described in Table 1.

Because the sentiment phases for cell phones are specific, an existing sentiment lexicon such as ICTCLAS (Institute of Computing Technology, Chinese Lexical Analysis System) is not suitable for this study. Therefore, a sentiment lexicon is built in a similar way to the attribute lexicon in Table 2 below. The sentiment lexicon, including 909 words, is built.²

From the attribute terms, we conduct a backwards search on the review to find the sentiment words, due to Chinese grammar. These technologies have been extensively researched [22–25,34,35]. The sentiment analysis results of the review are displayed in Table 3.

By transforming the results in Table 3, we obtain Table 4. Here, the labels “pos” and “neg” indicate whether attribute j ($j = 1, \dots, 15$) obtains a positive or negative sentiment in review i ($i = 1, \dots, 679,422$) of a specific product (A_1, \dots, A_{757}). Otherwise, attribute j is coded as a missing value, “mv.”

Let $h \in \{\text{pos}, \text{neg}\}$ be the respective sentiment of attribute j . Using this notation, the nominally coded data (seen in Table 4) can be converted into binary data, with:

$$X_{ij}^h = \begin{cases} 1, & \text{if attribute } j \text{ takes level } h \text{ in review } i \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

As Table 5 shows, we have two sentiment variables for each attribute j . If both X_{ij}^{pos} and X_{ij}^{neg} equal 0, then review i obtains a missing value for attribute j .

¹ The attribute lexicon can be found at http://www.datasci.icoc.cc/nd.jsp?id=30&_np=2_363.

² The sentiment lexicon can be found at http://www.datasci.icoc.cc/nd.jsp?id=30&_np=2_363.

We obtain dataset $D_1 = \{x\}$, where x is a 30-dimension vector representing consumers' sentiment orientation towards different attributes.

Table 1

Classified attributes and relative number of terms.

Attribute	Number of terms	Attribute	Number of terms
Edition	29	Feeling	33
CPU	105	Appearance	208
Battery	46	Logistics	61
Function	201	System	166
Music	18	Signal	40
Price	57	Camera	37
Compatibility	40	Rest	23
Screen	59		

Table 2

Classified sentiment and relative number of terms.

Sentiment	Number of terms
Positive sentiment	554
Negative sentiment	355

Table 3

Results of attribute extraction and sentiment analysis.

Review	Positive	Negative
$i=1$	CPU	Feeling
$i=2$	Appearance	Battery
\dots	\dots	\dots
$i=679,422$	Null	Camera

Table 4

Structure data of online reviews.

Attribute				
Review	j=1	j=2	\dots	j=15
$i=1$	Pos	Neg	\dots	mv
\dots	\dots	\dots	\dots	\dots
$i=679,422$	Mv	Pos	\dots	neg

Table 5

Structure data of online reviews.

Attribute					
Review	$X_{ij}^{\text{pos}} (j=1)$	$X_{ij}^{\text{neg}} (j=1)$	\dots	$X_{ij}^{\text{pos}} (j=15)$	$X_{ij}^{\text{neg}} (j=15)$
$i=1$	1	0	\dots	0	0
\dots	\dots	\dots	\dots	\dots	\dots
$i=679,422$	0	0	\dots	0	1

4.4.2. Sentiment orientation of review text

To measure customer requirements, we must measure the utility of the customers. Previous research takes the overall rating as the gold standard, where 1 star represents most unsatisfied and 5 stars represent most satisfied [7]; however, customers' ratings vary. For example, when customers feel satisfied, some will use a 3 rating, while others use a 4 or 5. As word of mouth becomes more and more important, some vendors tend to give gifts or rebates to obtain high ratings.

Among the 757 products reviewed in this investigation, the average rating was found to be as high as 4.7. Compared to ratings, the sentiment in the text is more consistent among the consumers, as consumers must recall the purchase and use experience carefully and write down the most impressive and/or relevant parts. Therefore, the NLP technique is applied to infer the sentiment of the review text as the utility of the customers, which is measured on a three-point scale: 1 denotes positive, -1 denotes negative and 0 denotes neutral. We then obtain a new dataset $D_2 = \{y\}$ that denotes the sentiment orientation of the review text.

5. Helpfulness of online reviews

5.1. Feature extraction

5.1.1. Linguistic features

Linguistic features capture the linguistic aspect of the review text. Liu et al. (2013) found that product designers prefer reviews with more useful information. More useful reviews are determined by the length and volume of the sentences [21]. In addition, reviews with proof and evidence, and therefore, more adjectives and adverbs, are more likely to be trusted by designers.

5.1.2. Features based on information quality

Features based on information quality are those regarding information coverage and information accuracy [21]. In an interview, it was found that product designers have different judgments of reviews, even if the reviews are of a similar length.

This is because the number of product features mentioned in the reviews is largely differentiated. The time of the review posting can also affect designers' judgment, so the timeliness of the reviews is also included.

5.1.3. Features based on information theory

Features based on information theory measure the information gain of the review [21]. Product designers trust the reviews that simultaneously state the pros and cons; they also trust reviews with novel opinions. Accordingly, Liu et al. (2013) constructed 3 features based on information theory. These features measure the information gain of the review. For example, these features will be not equal 0 only when negative and positive attitudes towards a specific product attribute both exist. As such, we can identify whether a reviewer assesses both the pros and cons of the attribute.

5.1.4. Reviewer features

Reviewer features are the features (e.g., the expertise and activeness) about the reviewers. E-commerce websites typically use membership or score systems, so they gather a great amount of information about a reviewer, including the volume of the reviews posted in the past by a specific reviewer and the grade of the reviewer. The volume of reviews posted indicates the expertise of the reviewer, while the grade of a reviewer indicates the reviewer's activeness on the website. If a reviewer is highly active, they are more likely to provide thorough explanations of their viewpoints.

5.1.5. Metadata features

Metadata are data about data. Metadata features are the descriptions of the review text (e.g., pros, cons and labels) that are filled in by reviewer. These descriptions are concerned with the reviewers' involvement. The number of helpful votes and the number of replies indicates the evaluation level from other consumers, while the rating is the overall evaluation of the product from the reviewer.

In conclusion, we obtain five categories of 25 features as shown in Table 6.

Table 6
Five categories of features.

Category	NO.	Feature alias	Description
Linguistic features (Liu et al., 2013)	1	L-NW	# of words
	2	L-NS	# of sentences
	3	L-ALS	Average length of sentence
	4	L-NADJ	# of adjectives
	5	L-NADV	# of adverbs
Features based on information quality (Liu et al., 2013)	6	IQ-NSS	# of subjective sentences
	7	IQ-NOS	# of objective sentence
	8	IQ-TIM	# of total elapsed days
	9	IQ-NRP	# of referred products
	10	IQ-NPF	# of product features
	11	IQ-NSPF	# of sentences referring to product features
	12	IQ-RPFR	# of product features/# of sentences referring to product features
	13	IQ-RPFS	# of product features/# of sentences
	14	IQ-RRS	# of sentences referring to product features/# of sentences
Features based on information theory (Liu et al., 2013)	15	IT-SI	The self-information sum of product features
	16	IT-DS	The divergence of sentiment sentences
	17	IT-SS	The strength of sentiment sentences
Reviewer features	18	R-NR	# of reviews
	19	R-TGR	The grade of reviewer
Metadata features	20	M-WPF	Whether pros is filled or not
	21	M-WCF	Whether cons is filled or not
	22	M-NL	# of labels
	23	M-NHV	# of helpful votes
	24	M-NR	# of replies
	25	M-NS	# of stars

5.1.6. Helpfulness prediction model and input variable determinations

The feature extraction resulted in five categories with 21 features for the helpfulness prediction (because of the website alteration of JD.com, 4 features were deleted from the webpage when we started our data crawling). The process of our analysis is executed in three steps, described below.

Step 1: Helpfulness rating. We randomly select one twentieth of the 679,422 online reviews as a training sample; 27,929 reviews were selected. Every review is scored on a 5-scale Likert, from 1 to 5, by two product designers, with 1 representing the least helpful and 5 representing the most helpful. The mean of their scores is calculated as the final score of the review.

Step 2: Model selection. According to the five categories of features discussed above, we extract 21 features from the reviews to predict the helpfulness of the online reviews. Three text features (e.g., linguistic features, features based on information quality, features based on information theory) have been proven to be useful for predictions [21]. Therefore, in this paper, we only test whether the reviewer features and metadata features are reliable for the helpfulness prediction.

Step 3: Significance analysis and helpfulness prediction. The significance of the features is analyzed, and the significant features are chosen for the helpfulness prediction.

5.2. Experiment results

The results of the helpfulness rating are presented in Table 7. In total, 67.98 percent of the reviews were rated below 1.5; 82.36 percent were rated below 2. Hence, most of the reviews were rated as unhelpful by the product designers. This demonstrates the necessity and urgency of spam detection.

Liu et al. (2013) inferred a reviews' helpfulness entirely from the text (see Model I):

$$y = \alpha \text{Text}_{\text{feature}} \quad (2)$$

In this paper, we propose two new categories of features: reviewer features and metadata features. With these two categories of features included, we obtain Model II:

$$y = \alpha \text{Text}_{\text{feature}} + \beta \text{Reviewer}_{\text{feature}} + \gamma \text{Metadata}_{\text{feature}} \quad (3)$$

As Table 7 illustrates, with reviewer features and metadata features included, the AIC decreases from 1.852 to 1.831 and the R-square increases from 0.321 to 0.336. Therefore, Model II is selected for a further analysis, with all five categories of features included.

The results of Model II are described in Table 8. Most of the variables are significant; the IQ-NOS (number of objective sentence) is omitted due to co-linearity.

As we can infer from Table 9, linguistic features are critical for helpfulness ratings. If the review is full of description, the review is likely to be rated as helpful. The helpfulness is negative in terms of the length and number of sentences, however, implying that long

Table 8

Comparison of Models I and II.

Model	AIC	R-square
I	1.852	0.321
II	1.831	0.336

reviews are not necessarily mean they are helpful for product designers. Rather, it is what they say that matters.

Eight out of nine information quality features are significant. In this way, the review that mentions product features in different sentences would be more helpful for product designers. Third, all of the information theory-based features are significant and the review mentioning both the pros and cons would be chosen. Finally, the reviewer features and one out of three metadata features are found to be significant.

Helpfulness is negative to the grade of the reviewer because the grade on JD.com is calculated merely by accumulated consumption. As such, high purchasing power does not mean helpful feedback. The effect of helpful votes and replies are also not significant, quite differently from the customer perspective. The review of low star ratings tells us more about the product defect and would thus be more likely to be regarded as helpful.

The significant variables are chosen as inputs for the helpfulness prediction. Therefore, the 17 features (i.e., 5 linguistic features, 7 features based on information quality, 3 features based on information theory, 1 reviewer feature and 1 metadata feature) are chosen to predict reviews' helpfulness.

Finally, we take the 17 features into account as inputs by using the estimated coefficients as the weights to predict reviews' helpfulness. We then rank the reviews in the reverse order of the helpfulness score. As the manual rating sample is the random sample of the reviews, we assume that the helpful review ratio is the same in both sets. Therefore, we select the reviews according to the helpful ratio in the manual rating sample. Overall, 55,000 or approximately 7.97% of the reviews were considered helpful and selected for further analysis.

6. Product improvement strategies analysis

6.1. Customer requirement weight analysis based on the conjoint analysis model

In this section, we propose an approach based on conjoint analysis to measure the impact of product attributes on consumers' satisfaction. There are three types of preference models in conjoint analysis: vector models, ideal-point models and part-worth function models [28,29].

Let $j = 1, 2, \dots, 15$ denote the set of 15 attributes, or factors, that have been chosen. The vector model posits that the consumer preference is given by:

$$y = \sum_{j=1}^{15} \beta_j X_j \quad (4)$$

where β_j is the weight for the j attribute and X_j denotes the performance of the j attribute. In this paper, there are three values for each attribute; X_j obtains a value of 1 if one consumer holds a positive sentiment for the j attribute. It obtains a value of -1 for a negative sentiment and 0 for a missing value.

The ideal-point model posits that preference y is negatively related to the squared (weighted) distance d^2 of the location $\{X_j\}$

Table 7

Results of the helpfulness rating evaluated by the product designers.

Helpfulness	Frequency	Percent	Cumulation
1	12,513	44.8	44.8
1.5	6472	23.17	67.98
2	4016	14.38	82.36
2.5	2703	9.68	92.03
3	1118	4	96.04
3.5	630	2.26	98.29
4	284	1.02	99.31
4.5	128	0.46	99.77
5	65	0.23	100
Total	27,929	100	

Table 9

The estimated results of Model II.

Variable	Estimate	Std. error	p value	Variable	Estimate	Std. error	p value
L-NW	−0.00525	0.00062	***	IQ-RPFR	0.1429	0.00961	***
L-NS	−0.01756	0.00562	**	IQ-RPFS	−0.11785	0.01451	***
L-ALS	−0.00385	0.00181	*	IQ-RRS	0.09992	0.02795	***
L-NADJ	0.09301	0.00288	***	IT-SI	0.20524	0.01761	***
L-NADV	0.02792	0.00199	***	IT-DS	0.18916	0.05096	***
IQ-NTD	0.02741	0.00517	***	IT-SS	−0.42817	0.04942	***
IQ-NSS	−0.00011	0.00002	***	R-TGR	−0.00091	0.0002	***
IQ-NOS	0	(omitted)		M_NHV	−0.00078	0.00069	0.256
IQ-NRP	−0.05823	0.01196	***	M_NR	0.00054	0.00176	0.759
IQ-NPF	−0.00883	0.009	0.327	M_NS	−0.10975	0.00462	***
IQ-NSPF	0.15439	0.01279	***	Intercept	1.82951	0.03495	***

AIC=1.831; R²=0.336; Significance levels (two-tailed p values): ***, **, and * are statistically significant at the 0.001, 0.01, and 0.05 level, respectively.

from the individual's ideal point $\{X_{ideal}\}$, where d^2 is given by:

$$d^2 = \sum_{j=1}^{15} \beta_j (X_j - X_{ideal})^2 \quad (5)$$

Therefore, the preference will increase as the product comes closer to the ideal point.

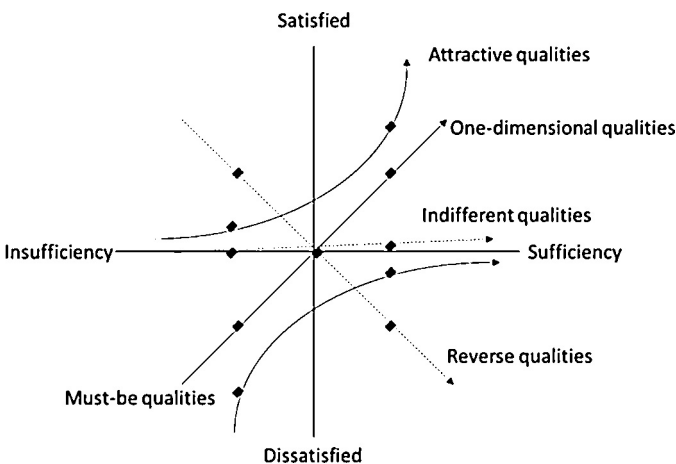
The part-worth function model posits that:

$$y = \sum_{j=1}^{15} f_j(X_j) \quad (6)$$

where f_j is the function denoting the part worth of the different levels of f_j for the j^{th} attribute. In practice, $f_j(X_j)$ is calculated for a selected set of levels for X_j (usually three or four), with the part worth for intermediate X_j obtained by a linear interpolation. Thus, the part-worth function is represented as a piecewise linear curve.

In the vector model, preference changes linearly with performance. In an ideal point model, preference first increases with the performance and then decreases when the performance exceeds the ideal point. The part-worth function model provides the greatest flexibility in allowing different shapes for the preference function along each of the attributes.

In Section 4, we transform the reviews expressed in natural language into structured data. There are 3 discrete levels for each attribute: positive, negative and missing values. As the famous KANO model shows, customer satisfaction may not change linearly with the attribute level. Hence, it is not appropriate to determine an ideal product in a rapidly changing market, so the part-worth model is chosen for our study.

**Fig. 3.** Mapping the attributes according to the KANO model.

Taking each review as a stimuli perceived by the consumers and sentiment of text as consumer utility, our model is:

$$y = \alpha + \sum_{j=1}^{15} (\beta_j^{\text{pos}} X_j^{\text{pos}} + \beta_j^{\text{neg}} X_j^{\text{neg}}) \quad (7)$$

where y is consumers' utility and $X_j^{\text{pos}} = 1$ denotes a consumer holding a positive sentiment for attribute j and $X_j^{\text{neg}} = 1$, respectively. If the j attribute obtains a missing value, both X_j^{pos} and X_j^{neg} equal 0. β_j^{pos} is the preference for the positive sentiment and β_j^{neg} is the preference for the negative sentiment. The preference for a missing value is 0, by default.

Based on Eq. (7), the parameters can be estimated. For each attribute, we have three observation points, that is $(-1, \beta_j^{\text{neg}})$, $(0, 0)$ and $(1, \beta_j^{\text{pos}})$. As illustrated in Fig. 4, with these three observation points, we can map the attributes to different categories according to the KANO model.

Here, we try to provide the mapping rules. First, we construct two basic variables as:

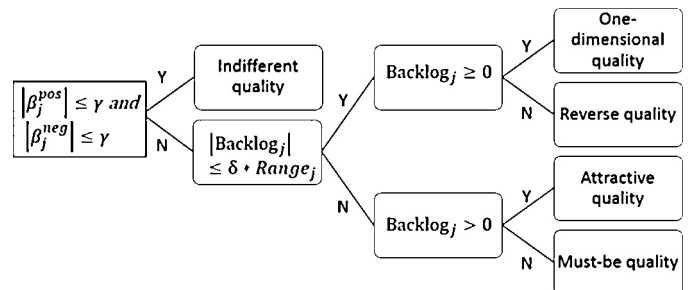
$$\text{Range}_j = |\beta_j^{\text{pos}} - \beta_j^{\text{neg}}| \quad (8)$$

$$\text{Backlog}_j = \beta_j^{\text{pos}} + \beta_j^{\text{neg}} \quad (9)$$

Based on the calculation of range, the weight of each attribute can be given as:

$$\text{Weight}_j = \text{Range}_j / \sum_{j=1}^{15} \text{Range}_j \quad (10)$$

If the preference stays close to zero at all times, $|\beta_j^{\text{pos}}| \leq \gamma$ and $|\beta_j^{\text{neg}}| \leq \gamma$, the attribute can be mapped to indifferent qualities and γ is the threshold that must be set. Otherwise, if the three observation points look like a straight line, $|\text{Backlog}_j| \leq \delta * \text{Range}_j$,

**Fig. 4.** Mapping rules according to the KANO model.

the attributes can be mapped to one-dimensional qualities when $Backlog_j \geq 0$ and mapped to reverse qualities when $Backlog_j < 0$. Otherwise, the attributes can be mapped to attractive qualities when $Backlog_j > 0$ and must-be qualities when $Backlog_j < 0$. The inference logic is shown in Fig. 3. In conclusion, the mapped rules can be constructed as a decision tree (Fig. 4).

6.2. Experiment results

We apply product attribute identification and sentiment analysis to the selected reviews. To validate the robustness of the conjoint analysis model, an SVM-based model is applied.

6.2.1. Experiment results of a conjoint analysis-based model

The 55,000 helpful reviews are used for inputs into the parameter estimation. The conjoint analysis-based model was estimated using Stata 12.0. AIC is 2.325 and R-squared equals 0.195. The parameter estimate results of the model are presented in Table 9. Most of the parameters (26/30) are statistically significant.

Based on the estimated results, we calculated the range and the weight (Table 10). The top 4 attributes are: battery life, signal, camera, and appearance. A cell phone with a low quality battery, unclear voice service, low performing camera, or poor appearance is more likely to appear as unsatisfactory. The two least important attributes are edition and price. As multi-carrier phones develop and the significant indication of a support edition on the product detail page of the cell phones appears, edition becomes less important. Further, as competition increases, the price for each market segment is transparent, which gives consumers stable expectations (Table 11).

6.2.2. Robustness and performance analysis

(1) Robustness checks

In a previous section, 55,000 helpful online reviews were selected for model training. To test the robustness of our model, a different proportion of reviews were chosen to estimate the parameters. We compare the results for a robustness checks.

A total of 55,000 reviews were divided into 5 random samples and added to the training set one by one. Similarly, the range and weight were calculated and the ranks of the attributes were compared. As Table 12 illustrates, with the sample added in, the performance stayed consistent and the ranks of the attribute became increasingly stable.

Table 10

Estimated results of the parameters.

Variable	Estimate	Std. error	p value
Intercept	−0.066	0.007	***
Variable	Estimate	Std. error	p value
pos_edition	−0.059	0.06	0.323
neg_edition	−0.161	0.021	***
pos_cpu	0.023	0.028	0.421
neg_cpu	−0.227	0.01	***
pos_battery	0.249	0.013	***
neg_battery	−0.181	0.01	***
pos_function	0.236	0.017	***
neg_function	0.006	0.009	0.493
pos_music	0.356	0.021	***
neg_music	0.191	0.008	***
pos_price	0.445	0.015	***
neg_price	0.356	0.01	***
pos_compatibility	−0.273	0.113	*
neg_compatibility	−0.403	0.026	***
pos_screen	0.312	0.016	***

AIC=2.325; R²=0.195; Significance levels (two-tailed p values): ***, **, and * are statistically significant at the 0.001, 0.01, and 0.05 level, respectively.

Table 11

Range and weight of each attribute.

attribute	range	weight	attribute	Range	weight
battery	0.431	0.165	system	0.163	0.062
signal	0.304	0.116	compatibility	0.13	0.05
camera	0.29	0.111	feeling	0.121	0.046
appearance	0.29	0.111	logistics	0.11	0.042
cpu	0.25	0.096	edition	0.102	0.039
screen	0.247	0.095	price	0.09	0.034
function	0.23	0.088	rest	−0.313	−0.12
music	0.165	0.063			

Table 12

Comparison of ranks.

Attribute	Number of samples				
	1/5	2/5	3/5	4/5	5/5
battery	2	1	1	1	1
signal	7	2	4	2	2
camera	5	3	3	3	3
appearance	3	4	5	4	4
cpu	4	7	7	5	5
screen	9	6	6	6	6
function	8	5	8	7	7
music	11	10	9	9	8
system	10	11	10	10	9
compatibility	1	8	2	8	10
feeling	13	12	12	12	11
logistic	12	13	11	11	12
edition	6	9	13	14	13
price	14	14	14	13	14
rest	15	15	15	15	15
R-square	0.199	0.198	0.197	0.197	0.195
AIC	2.328	2.329	2.326	2.324	2.325

The correlation coefficient increased with the samples added; the average of the correlation coefficient is 0.89. The robustness of our model is checked as shown in Fig. 5.

(2) Performance analysis

Jin's (2013) SVM-based method was used to check the performance of our model [33]. Because the SVM-based method is time-consuming and uses an algorithm whose complexity increases non-linearly with the number of reviews, a smaller dataset of 6941 helpful reviews of 4 mobile phones is used for

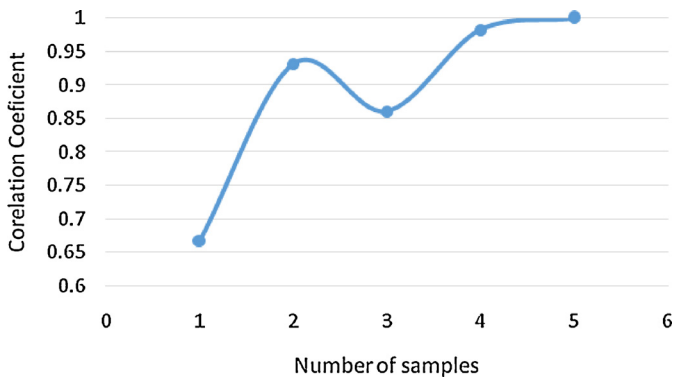


Fig. 5. Correlation coefficient of ranks.

performance analysis. Similar to Section 3, the data are pre-processed. As these reviews largely ignore the music and compatibility of the phones, 13 out of 15 attributes are identified in these reviews.

Let X_{ij} be the respective sentiment value of attribute j ; we transform the data in Table 5 by:

$$x_{ij} = \begin{cases} 1, & \text{if attribute } j \text{ takes positive in review } i \\ 0, & \text{if attribute } j \text{ takes missing value in review } i \\ -1, & \text{if attribute } j \text{ takes negative in review } i \end{cases} \quad (11)$$

Similarly, we can obtain a dataset $D_3 = \{(x, y), X^*Y\}$; x is a 13-dimension vector representing consumers' sentiment orientation towards different attributes. y is the sentiment of the review text, which is a three scale value. As the SVM is usually used for a binary classification [31,32], the data should be transformed by:

$$y_m - y_n = \begin{cases} 1, & y_m > y_n \\ 0, & y_m = y_n \\ -1, & y_m < y_n \end{cases}, \text{ while } n \quad (12)$$

By exchanging y_m and y_n , when $y_m < y_n$, we can obtain data in an induced form:

$$y_m - y_n = \begin{cases} 1, & y_m y_n \\ -1, & y_m = y_n \end{cases}, \text{ while } n \quad (13)$$

Finally, we obtain a new dataset $D_4 = \{(\hat{x}, \hat{y}), X^*Y\}$, while $\hat{x} = x_m - x_n$ and $\hat{y} = y_m - y_n$ ($y_m \geq y_n$). The weighting problem can be solved as a binary classification problem. Taking each transformed review as a classification sample, the weighting problem can be solved as an SVM-based model, like:

$$\begin{aligned} \min & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.}, & \hat{y}_i = \omega^T \hat{x}_i \\ & y_i \hat{y}_i \geq 1 - \xi_i \\ & \xi_i \geq 0 \\ & \omega_i \geq 0 \end{aligned} \quad (14)$$

Table 13
Weight of each attribute.

Attribute	Weight	attribute	weight
Logistics	0.12	system	0.071
Edition	0.108	signal	0.067
Screen	0.094	function	0.062
Appearance	0.078	battery	0.061
Cpu	0.077	rest	0.056
Price	0.077	camera	0.056
Feeling	0.074		

ξ_i is a slack variable and C is a penalty coefficient. Because the weight is non-negative, we add a new constraint: $\omega_i \geq 0$.

The SVM-based method is run on Matlab 8.1. As shown in Eq. (12), when we have N reviews, $N * (N - 1)/2$ constraints would be generated. As a result, the algorithm complexity grows non-linearly with the number of reviews. For convenience, the reviews are equally divided into 6 random samples and the model is trained for each sample. The weights for the 6 samples are averaged as the final score of each attribute in Table 13 (the penalty coefficient $C = 1000$).

The top 3 attributes are logistics, edition, and screen. A cell phone that has an unexpected edition, low-performance screen, or is distributed to the customer slowly, is more likely to arise as unsatisfactory. The two least important attributes are rest and camera. The attribute rest concerns the operating system's update and network resources; camera is mainly about camera pixels. Because these two attributes are always mentioned on the product details page and the network resources are open to everyone, customers are more familiar with them and they produce less uncertainty. Other attributes are between these 5.

We then compare the conjoint analysis-based model with the SVM-based model. First, the correlation between the results of the two methods is calculated. As Table 13 illustrates, the ranking of the attributes is quite similar, with a correlation coefficient of up to 0.84, which validates the robustness of our model. As we compare Eqs. (7) and (14), we see that the basic model of the SVM is a linear model, which is consistent with the part worth model (Table 14).

Second, we focus on the comparison of the sentiment and the prediction of our model. MAE, RMSE, and PMCC are calculated to measure the performance of the model. A ten-fold cross validation is adopted, which divides the dataset into ten folds; each fold is used for a test once with the other nine folds used for training. As shown in Fig. 6, the PMCC and MAE of the two models are nearly equal and a conjoint analysis dominates the SVM in the RMSE. In summary, the conjoint analysis performs better than the SVM.

Finally, in terms of execution time, a conjoint analysis-based model is much better than a SVM-based model. The conjoint analysis-based model takes only two seconds to estimate the parameters on a Dell Spring 1440 notebook with P7350 CPU and 4G memory, while the SVM-based model takes more than 24 h on a Dell Power Edge R710 server with 6 CPU and 16G memory.

In conclusion, the conjoint analysis-based model is consistent with the SVM-based model, with their rank correlation up to 0.84. The conjoint analysis-based model is more efficient and accurate than the SVM-based model.

(3) Reliability analysis

As indicated by the QFD (Quality Function Deployment) model, the product attributes must be mapped to engineering features

Table 14
Comparison of ranks.

Attribute	Conjoint Analysis	SVM
Logistics	1	1
Edition	2	2
Screen	3	3
Appearance	4	4
System	5	8
Cpu	6	5
Price	7	6
Signal	8	9
Battery	9	11
Rest	10	12
Function	11	10
Camera	12	13
Feeling	13	7
Correlation coefficient	0.84	

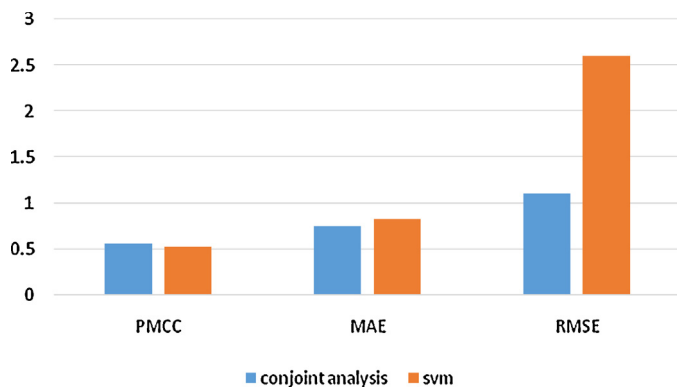


Fig. 6. Performance of conjoint analysis and SVM.

when product development decisions are made; however, the prioritization of the product attributes can tell product designers which features significantly affect the customer utility and how much the customer utility would be improved if the specific attribute was developed. The result can easily supplement the product design decision process or cooperate with other requirement measurement methods.

Our study was sponsored by the China Mobile Communications Corporation (CMCC). After we obtained the results, the Terminal Department of CMCC organized a panel meeting to evaluate our analysis results. The nine experts from the Terminal Department of the CMCC insisted that our final product improvement strategies reflected the actual situation. Hence, we developed a software tool that is now being used in the Beijing Branch of the CMCC.

We also performed an offline survey on the campus of Beijing University of Posts and Telecommunications on January 15, 2016. We administered a standard KANO questionnaire to 135 students and collected 123 effective response samples. The comparison of the offline results and our results indicates they are largely consistent (12/15). The non-consistent attributes are CPU, camera, and rest, which are classified as must-have quality based on online reviews but as attractive on the offline survey. The contradiction may result from the sampling bias. It is possible that many college students use low end cell phones, while our online reviews involve both low end and high end cell phones. This may cause the difference between the online analysis and offline survey.

Today, the battery, camera, and appearance, which rank first in our analysis, have become the highlight of cell phone advertisements (e.g., iPhone, OPPO, Huawei, Xiaomi) [74]. This can determine the accuracy of our analysis, to a certain extent. For example, the slogan “5 min charging, two hours calling,” which

illustrates the outstanding battery performance, has been popular among consumers.

6.2.3. Product improvement strategies

Our study has placed the estimated results of Table 12 into (8) and (9) and obtained the classification of attributes based on the KANO model (Table 15).

Attributes that are identified as must-have attributes include: edition, CPU, compatibility, signal, camera, and rest. These attributes can cause great dissatisfaction when they are lacking, so they should be considered a high priority in product improvement. The battery is a one-dimensional attribute whose variation has a positive linear relationship with satisfaction.

Attractive attributes include function, music, price, screen, feeling, appearance, logistics, and system. These attributes can greatly increase satisfaction when they are fulfilled, but will not cause a serious decline in satisfaction when they are not. As such, producers should first meet the needs of the must-have and one-dimensional attributes.

7. Discussion and conclusions

Online reviews have been studied by many scholars due to their rich content and high reliability. Unlike ample research from the consumer perspective, we take a product design perspective in this study. We use this perspective to develop strategies to improve product designs using online reviews. To accomplish this, we combine spam detection, consumer preference measurements, conjoint analysis, and the KANO model in a unique manner. Our work has useful implications for Big Data commerce.

7.1. Theoretical implications

Big data commerce relies on a large volume of data to glean valuable, real-time insights that drive smarter, more profitable business decisions. Taking the KANO model into consideration, based on online reviews, we show that big data analytics can obtain insights from the knowledge-based view to support big data commerce. Big data can be described as a holistic approach to manage, process, and analyze the 5Vs (i.e., volume, variety, velocity, veracity and value), which is important to create actionable insights for sustained value delivery, performance measurement, and the establishment of competitive advantage [66].

Understanding how to mine value from big data is essentially important for big data commerce. There are two ways to obtain insights: the knowledge-based view and the ignorance-based view [67]. Many computer researchers are well versed in obtaining novel discoveries from big data based on the view of ignorance. Although “what we don’t know” is much larger than “what we have known,” humans have accumulated an enormous amount of knowledge to better understand the outside world. Therefore, if we can find a way to transform existing knowledge to adapt to big data, we are more likely to gain reliable insights than when we simply rely on an ignorance-based view.

Xu et al. (2015) insist that firms adopting high levels of traditional marketing analysis and big data analysis have the highest level of new product success [68]. Our paper supports Xu et al. (2015)’s opinion. Thus, our research illustrates that the knowledge-based view is also effective for big data analytics, and thus supports big data commerce.

How should we realize the value of big data commerce from the knowledge-based view? Today, when people talk about big data analytics, they commonly expect the discoveries that were previously completely unexpected from the ignorance-based view. Data driven decision making is a hot topic in today’s research

Table 15
Category of each attribute.

attribute	range	backlog	Category
battery	0.431	0.068	one-dimensional
edition	0.102	−0.22	must-have
CPU	0.25	−0.205	must-have
compatibility	0.13	−0.676	must-have
signal	0.304	−0.199	must-have
camera	0.29	−0.084	must-have
rest	−0.313	−1.267	must-have
function	0.23	0.242	Attractive
music	0.165	0.546	Attractive
price	0.09	0.801	Attractive
screen	0.247	0.377	Attractive
feeling	0.121	0.399	Attractive
appearance	0.29	0.672	Attractive
logistics	0.11	0.534	Attractive
system	0.163	0.598	Attractive

environment. We must argue that the knowledge-based view is still relevant in the era of big data commerce. The data driven decision making pattern does not simply rely on data, it also relies on domain knowledge.

If we can combine big data with classical management models, then we yield twice the results with half the effort. This means we can obtain more reliable insights with much higher efficiency from big data if we combine big data and classical management models. However, how should we combine the two?

From our research trial, we generalize the following four steps. First, the classical models should be reformed under the context of big data. In general, most of the classical management models involve certain constructs that cannot be measured directly by objective data. Instead, they must be measured by a questionnaire from a subjective evaluation by respondents. As such, if we want to use classical models with big data, the first thing we should do is to transfer these models to other equivalent models that can use objective data directly. For example, in our research, we found that we could not use the KANO model directly with big data; however, if we connect the KANO with the conjoint model, we obtain an equivalent model that allows us to use the framework of the KANO with big data directly.

Second, we must use the equivalent model to guide the processing of big data. If we have no or little knowledge of big data analytics, it will be more difficult and fairly inefficient to mine insights. The equivalent models of the classical models can guide us to do a big data analytics. When we have the equivalent models, we will naturally know the dependent variables and the independent variables and will thus understand what type of information to extract from the big data. We also know the relationships among the variables that should be tested in the future.

Third, we must use the correct techniques to change the unstructured data into structured data. Sentiment analysis technology, which involves candidate term extraction, classification, identification, and sentiment inference, has been widely researched; however, as previous research found, sentiment analysis is highly domain-specific [30] and the general technique should be modified to adapt to the specific materials. For example, in our research, the attribute and sentiment lexicon are built from the cell phone reviews and used for further processing.

The last step is to find a computing model that allows the entire decision making process to run efficiently. As shown in Section 6.2.2, different parameters could be used to evaluate the models. Thus, we should carefully evaluate the performance of different methods and the most efficient method should be selected to be applied in the real situations.

Our research also illustrates that big data has big value. At the same time, it is also quite noisy. In this study, we obtained 679,442 online reviews from 757 types of cell phones. After the helpfulness analysis was conducted, there were just 55,000 online reviews left for further analysis. The helpfulness rate is 7.97%. Hence, filtering according to different research aims is crucial for big data analysis. Filtering removes noise and creates more possibilities to obtain insights.

7.2. Practical implications

Big data commerce is a new trend in e-commerce. Our results reveal that the big data commerce industry can harvest the following insights from our research.

First, our study shows that business insights can be discovered from big data with knowledge-based value. Traditional management models are still useful in the era of big data commerce. Big data commerce should not be solely dependent on data driven decisions: it should also rely on knowledge driven decisions. The

best way to accomplish this is by combining big data and knowledge.

Second, our study proposes a methodology for obtaining value from big data with a knowledge-based view. We learned how to mine customer requirements from online reviews from the perspective of product improvements as our research topic, but our methodology can be extended to other areas of big data commerce.

Third, in relation to the manufacturing industry, our research provides a practical procedure to improve the development of new products. The results are useful in filtering online reviews and handling big data. Our results are also helpful for producers seeking to improve their product designs. Moreover, our method helps producers better understand their consumers' feedback in a timely manner. This can result in a more rapid response to customers' requirements.

Although we use cell phones as a research object, our model was developed with a general purpose and can be easily extended to other industries (e.g., other digital products).

7.3. Limitations

One obvious limitation in our research is that we assume that preferences are consistent across all consumers in this study. As customer segmentation theory shows, there may be different types of preferences among different types of market segmentation. Thus, conducting a clustering process before a preference measurement may provide a way to improve our model in the future.

References

- [1] C. Lorenzo-Romero, E. Constantinides, L.A. Brünink, Co-creation: customer integration in social media based product and service development, *Procedia - Soc. Behav. Sci.* 148 (2014) 383–396.
- [2] N. Kano, N. Seraku, F. Takahashi, S.I. Tsuji, Attractive quality and must-be quality, *J. Jpn. Soc. Qual. Control* 14 (2) (1984) 147–156.
- [3] T. Jackson, Prosperity without growth: economics for a finite planet, *Inform. Theory IEEE T.* 56 (10) (2009) 4956–4980.
- [4] J. Guiltinan, Creative destruction and destructive creations: environmental ethics and planned obsolescence, *J. Bus. Ethics* 89 (1 Suppl) (2008) 19–28.
- [5] J. Chapman, Design for (emotional) durability, *Des. Issue* 25 (4) (2009) 29–35.
- [6] A. Leonard, A. Conrad, The story of stuff: how our obsession with stuff is trashing the planet, our communities, and our health-and a vision for change, *Libr. J.* 136 (4) (2010) (38–38).
- [7] R. Decker, M. Trusov, Estimating aggregate consumer preferences from online product reviews, *SSRN Electron. J.* 27 (4) (2010) 293–307.
- [8] E.D.D. Leeuw, W.D. Heer, Longitudinal and international comparison, *Chin. J. Electron.* 41 (3) (2002) 515–518.
- [9] R.M. Groves, Nonresponse rates and nonresponse bias in household surveys, *Public Opin. Q.* 70 (5) (2006) 646–675 (30).
- [10] J.A. Chevalier, D. Mayzlin, The effect of word of mouth on sales: online book reviews, *Nber Working Pap.* 43 (3) (2003) 345–354.
- [11] A. Ghose, P.G. Ipeirotis, Designing novel review ranking systems: predicting the usefulness and impact of reviews, *Int. Conf. Electron. Comm. ACM* (2007) 303–310.
- [12] T.Y. Lee, E.T. Bradlow, S.O. Kimbrough, B. Padmanabhan, Y. Wind, Automatic construction of conjoint attributes and levels from online customer reviews, *J. Mark. Res.* 1 (215) (2007) 898–3664.
- [13] N. Franke, F.T. Piller, Key research issues in user interaction with user toolkits in a mass customisation system, *Int. J. Technol. Manage.* 26 (26) (2003) 578–599.
- [14] A.L. Wiley, The voice of the customer, *Tech. Commun.* 40 (4) (1993) 774–777.
- [15] N. Korfiatis, D. Rodríguez, M.A. Sicilia, The Impact of Readability on the Usefulness of Online Product Reviews: a Case Study on an Online Bookstore, Springer, Berlin Heidelberg, 2008, pp. 423–432.
- [16] P.J. Sher, S.H. Lee, Consumer skepticism and online reviews: an elaboration likelihood model perspective, *Soc. Behav. Personal. Int. J.* 37 (1) (2009) 137–143.
- [17] Y. Liu, X. Huang, A. An, X. Yu, Modeling and predicting the helpfulness of online reviews, *IEEE International Conference on Data Mining (ICDM)*, Pisa, Italy, 2008, pp. 443–452.
- [18] Y.C. Ku, C.P. Wei, H.W. Hsiao, To whom should I listen? Finding reputable reviewers in opinion-sharing communities, *Decis. Support Syst.* 53 (3) (2012) 534–542.
- [19] S.M. Mudambi, D. Schuff, What makes a helpful online review? A study of customer reviews on amazon.com, *MIS Q.* 34 (1) (2010) 185–200.

- [20] S.M. Kim, P. Pantel, T. Chklovski, M. Pennacchiotti, Automatically assessing review helpfulness, Conference on Empirical Methods in Natural Language Processing (EMNLP), Sydney, Australia, 2006, pp. 423–430.
- [21] Y. Liu, J. Jin, J.A. Harding, R.F.K. Fung, Identifying helpful online reviews: a product designer's perspective, *Comput. -Aided Des.* 45 (2) (2013) 180–194.
- [22] B. Liu, M. Hu, J. Cheng, Opinion observer: analyzing and comparing opinions on the Web, International Conference on World Wide Web. ACM (2005) 342–351.
- [23] C. Fellbaum, G. Miller, WordNet, An electronic lexical database, *Cognition Brain & Behavior* (1998).
- [24] R. Agrawal, S. Rajagopalan, R. Srikanth, Y. Xu, Mining newsgroups using networks arising from social behavior, World Wide Web Conference Series, ACM (2003) 529–535.
- [25] M. Abulaish, J. Jahiruddin, M.N. Doja, T. Ahmad, Feature and opinion mining for customer review summarization, Pattern Recognition and Machine Intelligence (PRMI), New Delhi, India, 2009, pp. 219–224.
- [26] H. Sattler, S. Hensel-Börner, A comparison of conjoint measurement with self-explicated approaches, *Conjoint Meas* 29 (3) (2001) 121–133.
- [27] J. Jiao, T.W. Simpson, Z. Siddique, Product family design and platform-based product development: a state-of-the-art review, *J. Intell. Manuf.* 18 (1) (2007) 5–29.
- [28] P.E. Green, V. Srinivasan, Conjoint analysis in consumer research: issues and outlook, *J. Consum. Res.* 5 (2) (1978) 103–123.
- [29] P.E. Green, V. Srinivasan, Conjoint analysis in marketing: new developments with implications for research and practice, *J. Mark.* 54 (4) (1990) 3–19.
- [30] A.S. Abrahams, J. Jiao, G.A. Wang, W. Fan, Vehicle defect discovery from social media, *Decis. Support Syst.* 54 (1) (2012) 87–97.
- [31] C. Cortes, V. Vapnik, Support-vector networks, *Mach. Learn.* 20 (3) (1995) 273–297.
- [32] A.M. Deris, A.M. Zain, R. Sallehuddin, Overview of support vector machine in modeling machining performances, *Procedia Eng.* 24 (8) (2011) 308–312.
- [33] J. Jin, Information Mining from Online Reviews for Product Design, The Hong Kong Polytechnic University, 2013.
- [34] M. Hu, B. Liu, Mining and summarizing customer reviews, Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, Washington, USA, 2004.
- [35] B. Liu, Sentiment analysis and subjectivity, *Handbook of natural language processing*, 2010, pp. 627–666.
- [36] C. Dellarocas, The digitization of word of mouth: promise and challenges of online feedback mechanisms, *Working Pap.* 49 (10) (2003) 1407–1424.
- [37] D. Yin, S.D. Bond, H. Zhang, Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews, *MIS Q.* 38 (2013) 539–560.
- [38] J. Berger, A.T. Sorensen, S.J. Rasmussen, Positive effects of negative publicity: when negative reviews increase sales, *Soc. Sci. Electron. Publ.* 29 (5) (2007) 815–827.
- [39] P.K. Chintagunta, S. Gopinath, S. Venkataraman, The effects of online user reviews on movie box-office performance: accounting for sequential rollout and aggregation across local markets, *Market. Sci.* 29 (5) (2010) 944–957.
- [40] C. Dellarocas, X. Zhang, N.F. Awad, Exploring the value of online product reviews in forecasting sales: the case of motion pictures, *J. Interact. Mark.* 21 (4) (2007) 23–45.
- [41] W. Duan, B. Gu, A.B. Whinston, Do online reviews matter?—An empirical investigation of panel data, *Decis. Support Syst.* 45 (4) (2008) 1007–1016.
- [42] C. Forman, A. Ghose, B. Wiesenfeld, Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets, *Soc. Sci. Electron. Publ.* 19 (3) (2008) 291–313.
- [43] X. Li, L.M. Hitt, Self selection and information role of online product reviews, *Inform. Syst. Res.* 19 (4) (2007) 456–474.
- [44] Y. Liu, Word of mouth for movies: its dynamics and impact on box office revenue, *J. Mark.* 70 (3) (2006) 74–89.
- [45] E. Toby Stuart, S. Olav, Social networks and entrepreneurship, *Handbook of Entrepreneurship Research*, Springer US, 2005, pp. 233–252.
- [46] M. Sun, How does the variance of product ratings matter? *Manage. Sci.* 58 (4) (2012) 696–707.
- [47] F. Zhu, X. Zhang, Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics, *J. Mart. Q. Publ. Am. Mark. Assoc.* 74 (2) (2013) 133–148.
- [48] O. Netzer, R. Feldman, J. Goldenberg, M. Fresko, Mine your own business: market-structure surveillance through text mining, *Mark. Sci.* 31 (3) (2012) 521–543.
- [49] D.H. Park, J. Lee, I. Han, The effect of on-line consumer reviews on consumer purchasing intention: the moderating role of involvement, *Int. J. Electron. Comm.* 11 (4) (2007) 125–148.
- [50] M. Salehan, K. Dan, Predicting the performance of online consumer reviews: a sentiment mining approach, 2014.
- [51] N. Hu, N.S. Koh, S.K. Reddy, Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales, *Decis. Support Syst.* 57 (1) (2014) 42–53.
- [52] Q. Cao, W. Duan, Q. Gan, Exploring determinants of voting for the “helpfulness” of online user reviews: a text mining approach, *Decis. Support Syst.* 50 (3) (2011) 511–521.
- [53] P.Y. Chen, S. Dhanasobhon, M.D. Smith, All reviews are not created equal: the disaggregate impact of reviews and reviewers at Amazon.com, *Soc. Sci. Electron. Publ.* 10 (11) (2008) 396–401.
- [54] D. Yin, S.D. Bond, H. Zhang, Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews, *MIS Q.* 38 (2013) 539–560.
- [55] C. Forman, A. Ghose, B. Wiesenfeld, Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets, *Soc. Sci. Electron. Publ.* 19 (3) (2008) 291–313.
- [56] K.K.Y. Kuan, K.L. Hui, P. Prasarnphanich, H.Y. Lai, What makes a review voted? An empirical investigation of review voting in online review systems, *J. Assoc. Inform. Syst.* 16 (1) (2015) 48–71.
- [57] L. Zhu, G. Yin, W. He, Is this opinion leader's review useful? Peripheral cues for online review helpfulness, *J. Electron. Comm. Res.* (2014) 15 p..
- [58] S. Basuroy, S.A. Ravid, How critical are critical reviews? The box office effects of film critics, *Star Power, and Budgets*, *J. Mark.* 67 (4) (2003) 103–117.
- [59] N. Hu, J. Zhang, P.A. Pavlou, Overcoming the J-shaped distribution of product reviews, *Commun. ACM* 52 (10) (2009) 144–147.
- [60] T.L. Ngo-Ye, A.P. Sinha, T.L. Ngo-Ye, The influence of reviewer engagement characteristics on online review helpfulness: a text regression model, *Decis. Support Syst.* 61 (4) (2014) 47–58.
- [61] Q. Jones, G. Ravid, S. Rafaeli, Information overload and the message dynamics of online interaction spaces: a theoretical model and empirical exploration, *Eng. Manage. Rev. IEEE* 38 (1) (2010) 91–109.
- [62] Z. Xu, G.L. Frankwick, E. Ramirez, Effects of big data analytics and traditional marketing analytics on new product success: a knowledge fusion perspective, *J. Bus. Res.* 69 (5) (2015) 1562–1566.
- [63] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, The KDD process for extracting useful knowledge from volumes of data, *Commun. ACM* 39 (11) (1996) 27–34.
- [64] M. Lycett, Data fication: making sense of (big) data in a complex world, *Eur. J. Inform. Syst.* 22 (4) (2013) 381–386.
- [65] K. Ebner, T. Buhnen, U. Nils, Think big with big data: identifying suitable big data strategies in corporate environments, Hawaii International Conference on System Sciences (HICSS), Hawaii, USA, 2014.
- [66] S.F. Wamba, S. Akter, A. Edwards, G. Chopin, D. Gnanzou, How ‘big data’ can make big impact: findings from a systematic review and a longitudinal case study, *Int. J. Prod. Econ.* 165 (2015) 234–246.
- [67] S. Erevelles, N. Fukawa, L. Swayne, Big data consumer analytics and the transformation of marketing, *J. Bus. Res.* (2015).
- [68] Z. Xu, G.L. Frankwick, E. Ramirez, Effects of big data analytics and traditional marketing analytics on new product success: a knowledge fusion perspective, *J. Bus. Res.* 69 (5) (2015) 1562–1566.
- [69] J.W. Cantile, System for in-line processing of envelopes and the like: US, 1986.
- [70] IBM, What Is Big Data? [online], (2012) . (accessed 12.05.24) <http://www-01.ibm.com/software/data/bigdata/>.
- [71] Oracle White Paper Oracle, Big data for the enterprise. [online] 2012 Available: <http://www.oracle.com/us/products/database/big-data-forenterprise-519135.pdf> (accessed 14.02.01).
- [72] IDC, The Digital Universe of Opportunities: Rich Data & the Increasing Value of the Internet of Things. [online], (2014) . (accessed 14.07.01) <http://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf>.
- [73] Integreon Insight, Big just got bigger. [online] 2012 Available: http://www.integreon.com/pdf/Blog/Graill-Research-Big-Data-Just-Got-Bigger_232.pdf (accessed 12.05.17).
- [74] ZDC, How Do People like iPhone 6s/plus? An Survey of 958 Chinese Consumers. [online], (2015) . (accessed 16.01.12) <http://zdc.zol.com.cn/543/5432630.html>.



Jiayin Qi is a professor of Management School, Shanghai University of International Business and Economics. Her research interests are social network analysis, user generated contents management and emergent event management etc.

Zhenping Zhang is a graduate student of School of Economics and Management, Beijing University of Posts and Telecommunications (BUPT). His research interest is user generated contents management and customer relationship management.

Seongmin Jeon is an Assistant Professor at Gachon University in South Korea. He has two major areas of expertise: 1) empirical analysis on electronic marketplace, and 2) the use of data science methods to extract, process, and analyze online data. His works have been accepted in peer-reviewed journals including *Journal of Interactive Marketing*, *Information Systems* and *e-Business Management and Carbon Management*.

Yanquan Zhou is a professor of School of Computer Science, Beijing University of Posts and Telecommunications (BUPT). Her research interests are natural language processing and social media mining etc.