Proceedings of the ASME 2022
International Mechanical Engineering Congress and Exposition
IMECE2022
October 30-November 3, 2022, Columbus, Ohio

IMECE2022-95362

APPROACH FOR KANO-IPA ANALYSIS OF PRODUCT ATTRIBUTES FROM ONLINE REVIEWS AND PRODUCT MAINTENANCE RECORDS

Aoxiang Cheng[1], **Mengyuan Shen**[1], **Youyi Bi**[1*]

[1] University of Michigan – Shanghai Jiao Tong University Joint Institute
Shanghai Jiao Tong University
Shanghai, China

ABSTRACT
Kano analysis and importance-performance analysis (IPA) are widely used for needs analysis, product positioning, and strategic planning in product design. Previous research uses customer surveys and online reviews as the main data sources. However, these data carry inevitable subjective bias. In contrast, product maintenance records provide objective information on product quality issues and failure patterns, which can be cross-validated with customers' personal experience from online reviews. In this paper, we propose a systematic approach for conducting Kano-IPA analysis from online reviews and product maintenance records synthetically. An attribute-keyword dictionary is first established using keyword extraction and clustering methods from online reviews and maintenance records. After that, semantic groups including product attributes and associated descriptions are extracted by dependency parsing analysis. The sentiment scores of identified attributes are calculated by a self-supervised representation learning approach (Sentiment Knowledge Enhanced Pre-training, SKEP) from the built semantic groups. Sentiment scores and occurrence frequencies of attributes in online reviews are utilized for Kano analysis. The importance of product attributes in IPA is estimated from the impact of sentiments of each product attribute on product ratings, while the performance is estimated from the sentiment scores of online reviews or the quality statistics from maintenance records. A case study of passenger vehicles shows that integrated data can provide more comprehensive results and richer insights. The proposed approach enables automatic data processing and can support companies to make efficient design decisions with broader perspectives from multi-source data.

Keywords: Kano analysis, Text mining, Maintenance records, Importance-Performance Analysis, Online reviews

1. INTRODUCTION
In recent years, Kano analysis and importance performance analysis (IPA) are widely used methods for needs analysis, product positioning, and strategic planning in product design [1,2]. These methods can divide product attributes into different categories so that designers can take tailored actions when design improvements are needed. For those product attributes that matter to customers greatly but with poor performance or satisfaction levels in market, designers are suggested to prioritize optimizing them. These applications can be found in design of both products and services [3,4]. For example, Bi et al. performed IPA with online reviews of accommodation service and discussed those service attributes with urgent need of improvement [5]. Joung and Kim conducted IPA of online reviews for the design improvement of smart phones [6]. Yao et al. analyzed the survey data on mobile security applications with Kano analysis and discussed the impact of different design features on customers' satisfaction [7].

Previous research of Kano analysis and IPA uses customer surveys as the main data source. Surveys can provide detailed information about customers' opinions towards products, but they usually cost certain amount of time, money, and human resources. Online reviews are becoming more popular in recent studies as they can be collected at a lower cost [8]. These data reflect customers' thoughts and attitudes from which insights can be extracted to guide better design decisions. However, both customer surveys and online reviews carry inevitable subjective bias. In addition, customers with complaints on products are more likely to express their opinions online, while most of those satisfied with products may not participate in such activities, which can weaken the representativity of online reviews [9]. In contrast, product maintenance records provide information on product quality issues and failure patterns, which can reveal the performance of a product in a more objective way. These

[*] Corresponding author, Assistant Professor in Mechanical Engineering, Shanghai Jiao Tong University
Email: youyi.bi@sjtu.edu.cn

insights can be cross-validated with customers' personal experience from online reviews, which will contribute to a more comprehensive analysis on product attributes.

In this paper, we propose an approach for conducting Kano-IPA analysis from online reviews and product maintenance records synthetically. The textual information from these two data sources can be processed and analyzed in a unified and systematic framework. Keywords on product attributes are first identified from online reviews using a keyword extraction method based on Term Frequency-Inverse Document Frequency (TF-IDF) algorithm after sentence segmentation and part-of-speech tagging. Then, these keywords are embedded and classified into product attributes by X-means clustering method. A similar procedure is applied to the product maintenance records, and the identified product attributes are corrected by fusing the results of these two data sources. After that, semantic groups including product attributes and associated descriptions are established by dependency parsing analysis. The sentiment scores of identified attributes are calculated by a self-supervised representation learning approach (Sentiment Knowledge Enhanced Pre-training, SKEP) from the built semantic groups.

In Kano analysis, the adequacy and satisfaction of product attributes are obtained by the occurrence statistics of attributes and sentiment scores, respectively. While in IPA, the importance of the product attributes can be estimated from the impact of sentiments of each product attribute on the product rating using regression methods. The performance of product attributes in IPA can be the sentiment scores from online reviews or the quality statistics from product maintenance records.

We present a case study of passenger vehicles in China's auto market to demonstrate the proposed approach. Our study shows that integrated data can provide more comprehensive results compared to using only online reviews. Vehicles with varying sales volumes show notable differences in the results of Kano-IPA analysis. Such differences can assist automakers to further improve their designs and enhance their market competitiveness. The proposed approach enables automatic data processing and can support companies to make efficient design decisions from multi-source data.

The remaining parts of this paper are organized as follows. Section 2 explains the proposed approach including data preprocessing, attribute identification, semantic group generation, sentiment analysis and Kano-IPA analysis. Section 3 shows the results of the case study and discusses potential implications to vehicle designers. Section 4 concludes the study and provides suggestions for future work.

2. METHODOLOGY

We propose an approach for Kano-IPA analysis of product attributes from online reviews and product maintenance records as shown in Figure 1. The approach includes three stages: Data Collection and Preprocessing, Key Information Extraction (product attribute identification, semantic group generation and sentiment analysis), and Kano-IPA Analysis. We first collect online reviews and product maintenance records from product forums and data service companies, respectively. We extract key information from the preprocessed data, including the product attributes mentioned, and customers' attitudes towards these attributes. Several methods of product attribute identification, semantic group generation and sentiment analysis are developed or leveraged to support this process. After that, we can build Kano models and perform IPA based on occurrence statistics and sentiment scores of product attributes, and the quality statistics from maintenance records. The Kano-IPA analysis results can support making strategies for improving product design. The detailed procedures of each stage are provided in the following subsections.

2.1 Data collection and preprocessing

Online reviews of targeted products can be collected from e-shopping websites and online forums using web crawling techniques. Usually these reviews include posting time, word of mouth, product rating, and customer location. Product maintenance records can be collected from data service companies, which include basic specifications of products, fault descriptions, and maintenance procedures. Both online reviews and product maintenance records contain large amount of unstructured textual data, which need to be preprocessed for further analysis. Common preprocessing steps include data cleaning (e.g., eliminating the redundant, erroneous, missing, duplicate data and unnecessary symbols from the raw data), sentence splitting and word tokenization. These operations enable smooth extraction of key information from the raw data.

2.2 Key Information Extraction

2.2.1 Attribute identification of product

After preprocessing the online reviews and maintenance records, we need to identify those product attributes that most matter to customers. We first filter out the most significant words in the corpus (i.e., a collection of online reviews and maintenance records) with the common TF-IDF algorithm. Then the Skip-Gram model is leveraged to embed these words into a low-dimension vector space, and the obtained word vectors are clustered with the X-means clustering technique. These clusters can be considered as the main product attributes identified. Based on the functional structure of the product (e.g., a motor vehicle system usually consists of subsystems such as power, chassis, body and electronics), these attributes will be further corrected, i.e., those clusters sharing same product functions will be merged. Finally, the name of each cluster (i.e., identified attribute) and the significant words in each cluster form an attribute-keyword dictionary. Note that the one significant word can only be found in one cluster (i.e., one product attribute). Thus, once a significant word is detected from a piece of textual information from either online reviews or product maintenance records, we can determine what product attribute is mentioned by looking up the attribute-keyword dictionary.

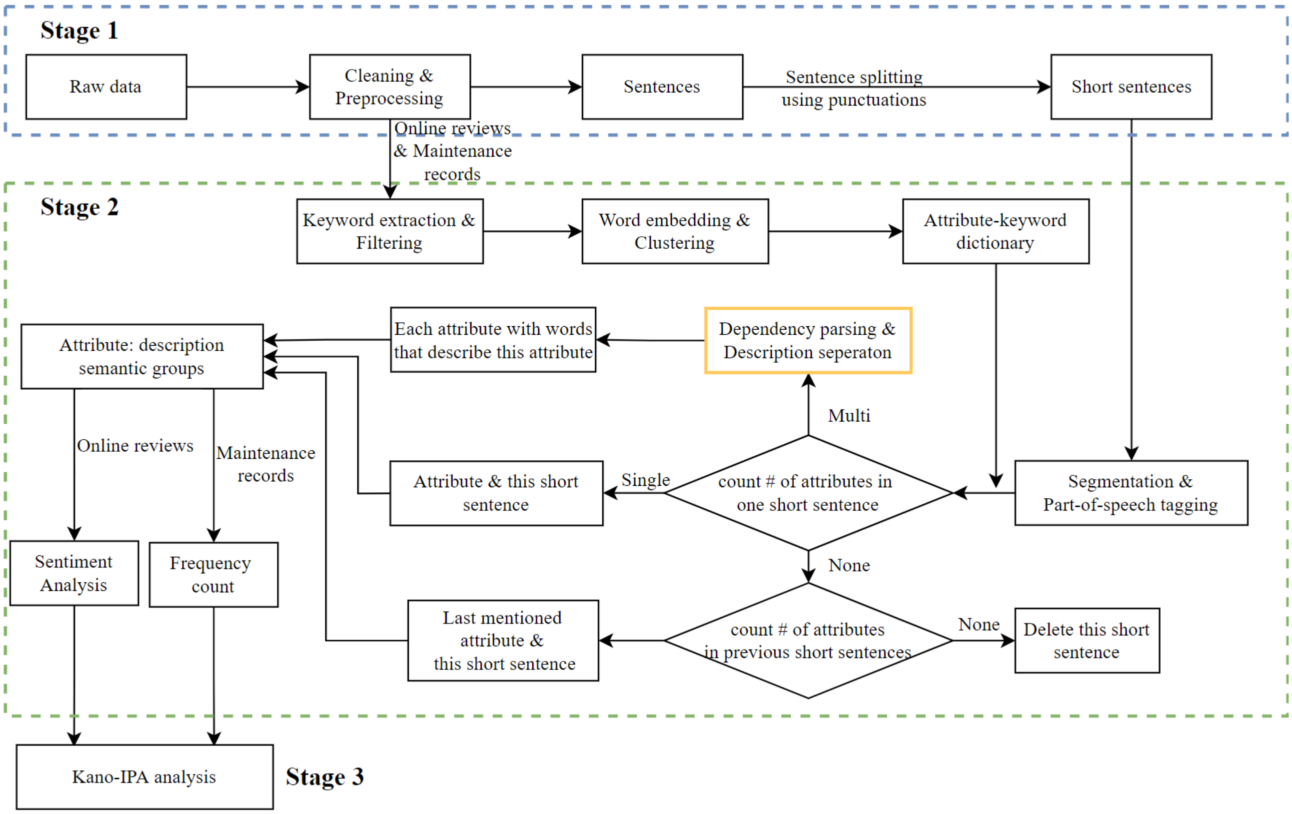

FIGURE 1: OVERALL FRAMEWORK OF THE PROPOSED APPROACH

2.2.2 Semantic group generation

A tricky issue in the analysis of customer reviews is that customers occasionally express their opinions towards one product attribute in multiple sentences, and sometimes they mention multiple attributes in one sentence. Thus, we cannot simply perform the sentiment analysis of customer reviews just by the sentences separated with periods or commas, which will bring inevitable errors and confusions. To overcome this issue, we develop a method named as semantic group generation. We define a semantic group as a collection of descriptions on a single product attribute. Here the descriptions can include one or multiple phrases and sentences. In other words, a semantic group is an attribute-description pair. One customer review includes at least one semantic group. Then in the counting of identified product attributes and calculation of sentiment scores, semantic groups can be considered as the basic textual units. The general process for semantic group generation is as shown in the green dashed box in Figure 1.

Specifically, customer reviews are first to split into short sentences separated by punctuations such as commas, semicolons, and tilde. For each short sentence, sentence segmentation and part-of-speech (POS) tagging are applied. Nouns in each sentence are extracted and searched in the attribute-keyword dictionary built in previous step to find the corresponding product attributes mentioned by customers. Based on the number of attributes identified in each short sentence, semantic groups are generated in three ways.

(1) If there is no attribute identified in the current sentence, the last-mentioned attribute in the previous sentence of the review will be assigned as the attribute of this sentence. This assigned attribute and the current sentence are combined as a semantic group. If there is no last-mentioned attribute, the current sentence will not be considered.

(2) If there is only one attribute in the current sentence, the semantic group is automatically generated by paring the identified attribute and the current sentence.

(3) If there are multiple attributes identified in one sentence, dependency parsing is applied to separate the corresponding descriptions of each identified attribute. Then each pair of identified attribute and corresponding descriptions form a semantic group. The detailed procedures are explained in following subsection.

After all semantic groups are generated from a customer review, those semantic groups describing the same product attribute will be merged.

2.2.3 Separation of descriptions for sentences with multiple attributes identified

The separation of descriptions in sentences with multi-attributes is realized based on dependency parsing. The dependency parsing can provide the syntactic structure of sentences in terms of dependency relations. An open natural language processing platform, Language Technology Platform Cloud (LTP-Cloud), is utilized in this operation [10,11]. There are totally 14 dependency relations that can be identified. Their descriptions and tags are listed in Table 1. These dependency relations are analyzed based on Chinese language grammar and more detailed explanations of these relations can be found in [11].

TABLE 1: DEPENDENDY RELATIONS FROM DEPENDENCY PARSING

Tag	Description	Tag	Description
SBV	Subject-verb	CMP	Complement
VOB	Verb-object	COO	Coordinate
IOB	Indirect-object	POB	Preposition-object
FOB	Fronting-object	LAD	Left adjunct
DBL	Double	RAD	Right adjunct
ATT	Attribute	IS	Independent structure
ADV	Adverbial	HED	Head

Figure 2 shows an example of dependency parsing. The sentence is segmented as “这个/手机/的/电池/很/耐用”, which means “This /phone /’s / battery is /very /durable”.[1] Usually SBV (Subject-verb) relation is the most common relation between keywords (e.g., “battery”) and descriptions (e.g., “durable”), which are often adjectives. Also, the ADV (Adverbial) relation modifies or qualifies the descriptions (e.g., “very durable”) and is quite useful in the sentiment analysis of sentences. Based on our preliminary study, five most important dependency relations (SBV, VOB, ATT, ADV, and COO) are selected for further analysis.

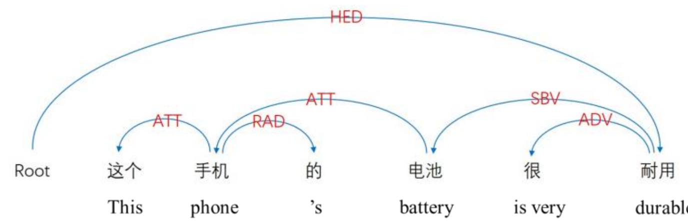

FIGURE 2: AN EXAMPLE OF DEPENDENCY PARSING. THE ORIGINAL SENTENCE IS IN CHINESE

Figure 3 illustrates the overall process of description separation for sentences with multiple attributes, which corresponds to the orange box in Fig. 1. After dependency parsing, the obtained various dependency relations (e.g., one sentence can include both ATT, COO and other relations) need to be analyzed one by one until all key product attributes with associated descriptors are extracted, from which the semantic groups can be naturally generated. In this process, some special cases must be treated first.

Case 1 is when two or more keywords appear in an ATT relation. In this case, the latter keyword is retained as it usually provides more specific information. For instance, as shown in Figure 2, both “phone” and “battery” appear in the ATT relation, and we choose to keep “battery” as the keyword. After this treatment, if there is only one attribute in the sentence, a semantic group can be generated. Otherwise, the sentence is passed to the treatment of Case 2.

In Case 2 of COO relations, usually there are multiple keywords and adjectives. We first count the number of adjectives in this relation. If one or no adjectives are mentioned in the rela-

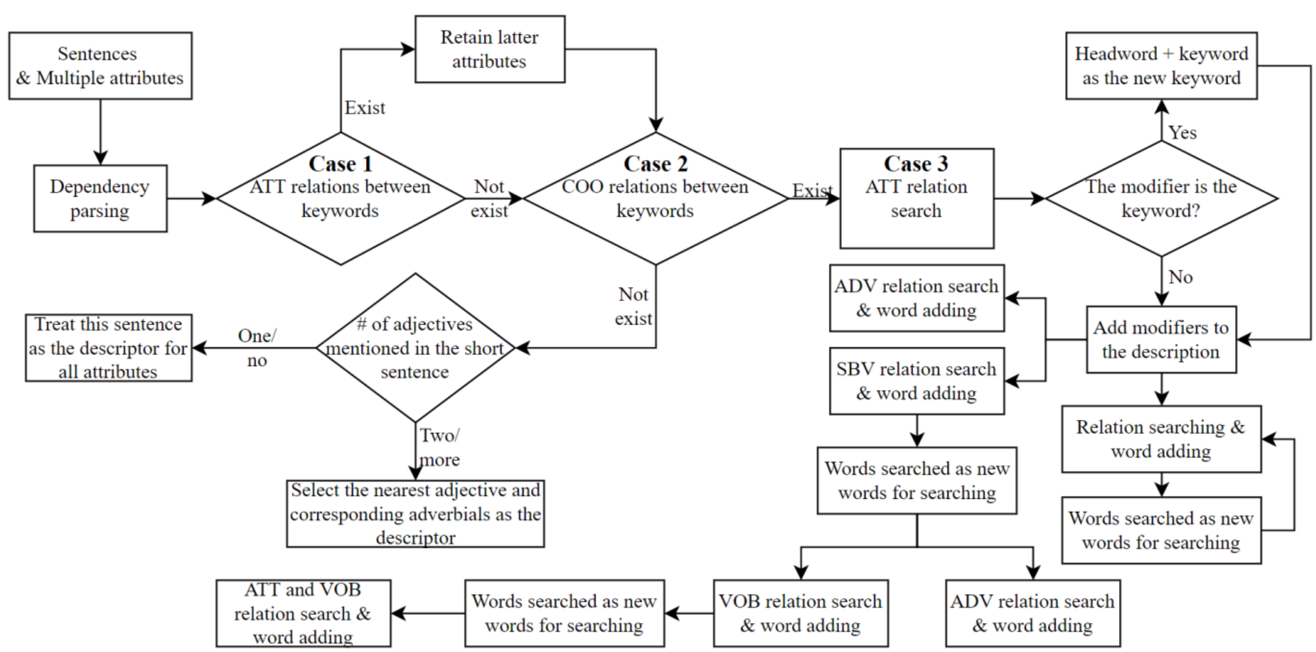

FIGURE 3: THE OVERALL PROCESS OF DESCRIPTION SEPARATION FOR SENTENCES WITH MULTIPLE ATTRIBUTES

[1] Due to the grammatical difference between Chinese and English, the final resulting structure may be different. The separation rule defined in this research can be extended to fit the English language as well.

tion, we treat this sentence as the description for all keywords. If the number is two or above, for each keyword, the adjective closer to the first identified keyword is treated as the corresponding descriptor. After that, this adjective will be deleted from the sentence to ensure the correction of the next search.

Case 3 is the general case where the ATT relations are first treated. In an ATT relation, if the keyword is the center word for this relation (e.g., "durable battery"), the modifier and its adverbial phrase are added to the related words. If the keyword is the modifier (e.g., "battery capacity"), the center word is combined with this keyword as the new keyword and is added to the related words. The following searches are applied based on the new keyword. The same procedures are iterated until all relations are analyzed.

2.2.4 Sentiment analysis.

In this study, we choose Sentiment Knowledge Enhanced pre-training (SKEP) [12], a deep learning-based model for sentiment analysis of online reviews. SKEP is a pre-trained model on a corpus containing over 3.2 million documents with excellent performance on various evaluation tests [13]. For example, the SKEP pre-trained model achieves 90.06% accuracy in a sentiment analysis task with the open dataset SE-ABSA16_CAME [14]. The SKEP model takes a semantic group as input and outputs its sentiment polarity, i.e., "positive" or "negative".

2.3 Kano Analysis

After obtaining the identified product attributes and associated sentiment polarities from previous stages, we can then perform Kano analysis to classify the product attributes.

In a dataset with N customer reviews, a review j ($j \in \{1,2,\dots,N\}$) includes a semantic group for each identified product attribute A_i, where i is the index of product attributes. After sentiment analysis, the semantic group will be assigned a sentiment polarity. We define S_i^{pos}, S_i^{neg} and SF_i as the average positive, negative and overall sentiment intensity of product attribute i. Equation (1) and Equation (2) show the calculation of the positive and negative sentiment intensity of the reviews on attribute i, respectively. Here S_{ij}^{pos} / S_{ij}^{neg} represent the positive/negative sentiment polarity on attribute i in review j, and its value can be either 1 or 0.

$$S_i^{pos} = \frac{1}{N}\sum_{j=1}^{N} S_{ij}^{pos} \tag{1}$$

$$S_i^{neg} = \frac{1}{N}\sum_{j=1}^{N} S_{ij}^{neg} \tag{2}$$

$$SF_i = \frac{1}{N}\sum_{j=1}^{N} \frac{S_{ij}^{pos} - S_{ij}^{neg}}{S_{ij}^{pos} + S_{ij}^{neg}} \tag{3}$$

Equation (3) shows the calculation of the overall sentiment intensity of product attribute A_i. Here $S_{ij}^{pos} + S_{ij}^{neg}$ denote the number of semantic groups on attribute i in review j, and $S_{ij}^{pos} - S_{ij}^{neg}$ is the difference between the number of positive and negative semantic groups on attribute i in review j. If $S_{ij}^{pos} + S_{ij}^{neg}$ equals 0 on attribute i in a review, then this review will not be considered in the calculation. The ratio in Equation (3) represents the overall sentiment intensity of attribute i.

We define a metric SA_i in Equation (4) to facilitate the Kano analysis. The SA_i index is the ratio of the average positive sentiment intensity to the average negative sentiment intensity comparing with the overall sentiment intensity of the attribute i.

$$SA_i = \frac{S_i^{pos} - SF_i}{SF_i - S_i^{neg}} \tag{4}$$

To determine the Kano category of product attributes, we follow the assignment method in [15] and define a cut-off point β as shown in Equation (5):

$$\beta = \frac{SA^{MAX} - SA^{MIN}}{I} \tag{5}$$

Here SA^{MAX} and SA^{MIN} represent the largest and smallest values of the SA_i among the all attributes, and I is the total number of identified product attributes. Thus, let $\overline{SA}$ represent the average values of the SA_i among all attributes, and a product attribute can be categorized as follows:

I. If $SA_i < \overline{SA} - \beta$, then attribute A_i is regarded as a basic attribute. For customers, this product attribute is must-be and taken for granted.

II. If $\overline{SA} - \beta \le SA_i \le \overline{SA} + \beta$, then attribute A_i is regarded as a performance attribute. When this attribute is provided, customer satisfaction will be improved. Otherwise, it will be reduced.

III. If $SA_i > \overline{SA} + \beta$, attribute A_i is regarded as an excitement attribute. If this attribute is improved, customer satisfaction can be greatly improved.

2.4 Importance and Performance Analysis (IPA)

In the Importance and Performance Analysis, the importance is estimated by measuring the influence of product attributes on customers' overall rating of the product from online reviews. The performance is estimated by calculating the occurrence frequency of the product attributes in product maintenance records or sentiment scores from online reviews. The detailed procedures are provided as follows.

2.4.1 Estimation of importance

In our study, logistic regression is leveraged to measure the influence of product attributes on customers' overall rating about the product as shown in Equation (6).

$$y \sim \alpha + \sum_{i=1}^{I}(\phi_i^{pos} S_i^{pos} + \phi_i^{neg} S_i^{neg}) \tag{6}$$

Here y is customers' overall rating about the product, and $y = 0$ represents the customer's rating is lower than a threshold value (usually the average value of ratings), while $y = 1$ represents the remaining situations. $S_i^{pos} = 1$ and $S_i^{neg} = 1$ denotes a customer holding a positive or negative sentiment for attribute i respectively. ϕ_i^{pos} and ϕ_i^{neg} are estimated weights. If an attribute is not included in the review, then $S_i^{pos} = S_i^{neg} = 0$. α

is the intercept term. All data in the regression model need to be normalized into the range of [0, 1]. In addition, if a customer's review shows positive (negative) sentiment on all product attributes, however his/her overall rating is below (higher than) the threshold, then this review is not included in the regression process.

Let $\left|\phi_i^{pos}\right|$ and $\left|\phi_i^{neg}\right|$ represent the absolute value of the estimated weight of the positive and negative sentiment feature in product attribute i, then the importance of a product attribute i can be calculated by Equation (7):

$$Imp_i = \left|\phi_i^{pos}\right| + \left|\phi_i^{neg}\right| \tag{7}$$

2.4.2 Estimation of Performance

Product maintenance records include product-related attributes (e.g., model, price, mileage of a vehicle) and maintenance information (e.g., fault description, maintenance item, repair date). We propose that the performance of a product attribute can be quantified by the occurrence frequency of its corresponding maintenance items (e.g., a fix job on vehicle engine). The more maintenance items associated with a product attribute, the higher the cost of repairing and maintaining the product. Therefore, we assume that attributes with a higher proportion of maintenance items perform worse. Equation (8) shows how to calculate the performance of attribute. Equation (9) normalizes the performance of each product attribute.

$$P_i = \frac{1}{M}\sum_{i=1}^{M} P_{ik} \tag{8}$$

$$\bar{P}_i = \frac{P^{MAX} - P_i}{P^{MAX} - P^{MIN}} \tag{9}$$

Here P_i represents the average number of the maintenance items of attribute i in M maintenance records. P_{ik} is the number of related maintenance items regarding product attribute i in maintenance record k. $\bar{P}_i$ is the normalized result, and the closer value to 1, the better the performance of the attribute i.

Customer's sentiment intensity revealed from online reviews can also be used to measure the performance of product attributes. We propose to use $\overline{SF_i}$ as another metric of performance in IPA as shown in Equation (10):

$$\overline{SF_i} = \frac{SF_i - {SF_i}^{MIN}}{{SF_i}^{MAX} - {SF_i}^{MIN}} \tag{10}$$

$\overline{SF_i}$ is a normalization result. SF^{MAX} and SF^{MIN} represent the largest and smallest values of the SF_i (i.e., the overall sentiment intensity of product attribute i) among all attributes.

3. A CASE STUDY OF PASSENGER VEHICLES

To demonstrate the proposed approach, a case study of passenger vehicles is presented. We collected online reviews for two typical sport utility vehicles, denoted as SUV A (Buick Envision), SUV B (Toyota Highlander), from a popular auto forum in China (autohome.com). The sales of SUV A is higher than the sales of SUV B in China. All reviews are posted between January 2019 and January 2022. Each customer review includes ratings on eight dimensions (space, power, fuel consumption, comfort, appearance, interior design, value for money, maneuverability), and customer's opinions and thoughts. For a fair comparison between two vehicles, 3000 reviews were randomly sampled for each vehicle model. We collaborate with a data service company to obtain the maintenance records of these two vehicle models during the same time period, and 4500 records were randomly sampled for each vehicle model. Each maintenance record contains vehicle information (e.g., fuel type, mileage) and specific maintenance items (e.g., filter change, engine check). In the following sections, the results of attribute identification, sentiment analysis and Kano-IPA analysis are presented.

3.1 Attribute identification and sentiment analysis.

Following the procedures described in Sec. 2, after preprocessing the online reviews and maintenance records, word embeddings of key words from these two data sources are created. Then the word vectors are clustered into product attributes using the X-means algorithm. Based on the functional structure of motor vehicle system, these attributes are further corrected. Finally, twelve vehicle attributes that are most concerned by customers are identified.

TABLE 2: PRODUCT ATTRIBUTES IDENTIFIED FROM ONLINE REVIEWS AND MAINTENANCE RECORDS

Attribute	Sample keywords	Num
Power (A_1)	Power, speed...	40
Fuel (A_2)	Fuel, oil...	46
Control (A_3)	Brake, clutch…	36
Electronics (A_4)	Electron, Sensor…	44
Operation (A_5)	Tire, vibration…	41
Conditioner (A_6)	Heat, cool…	24
Interior (A_7)	Material, chair…	25
Comfort (A_8)	Shake, smell…	28
Light (A_9)	Fog light, light…	35
Structure (A_{10})	Door, bumper…	53
Exterior (A_{11})	Shape, captain…	23
Maintenance (A_{12})	Purifier, seal…	53

Table 2 shows the identified vehicle attributes with sample keywords, and Num represents the number of extracted keywords for each attribute. Some product attributes reflect the performance of the automotive mechanical system. For example, Power and Fuel are more related to the engine performance. Control and Operation are related to the braking property, and shock absorption of the vehicle chassis. In addition, customers' perceived attributes are also included. For example, the Comfort contains keywords such as noise level, smell, etc. Exterior refers to the shape design of the vehicle, such as aesthetics and fashion style.

After obtaining the identified attributes and generating corresponding semantic groups as described in Sec. 2.2, we utilize SKEP model to analyze the sentiment polarity of all semantic groups. Table 3 shows sample results of the sentiment analysis on customers' 3000 reviews on SUV A. S_{ij}^{pos} and S_{ij}^{neg} represent the sentiment polarity of product attribute i in review j. y_j is the classified rating, $y_j = 0$ represents the customer's rating is lower than average rating, while $y_j = 1$ represents the opposite situation.

TABLE 3: SAMPLE RESULTS OF THE SENTIMENT ANALYSIS ON CUSTOMERS' REVIEWS ON SUV A

	Sentiment Intensity of Product Attribute i ($i = 1, 2, \ldots, 12$)					
Review ID (j)	S_{1j}^{pos}	S_{1j}^{pos}	…	S_{12j}^{pos}	S_{12j}^{pos}	Classified rating y_j
1	1	0	…	1	0	0
…	…	…	…	…	…	…
3,000	0	1	…	1	0	1

3.2 Results of Kano Analysis

Following the procedures in Sec.2.3, the values of average positive (S_i^{pos}), negative (S_i^{neg}) and overall (SF_i) sentiment intensity of product attribute i can be calculated, and the results are shown in Table 4. The SF_i index can reflect the average customers' satisfaction with the vehicle model. The SA_i index is the ratio of the mean positive sentiment intensity to the mean negative sentiment strength compared to the overall sentiment intensity of attribute A_i. $\overline{SA}$ is the average value of SA_i index of all product attributes of this model. β is the partition divided by calculating SA^{MAX} and SA^{MIN}. Type is the three Kano categories of product attributes based on the rules defined in Sec. 2.3.

As we can observe from Table 4, most of the SUV A's attributes are categorized as performance and excitement attributes, while SUV B has more basic attributes. This may indicate that SUV A is generally favored in many aspects, which is consistent with the sales difference of these two vehicle models. The attributes with consistent categorizations between two SUVs are Power, Fuel, Electronics, Conditioner, Interior, Maintenance, and the inconsistent attributes include Control, Operation, Comfort, Light, Structure and Exterior. This may indicate that the design of attributes like Comfort and Light shows more variations between two SUVs. By comparing the average satisfaction SF_i, we find SUV A is higher than SUV B in Power, Control, Electronics, Operation, Interior, Comfort, Light, Exterior and Maintenance. For SUV B, these attributes are needed to be optimized with higher priority to improve customers' satisfaction.

3.3 Results of Importance and Performance Analysis.

Following the procedures in Sec. 2.4, we build a logistic regression model to calculate the importance of product attributes. The input variables are the sentiment intensities for 12 vehicle attributes, and the output variable is the overall vehicle ratings from customers. The estimated weight of each attribute is shown in Table 5. We add "+"and "-" after each attribute to distinguish the positive and negative polarity of the attribute. We find most "+" attributes have positive weights, while the contributions of ""-" attributes are generally negative, which is consistent with our expectation that if a customer gives a positive review for an attribute, this attribute generally favors the customer's rating.

TABLE 4: RESULTS OF KANO ANALYSIS ON TWO VEHICLE MODELS. (B: BASIC ATTRIBUTE, P: PERFORMANCE ATTRIBUTE, E: EXCITEMENT ATTRIBUTE)

	SUV A					**SUV B**				
Attribute	S_i^{pos}	S_i^{neg}	SF_i	SA_i	Type	S_i^{pos}	S_i^{neg}	SF_i	SA_i	Type
Power	0.779	0.106	0.673	0.187	P	0.680	0.106	0.575	0.225	P
Fuel	0.595	0.224	0.370	1.537	P	0.648	0.141	0.507	0.385	P
Control	0.542	0.171	0.371	0.856	P	0.419	0.182	0.236	3.377	E
Electronics	0.242	0.125	0.116	-13.926	B	0.165	0.144	0.021	-1.168	B
Operation	0.326	0.123	0.203	1.548	P	0.292	0.113	0.178	1.744	E
Conditioner	0.164	0.079	0.086	11.238	E	0.160	0.055	0.106	1.072	E
Interior	0.857	0.087	0.770	0.127	P	0.643	0.179	0.463	0.631	P
Comfort	0.542	0.113	0.429	0.358	P	0.423	0.142	0.281	1.014	E
Light	0.222	0.103	0.119	6.596	E	0.130	0.211	-0.081	-0.723	B
Structure	0.804	0.118	0.686	0.208	P	0.869	0.067	0.802	0.092	B
Exterior	0.716	0.067	0.648	0.116	P	0.662	0.055	0.608	0.099	B
Maintenance	0.717	0.109	0.608	0.219	P	0.590	0.144	0.445	0.475	P
$\overline{SA}$				0.755					0.602	
SA^{MAX}				11.238					3.377	
SA^{MIN}				-13.926					-1.168	
β				2.097					0.379	

TABLE 5: THE RESULTS OF LOGISTIC REGRESSION BETWEEN CUSTOMERS' OVERALL RATING AND SENTIMENT LEVELS OF PRODUCT ATTRIBUTES

Variable	SUV A Estimate	SUV B Estimate	Variable	SUV A Estimate	SUV B Estimate
Intercept	-0.301	-0.711***	Interior+	0.294**	0.345**
Power+	0.395**	0.198.	Interior-	-0.156	0.109
Power-	-0.315	-0.337*	Comfort+	0.001	0.126
Fuel +	0.252*	0.391***	Comfort-	-0.363	-0.476***
Fuel-	-0.180	-0.076	Light+	0.110	0.143
Control+	0.188	0.669***	Light-	-0.266	-0.148
Control-	-0.622***	-0.175	Structure+	0.050	-0.353*
Electronics+	0.417***	0.442***	Structure-	-0.557**	-0.530*
Electronics-	-0.052	-0.026	Exterior+	0.224*	0.367***
Operation+	0.316**	0.238*	Exterior-	-0.360*	-0.625**
Operation-	0.060	-0.383**	Maintenance+	0.130	0.341***
Conditioner+	0.350**	0.532***	Maintenance-	-0.538***	-0.280*
Conditioner-	0.026	-0.302	AIC	3606.000	3634.800

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. AIC is used to evaluate the goodness of model fitting. The smaller, the better.

TABLE 6: IMPORTANCE AND PERFORMANCE OF THE TWELVE PRODUCT ATTRIBUTES FOR TWO VEHICLE MODELS

Variable	SUV A				SUV B			
	Imp	P_i	$\overline{P_i}$	$\overline{SF_i}$	Imp	P_i	$\overline{P_i}$	$\overline{SF_i}$
Power	0.710	671	0.211	0.858	0.535	287	0.700	0.743
Fuel	0.432	539	0.389	0.415	0.467	68	1.000	0.666
Control	0.810	655	0.233	0.417	0.844	522	0.377	0.359
Electronics	0.469	828	0.000	0.044	0.468	797	0.000	0.116
Operation	0.376	632	0.264	0.171	0.621	514	0.388	0.293
Conditioner	0.376	182	0.869	0.000	0.834	126	0.920	0.212
Interior	0.450	85	1.000	1.000	0.454	134	0.909	0.616
Comfort	0.364	267	0.755	0.501	0.602	279	0.711	0.410
Light	0.376	195	0.852	0.048	0.291	200	0.819	0.000
Structure	0.607	308	0.700	0.877	0.883	507	0.398	1.000
Exterior	0.584	—	—	0.822	0.992	—	—	0.780
Maintenance	0.668	454	0.503	0.763	0.621	278	0.712	0.596

Table 6 shows the calculated importance and performance of the twelve attributes for two vehicles following Equations (8) – (10). Since Exterior does not have matching items in the maintenance records, its performance is not calculated. For SUV A, the most important three attributes are Power, Control, and Maintenance. According to the P_i, the top three attributes that perform best are Interior, Conditioner, and Light. The three worst-performing attributes of SUV A are Electronics, Power and Control. Power and Control are of highest importance but their performance is relatively poor. It is suggested that the manufacturer of SUV A pays more attention to the engine system and chassis system related to these two attributes. For SUV B, Control, Structure, Exterior have the largest values of importance. Fuel, Conditioner, Interior are top three attributes that perform best among all attributes. Control, Electronics, Operation perform worst. Thus, it is suggested to prioritize optimizing the design of body and chassis system for SUV B.

Figure 4 and Figure 5 show the IPA Plots for SUV A with performance generated from maintenance records (see Equations (8)-(9)) and online reviews (see Equation (10)), respectively. Figure 6 and Figure 7 show the IPA Plots for SUV B with performance generated from two data sources. The dividing lines in blue color are obtained based on the average value of importance and performance. The resulted four quadrants contain eleven attributes that appear in both online reviews and maintenance records. We use squares, rhombuses and stars to represent the basic attributes, performance attributes and excitement attributes identified in the previous KANO analysis, respectively.

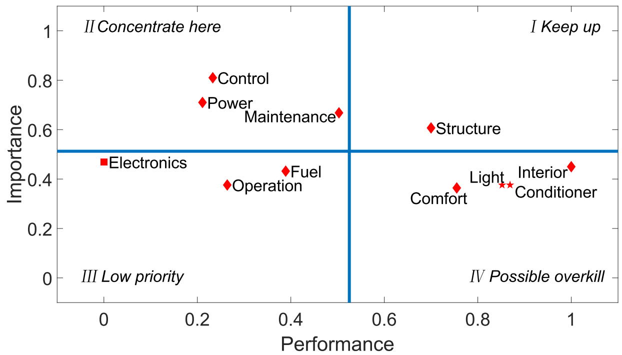

FIGURE 4: IPA PLOT FOR SUV A WITH PERFORMANCE GENERATED FROM MAINTENANCE RECORDS

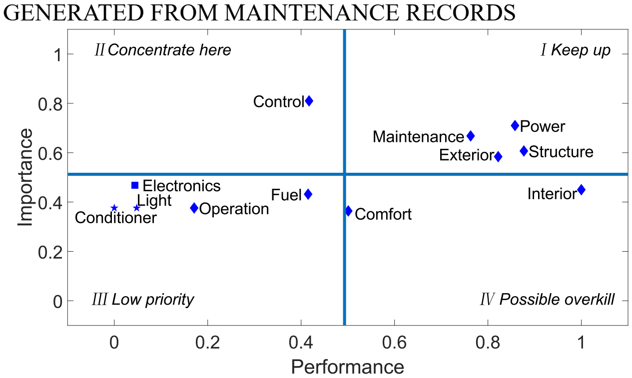

FIGURE 5: IPA PLOT FOR SUV A WITH PERFORMANCE GENERATED FROM ONLINE REVIEWS

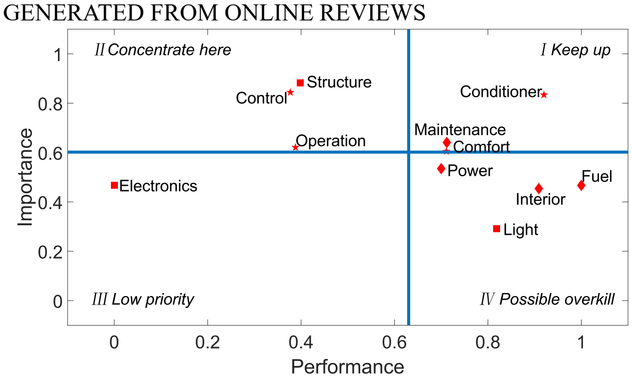

FIGURE 6: IPA PLOT FOR SUV B WITH PERFORMANCE GENERATED FROM MAINTENANCE RECORDS

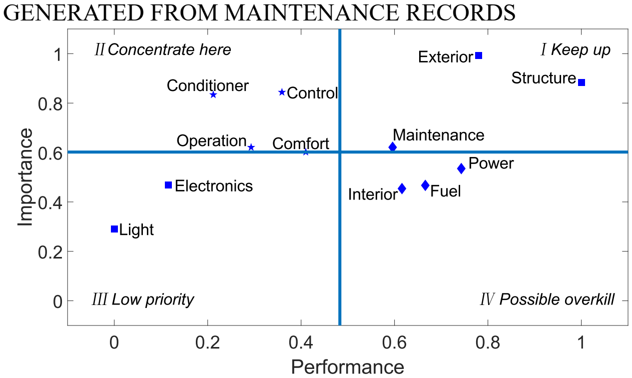

FIGURE 7: IPA PLOT FOR SUV B WITH PERFORMANCE GENERATED FROM ONLINE REVIEWS

As shown in the IPA plots, attributes in the Q1 quadrant should to be maintained, such as the Structure of SUV A, and the Conditioner, Maintenance, and Comfort of SUV B. Since customer satisfaction is related to the performance of product attributes, manufacturers should prioritize improving low-performance product attributes, i.e., Q2 quadrant attributes. Prioritized improving attributes for SUV A include Control, Power and Maintenance, and for SUV B, Structure, Control, and Operation are more important. Electronics in the Q3 quadrant is a low-priority optimization attribute for both SUV A and SUV B. Attributes in the Q4 quadrant are over-optimized. Manufacturers may consider relative adjustment of these attributes to maximize profits.

In addition, we are interested in whether the IPA results differ between two data sources, i.e., online reviews and maintenance records. For SUV A, we find that Structure (Q1 quadrant) and Control (Q2 quadrant) are consistent between Figure 4 and Figure 5. This indicates that the Structure of SUV A has reliable performance from both two data sources, and Control is the attribute to be improved. However, we find that Maintenance and Power belong to Q2 quadrant in Figure 4, but Q1 quadrant in Figure 5. Similarly, for SUV B, the Structure and the Conditioner are switching quadrants between Figure 6 and Figure 7. This result shows that customers' reviews do not always match with the real performance of a vehicle model obtained from the maintenance records. Certain product attributes need further inspection to get accurate evaluations of their performance. Integrated data (e.g., customer reviews and maintenance records) can provide more comprehensive understanding on the product's performance and guide the design improvement.

4. CONCLUSION

In this paper, we propose an approach for conducting Kano-IPA analysis of product attributes from online customer reviews and product maintenance records synthetically. Compared with previous research using single data source, our method integrates the objectiveness of product maintenance records with the subjective evaluation of customer reviews in Kano-IPA analysis. The textual information from both data sources can be processed in a unified and systematic framework. To smooth the sentiment analysis of customer reviews, we develop the method of semantic group generation. Specifically, to overcome the issue of processing sentences with multiple identified product attributes, we propose a dependency parsing-based method to facilitate the separation of descriptions in such sentences, which can further improve the accuracy of text analysis.

A case study of passenger vehicles in China's auto market is presented to demonstrate the proposed approach. Two typical SUVs with different sales are selected for Kano and IPA analysis. Our results show that the SUV model with lower sales owns more basic attributes from the Kano analysis, which have the highest priority in design improvement. The results of IPA show that the attributes in the Q2 quadrant (i.e., concentrate here) should be taken more care of. These results can provide designers with design insights such as which product attributes should be concerned with priority.

In addition, we find that the IPA results differ between online reviews and maintenance records, which indicates that customers' reviews do not always match with the real performance of a vehicle model. Thus, manufacturers should not rely solely on online reviews, and integrated data can provide more comprehensive understanding on the product's performance and guide the design improvement. Although vehicle data is used this case study, our proposed approach can be extended to the attribute analysis of other products, such as consumer electronics and engineering machinery.

One limitation of this study is that the Kano analysis and IPA are performed with data in a fixed time period, and the temporal change is not reflected, especially those latest trends in auto market. In the future, we will explore how to efficiently process online reviews and product maintenance records in time-series and get design insights from their dynamic features. Other novel techniques such as graph neural network can be used to model and predict the relations between customers and products and improve the quality of Kano and IPA analysis.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support from the National Natural Science Foundation of China (52005328) and Shanghai Science and Technology Commission "Yangfan" Program (20YF1419300). The authors would also like to acknowledge the technical support from Yilin Zhang, Zidong Huang, Xinyi Yang.

REFERENCES

[1] Matzler, K., Bailom, F., Hinterhuber, H. H., Renzl, B., and Pichler, J., 2004, "The Asymmetric Relationship between Attribute-Level Performance and Overall Customer Satisfaction: A Reconsideration of the Importance-Performance Analysis," Ind. Mark. Manag., **33**(4), pp. 271–277.

[2] Mikulić, J., and Prebežac, D., 2011, "A Critical Review of Techniques for Classifying Quality Attributes in the Kano Model," Manag. Serv. Qual., **21**(1), pp. 46–66.

[3] Mikulić, J., 2007, "The Kano Model–A Review of Its Application in Marketing Research from 1984 to 2006," … 1st Int. Conf. Mark. Theory …, (Table 1), pp. 1–10.

[4] Lee, Y. C., Sheu, L. C., and Tsou, Y. G., 2008, "Quality Function Deployment Implementation Based on Fuzzy Kano Model: An Application in PLM System," Comput. Ind. Eng., **55**(1), pp. 48–63.

[5] Bi, J. W., Liu, Y., Fan, Z. P., and Zhang, J., 2019, "Wisdom of Crowds: Conducting Importance-Performance Analysis (IPA) through Online Reviews," Tour. Manag., **70**(March 2018), pp. 460–478.

[6] Joung, J., and Kim, H. M., 2021, "Approach for Importance-Performance Analysis of Product Attributes from Online Reviews," J. Mech. Des. Trans. ASME, **143**(8), pp. 1–14.

[7] Yao, M. L., Chuang, M. C., and Hsu, C. C., 2018, "The Kano Model Analysis of Features for Mobile Security Applications," Comput. Secur., **78**(2018), pp. 336–346.

[8] Thakur, R., 2018, "Customer Engagement and Online Reviews," J. Retail. Consum. Serv., **41**(November 2017), pp. 48–59.

[9] Dellarocas, C., 2003, "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms," Manage. Sci., **49**(10), pp. 1407–1424.

[10] Che, W., Feng, Y., Qin, L., and Liu, T., 2021, "N-LTP: An Open-Source Neural Language Technology Platform for Chinese," pp. 42–49.

[11] Che, W., Li, Z., and Liu, T., 2010, "LTP: A Chinese Language Technology Platform," Coling 2010 - 23rd Int. Conf. Comput. Linguist. Proc. Conf., **2**(August), pp. 13–16.

[12] Tian, H., Gao, C., Xiao, X., Liu, H., He, B., Wu, H., Wang, H., and wu, feng, 2020, "SKEP: Sentiment Knowledge Enhanced Pre-Training for Sentiment Analysis," pp. 4067–4076.

[13] Dai, Y., Li, Y., Cheng, C. Y., Zhao, H., and Meng, T., 2021, "Government-Led or Public-Led? Chinese Policy Agenda Setting during the COVID-19 Pandemic," J. Comp. Policy Anal. Res. Pract., **23**(2), pp. 157–175.

[14] Wang, S., Sun, Y., Xiang, Y., Wu, Z., Ding, S., Gong, W., Feng, S., Shang, J., Zhao, Y., Pang, C., Liu, J., Chen, X., Lu, Y., Liu, W., Wang, X., Bai, Y., Chen, Q., Zhao, L., Li, S., Sun, P., Yu, D., Ma, Y., Tian, H., Wu, H., Wu, T., Zeng, W., Li, G., Gao, W., and Wang, H., 2021, "ERNIE 3.0 Titan: Exploring Larger-Scale Knowledge Enhanced Pre-Training for Language Understanding and Generation."

[15] Chen, Y., Zhong, Y., Yu, S., Xiao, Y., and Chen, S., 2022, "Exploring Bidirectional Performance of Hotel Attributes through Online Reviews Based on Sentiment Analysis and Kano-IPA Model," Appl. Sci., **12**(2).