

基于文本挖掘和IPA-KANO模型的电商产品需求层次研究

徐璐

江南大学商学院, 江苏 无锡

收稿日期: 2024年1月3日; 录用日期: 2024年1月11日; 发布日期: 2024年2月29日

摘要

在线评论作为用户偏好表达的主要途径之一, 成为企业发掘顾客产品需求的主要数据源泉。本研究通过文本挖掘在线评论取代传统需求分析中的问卷调查的方法, 采用机器学习算法提取有关产品需求的文本特征, 再基于各特征属性的情感分析结果, 使用改进后的IPA-KANO模型进行电商产品需求分类, 提出产品提升策略。使用京东平台6个猫粮品牌的在线评论对本文提出的研究思路进行验证, 结果表明, 对于适口性、性价比等特征属性, 企业应采取“继续保持”的策略; 对于香味、品牌效应等特征属性, 企业应采取“提升改进”策略。

关键词

IPA-Kano模型, 产品需求, 在线评论, 情感分析

Research on E-Commerce Product Demand Hierarchy Based on Text Mining and IPA-KANO Model

Lu Xu

School of Business, Jiangnan University, Wuxi Jiangsu

Received: Jan. 3rd, 2024; accepted: Jan. 11th, 2024; published: Feb. 29th, 2024

Abstract

As one of the main ways to express user preferences, online reviews have become the main data source for enterprises to explore customer needs. This study replaces the traditional question-

nnaire survey in demand analysis by text mining online review, and adopts machine learning algorithm to extract text features related to product demands. Then, based on the sentiment analysis results of each feature attribute, the improved IPA-KANO model is combined to classify e-commerce product demands and formulate product promotion strategies. The online reviews of 6 cat food brands on Jingdong platform were used to verify the research ideas proposed in this paper. The results showed that enterprises should adopt a “continue to maintain” strategy for features such as palatability and cost performance. For characteristic attributes such as fragrance and brand effect, enterprises should adopt the strategy of “upgrading and improving”.

Keywords

IPA-KANO Model, Product Demand, Online Review, Sentiment Analysis

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

根据 CNNIC 第 52 次《中国互联网发展状况统计报告》表示,截至 2023 我国网络购物用户规模达 8.84 亿人,较 2022 年 12 月增加 3880 万人,占网民整体的 82.0% [1]。而接近饱和的网购渗透率意味着消费者的需求将开始从低端基础服务向高端定制化服务转变,面对市场红海,电商企业更需要围绕顾客需求对自身产品进行更加精细的改进优化[2]。

目前,KANO 与 IPA 被普遍用于满意度相关研究中。Kano 模型基于顾客满意度感知情况来反映不同需求和顾客满意度之间的关系[3]。然而,KANO 模型由于受访者对某要素的感知和其占比的影响,存在主观性的不确定性,因此有时候难以准确反映出该要素的真实满意度[4] [5] [6] [7] [8]。而 IPA 根据用户的满意度及产品的重要度评价满意度,并制定产品/服务的改进优先次序及策略[9],尽管 IPA 广受欢迎,但仍存在若干缺点影响受访者对要素重要性的判断:产品属性重要度和客户满意度之间线性对称关系的错误假设、产品属性重要度测量存在误差、产品属性之间独立的错误假设等[10] [11] [12]。由于 IPA 与 KANO 均基于满意程度对指标进行分析,也有很多学者将两者进行结合[13],但相比机械组合两种模型在一起,Kuo 等提出了一种 IPA-KANO 模型[14],杜煜则利用该模型对地铁出入口设计要素进行了研究。IPA-KANO 模型结合了 KANO 模型对要素进行分类和 IPA 模型判定改进次序的优势,虽然提高了分析效率和决策效果,但分类规则模糊、自述评价度较主观等问题仍然存在。

而针对顾客需求研究,具有规模大且自发性特点的在线评论则为其提供了天然的理想数据,比起基础的调查问卷收集数据,在线评论作为用户直接生成内容更真实地反映了其体验[15],大大提高了研究可信度,已有研究通过在线评论挖掘完成了对消费者偏好的测度[16]-[21]。一些学者在对在线评论进行分析时结合了 Kano 模型来判断顾客需求,如赵宇晴等先将基于模糊理论的 Kano 模型与情感分析结合构建需求-满意量化模型,完成了客户需求的分类[22]。李贺等用 LDA 模型挖掘手机在线评论特征,有效识别需求要素,再根据 Kano 模型设计调查问卷,确定顾客各类需求重要度和供给优先顺序[23]。Bi 等基于 LDA 模型提取产品属性,使用 SVM 模型对产品属性进行情感分析,再使用神经网络模型来分析不同的产品属性对消费者满意度的影响,最后根据 Kano 模型将消费者进行产品属性的需求划分[24]。Xiao 等提出一个基于产品在线评论来衡量消费者需求偏好的经济计量模型。通过对手机在线评论内容的情感分析和改进的有序选择模型来确定消费者的总体偏好,并提出基于边际效应的 Kano 模型,增加了发散型需求、

逆向需求和需创新的需求类型[25]。而孙玲玲等在 LSTM-LDA 融合算法的基础上加入时间参数, 通过在 IPA 模型中计算主题热点的满意度和关注度得到了相关发展趋势[26]。Bi 等则通过 IOVO-SVM 确定用户需求及相应情感值, 以分析结果为依据带入 IPA 模型, 为酒店服务制定个性化改进建议及对策[27]。

基于此, 为改善传统满意度模型分类受主观影响的缺点, 本文结合 Ying F K 等人在 2012 年提出的 IPA-KANO 模型研究成果, 基于在线评论提出一种新的 IPA-KANO 模型改进方法, 旨在优化其指标分类能力, 并通过与在线评论的结合, 在特征挖掘的基础上研究电商产品顾客需求挖掘问题, 并设计对应改进方案。

2. 基于情感分析的 IPA-KANO 模型构建

本研究爬取相关在线评论数据, 并通过逆文档频率和 Word2vec 深入语义层面生成词向量, 利用改进后的 K-means 算法进行特征提取, 结合评论情感分析结果设计 IPA-Kano 模型指标, 最后结合情感度和关注度进行电商产品需求分析, 再针对模型结果给出相应的设计改进。具体设计流程如图 1 所示。

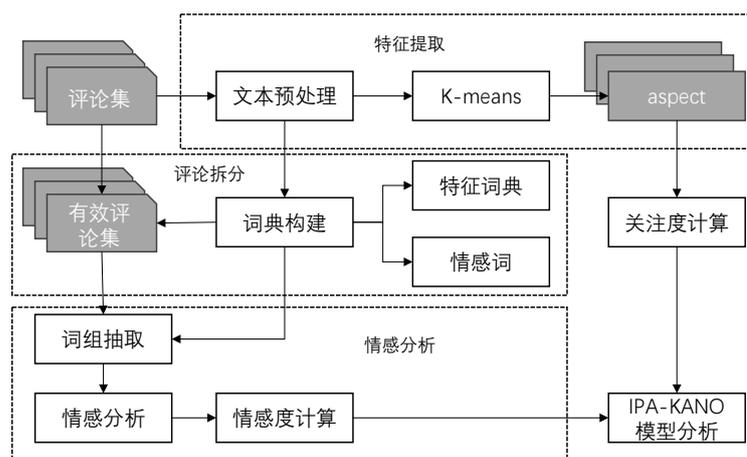


Figure 1. Text mining and requirements analysis process based on online reviews

图 1. 基于在线评论的文本挖掘及需求分析流程

2.1. 改进后的 K-means 特征属性提取

在通过初步的在线评论数据预处理后, 在使用 K-means 进行文本聚类时, K-means 的聚类结果受初始簇中心的随机确定的巨大影响, 聚类结果有时可能会收敛到局部最小, 产生空簇, 所以为防止聚类过迭代, 对传统 K-means 确定 k 簇的方法做出了改进, 简单概括就是: 依次选取 k 个簇中心, 判定依据为距离之前选定的簇中心越大的点被选为下一个簇中心的可能性越高。具体步骤为:

- 初始化质心, 从输入集合中随机选择一点作为第一个聚类中心;
- 对于剩余每一点 x_i , 计算它与最近聚类中心的距离 T_{xi} ;
- 为使聚类中心尽可能分散, 选择 T_{xi} 最大的数据点作为新聚类中心, 且 T_{xi} 越大, 被选取作为下一个聚类中心的概率越大:

$$P_{xi}^d = \frac{t_{xi}}{\sum_{i=1}^n t_{xi}} \quad (1)$$

其中, P_{xi}^d 为点 x_i 被选中为下一轮循环的簇中心的概率, t_{xi} 为每个点到所有类簇中心的最短距离。

- 重复上两步直至数据集中所有样本点到所属聚类质心的距离之和 J 为最小, 聚类结束:

$$J = \min \sum_{i=1}^n \|x_i - k_c^i\|^2 \quad (2)$$

经过这一改进后,可以有效降低簇内误差平方和得到的损失值,使数据最大化地收敛到全局最优减少迭代,还可以避免出现空簇。

2.2. 情感分析

本文基于 K-means 聚类后产品每个主题所表示的若干个特征属性词来组成产品的属性词典。先按基础标点符号把顾客在线评论划分为单元句,根据在线评论的语言特征,一般一个单元句里只包含消费者对一个产品属性的评价。再选用知网 HowNet 情感词典作为构建情感词典的基础词典,并在此基础上增加评论领域相关的词汇,情感词典中标记了情感词的极性和强度。其中,程度副词参考不同文献的研究成果,本文将程度副词强度划分为 6 个层级,否定副词划分为 2 个层级[28] [29] [30]。

以<特征词,情感词>词组的抽取在词标注的基础上,参考文献[31] [32] [33]情感分析算法设计进行情感强度计算。抽取出的词组中包括特征词、情感词以及与情感词相邻的副词(如果存在)。特征情感值的计算以单元句为基本单位,再把同属一条评论的属性情感进行汇总,具体公式如下:

$$S_{i_pos} = \sum_{j=1}^{n(S_i)} em_{pos}(S_i, j) \times adv \quad (3)$$

$$S_{i_neg} = \sum_{j=1}^{n(S_i)} em_{neg}(S_i, j) \times adv \quad (4)$$

其中, S_{i_pos} 表示评论中对第 i 个特征的正向情感值, $n(S_i)$ 表示 i 个特征的特征词总数, $em_{pos}(S_i, j)$ 表示抽取出的<特征词,情感词>词组中属性 i 的第 j 个特征词的正向情感值,由词组中情感词的情感强度来衡量。若词组中有否定副词则按负向情感计算。 adv 代表词组中的程度副词强度,若词组中无程度副词,则 $adv = 1$ 。

S_{i_neg} 表示评论中对第 i 个特征词的负向情感值, $em_{neg}(S_i, j)$ 表示<特征词,情感词>词组中特征 i 的第 j 个特征词的负向情感值,由词组中情感词的情感强度来衡量。

2.3. IPA-KANO 模型构建

假定 F 表示产品集合,即 $F = \{1, 2, 3, \dots, f\}$; I 表示产品特征属性的集合,即 $I = \{1, 2, 3, \dots, I\}$; i 表示产品特征属性下的特征词,即 $i = \{1, 2, 3, \dots, i\}$; T 表示评论的集合,即 $T = \{1, 2, 3, \dots, t\}$,则 F_I^k 为第 k 款产品的第 I 个特征属性, T_I^k 为第 k 款产品的第 I 个特征属性上评论数量,其中 $k \in F$ 。

为更契合在线评论文本数据特点,本文引入特征属性 I 的情感度 Sen_I 作为模型隐性重要性判断依据, Sen_I 越高,则特征属性 I 上顾客整体满意度越高,反之则越低。

$$Sen_I = \sum_{i=1}^{ni} \frac{S_{i_pos}}{(S_{i_pos} + S_{i_neg}) \times ni} \quad (5)$$

$$\bar{s} = \frac{1}{n} \sum_{I=1}^n Sen_I \quad (6)$$

其中, ni 为含有特征词 i 的评论数, $ni \in T_I^k$, $n \in I$ 。

引入消费者对指标 I 关注度作为显性重要性判断依据,先计算产品特征占比[34],即指标 I 在不同款式产品中被关注的平均程度,具体计算公式为:

$$\sigma_I = \frac{1}{F} \sum_{k=1}^f \frac{T_I^k}{T^k} \quad (7)$$

式中, T^k 为第 k 款产品所有特征属性上评论短句总数, T_I^k 表示关于第 k 款产品的第 I 个特征的评论数。

通过计算产品特征占比，可以看出在一系列产品中，该产品相应特征属性受顾客关注的关注水平。

基于上述产品属性占比数据，通过依次比较不同产品中的特征属性关注度，可以得到该产品特征属性的整体关注度 Att_I ，即评论中特征属性 I 在各款产品评论中被关注程度的均值：

$$Att_I = \sum_{f=1}^F \frac{\varphi_I^k}{F} \quad (8)$$

$$\bar{a} = \frac{1}{n} \sum_{i=1}^n Att_I \quad (9)$$

φ_I^k 为常数，若第 I 个产品特征在第 k 款产品的评论中占主要地位，则 $\varphi_I^k = 1$ ，反之为 0：

$$\varphi_I^k = \begin{cases} 1, T_I^k / T^k > \sigma_I \\ 0, T_I^k / T^k < \sigma_I \end{cases} \quad (10)$$

可以看到， Att_I 越高，则特征属性 I 上整体关注度越高，即该特征越受消费者关注，反之则越低，越不受消费者关注。

根据 KANO 模型的需求定义及 IPA 模型的矩阵，以归一化处理后的 Sen_I 为竖轴， Att_I 为横轴构建 IPA-Kano 的象限图，具体如图 2 所示。

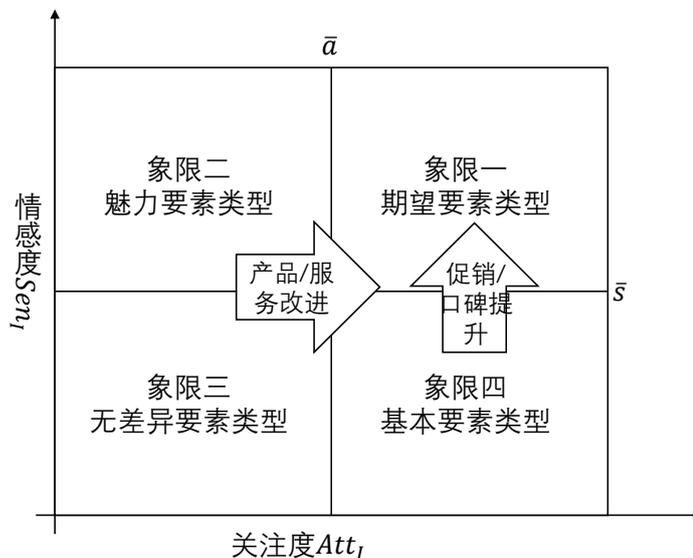


Figure 2. IPA-KANO model matrix diagram

图 2. IPA-KANO 模型矩阵图

(1) 象限一：期望要素类型。顾客对该象限的被测特征属性的关注度和情感度都很高，针对该区域，应继续保持高产品口碑和维持产品质量，保持顾客体验。

(2) 象限二：魅力要素类型。位于该象限的被测特征属性，顾客关注度较低，情感度很高，所以应采取根据消费者具体评价提升和改进产品细节质量的策略，逐步向象限一发展。

(3) 象限三：无差异要素类型。位于该象限的被测特征属性的关注度和情感度均低，应通过提升口碑或质量先向象限二或四优化靠拢，但在企业资源有限的情况下应先减少关注。

(4) 象限四：基本要素类型。位于该象限的被测特征属性的顾客关注度较高，情感度较低，可能过度表现，在保持高关注度的同时，可以采用更新、促销等手段提高顾客的实际体验。

3. 实证分析

本文选择具有代表性的宠物猫粮产品作为在线评论文本分析的研究对象，通过市场调查顾客大致品牌倾向，再运用基于 Python 开发的八爪鱼爬虫软件，分别获取京东网络购物平台上六个品牌的宠物猫粮在线评论，通过初步筛选，最终选出共 22,984 有效评论，评论时间约从 2018 年至 2023 年，如表 1 所示。

Table 1. Jingdong domestic cat food crawling data source
表 1. 京东国产猫粮爬取数据源及部分评论信息

品牌	产品	评论数	评论示例	评论星级
纯福	纯福无谷全期酶解幼猫成猫全价猫粮	4032	家里猫咪很喜欢吃,买了好几次了,吃过后长胖了,也不会有软便什么的,挺好的,配料表也干净,没有乱七八糟的东西,后续可以回购大包的了!	5
鲜朗	鲜朗低温烘焙猫粮无谷幼猫成猫全阶段	3959	吃的第二包啦,大便之类的都很正常,长肉也还 OK,最近我家猫咪食欲大增,马上又要买了,因为在定时喂养,所以也倒掉了一部分。	5
诚实一口	诚实一口 BK01 高蛋白全期全价增肥发腮烘焙猫粮	3527	猫猫们很喜欢吃,颗粒比较小,家里一大一小两只猫都很适合吃,吃了没有软便,切换的也比较顺畅,自然香味,没有刺鼻香精味。	5
蓝氏	蓝氏猎鸟乳鸽低温烘焙猫粮成猫幼猫粮全价鲜肉粮	4026	先买了试吃装,猫咪很喜欢吃,大小正好,烘焙粮,很干爽,不容易黑下巴,不添加肉粉,还添加了布拉氏酵母菌活菌,促进肠道健康,90%都是鲜肉+内脏,配料表干净,给毛孩子吃的也放心。	5
阿飞和巴弟	阿飞和巴弟 E76 益生菌冻干双拼猫粮	3603	这个猫粮真的很不错,适口性特别的好,颗粒不大不小刚刚好,毛毛本来有点厌食,无精打采的,吃了这个猫粮真的变精神了很多,大家赶紧给我冲啊!真的很好啊啊啊啊。	5
江小傲	江小傲无谷鲜肉全价冻干幼猫成猫专用营养全阶段猫粮	3837	仔细研究了这款猫粮,原料以鸡肉为主,用的是马铃薯,木薯粉等淀粉来黏合,适口性很不错,猫咪很喜欢吃,饭量也大了,除了贵再没有缺点。	5

顾客评论中使用的词汇具有明显的“长尾效应”，因此需要对评论数据进行预处理。首先使用基础 Python 代码对评论文本进行简单去重。其次，由于评论文本主围绕京东商城的某品牌猫粮进行评价，其中“京东”、“猫粮”等词出现的频数较大，但会干扰需求挖掘的准确性，因此同样使用代码去除。对文本进行清洗，最后用 JIEBA 进行分词和词性标注，使用百度中文停用词表对停用词进行删除。

3.1. 特征属性提取

猫粮评论特征词聚类的方法步骤如下：

- (1) 使用 TF-IDF 算法对文本进一步筛选得到词频高符合主题的优质词源。
- (2) 使用 Word2vec 模型，选择 CBOW 模式，windows = 5，size = 20，计算各个特征向量，通过基础语料库训练得到分词向量。
- (3) 基于模型构建中的 K-means 算法对词向量进行聚类，根据 SSE 肘函数算法确定 K = 5，提取聚类特征和每个特征下具体的特征词如图 3 和表 2 所示。

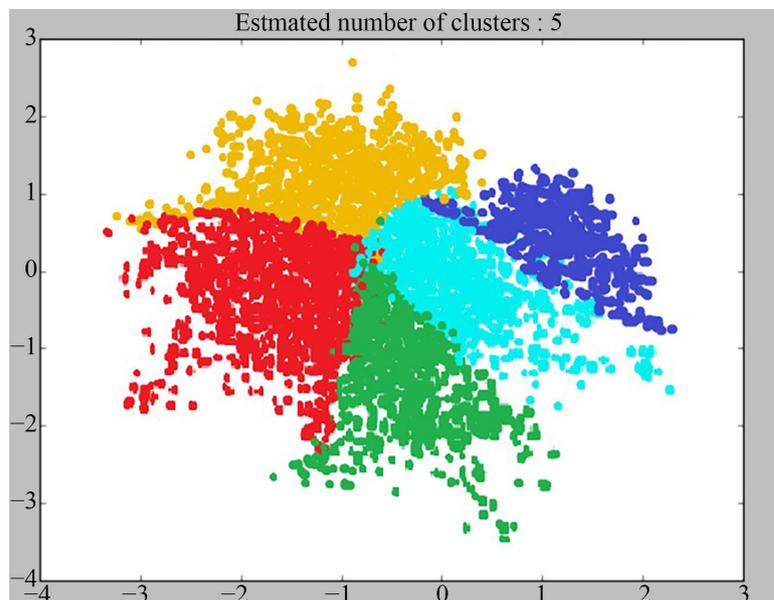


Figure 3. Improved K-means clustering plot
图 3. 改进后的 K-means 聚类图

Table 2. Preview of theme section feature words
表 2. 主题部分特征词预览

编号	主题	特征词
1	口感	适中, 大小, 颗粒, 适合, 合适, 偏软, 香, 香精味, 异味, 自然, 正好, 自然, 不规则……
2	食用反馈	吃, 毛发, 亮泽, 光滑, 肠胃, 拉稀, 排便, 便秘, 呕吐, 肥胖, 油腻, 不爱吃, 爱吃, 顺畅, 掉毛, 淡……
3	包装	包装, 大小, 严密, 漏, 好, 暴力运输, 过度, 难开, 包装垃圾, 难撕, 受潮, 防水, 防潮, 干燥剂……
4	服务	态度, 物流, 客服, 快, 小姐姐, 退货, 快递, 回复, 贴心, 速度, 恶劣, 吵架, 流程, 慢, 拖延, 小哥哥, 速度……
5	产品附加值	性价比, 喜欢, 品牌, 折扣, 赠品, 优惠, 打折, 力度, 信赖, 贵, 推荐, 买一送二, 垃圾, 值得, 便宜……

根据表 1 中每聚类词频前 10 的特征词, 经人工确定各聚类主题名称, 为了更加细分消费者需求, 归纳主题下相应特征属性名称, 分配对应特征词, 详见表 3。

Table 3. Feature attribute overview
表 3. 特征属性总览

特征属性	特征词数	特征属性	特征词数	特征属性	特征词数
适口性 A1	12	肠胃情况 B2	12	售后服务 D2	3
颗粒大小 A2	10	新鲜度 B3	15	物流服务 D3	6
香味 A3	7	产品内包 C1	6	性价比 E1	3
软硬程度 A4	5	产品外包 C2	8	赠品与折扣 E2	5
毛发情况 B1	10	客服服务 D1	5	品牌效应 E3	5

3.2. 情感分析

为对各类特征属性进行 IPA-KANO 模型分析, 本文对聚类主题中的特征词做进一步阐述, 使各个主题中各被测指标要素的涵义更加具体, 并利用 Python 程序代码计算情感值。首先读取对应的情感词典文件, 构建适应研究数据的词典, 其次导入语库将其转化为字典对象, 其中部分情感观点词和强度副词如下表 4、表 5 所示。

Table 4. Preview of selected opinion words

表 4. 部分观点词预览

情感词倾向	权重	部分情感词
积极	1	营养、健康、适口、美味、及时、新鲜……
负面	-1	垃圾、反胃、恶心、掉毛、拉肚子、破损……

Table 5. Preview of some adverbs

表 5. 部分副词预览

副词	权重	部分程度副词
程度副词	1.5	最为、极其、极为、无比、绝对……
	1.4	太、过、过分、过于、格外……
	1.3	很、挺、非常、特别、十分、尤其……
	1.2	真、真的是、那么……
	1.1	较、较为、比较、蛮……
	0.9	还、稍微、略微、有点、有些……
否定副词	-0.8	不、不大、不是很……
	-1.0	不是、没有、并非、从不、绝不……

最后输入预处理后的评论语料计算, 根据公式(3) (4)得到如表 6 的数据结构。

Table 6. Example of comment sentiment analysis results

表 6. 评论情感分析结果示例

评论	A1 _{pos}	A2 _{pos}	...	E3 _{pos}	A1 _{neg}	A2 _{neg}	...	E3 _{neg}
1	2.9	1.3	...	0	0	0	...	0
2	1	0	...	0	0	0	...	0.5
...
n	2	0	...	0	0	-0.3	...	0

3.3. 基于 IPA-Kano 模型的猫粮产品需求分析

基于上述, 根据公式(5)~(8), 计算各特征属性情感度和关注度得分如表 7 所示。

Table 7. Score for each characteristic attribute

表 7. 各特征属性分值

特征属性	关注度 Att_i	情感度 Sen_i	特征属性	关注度 Att_i	情感度 Sen_i
A1	0.8485	1.7687	D1	0.9617	1.9582
A2	0.3433	1.2109	D2	0.6830	-0.1812

续表

A3	0.0128	0.8193	D3	0.9445	1.0427
A4	0.8801	1.4541	E1	0.7404	1.4121
B1	0.6721	0.0243	E2	0.2619	0.0071
B2	0.6907	-0.0704	E3	0.3301	0.4719
B3	0.5387	1.5399	均值:	0.6001	0.9165
C1	0.2008	0.9841			
C2	0.8929	1.3052			

最后结合上述两项指标得分，绘制出基于 IPA-KANO 模型的猫粮产品需求分析象限图，如图 4 所示。

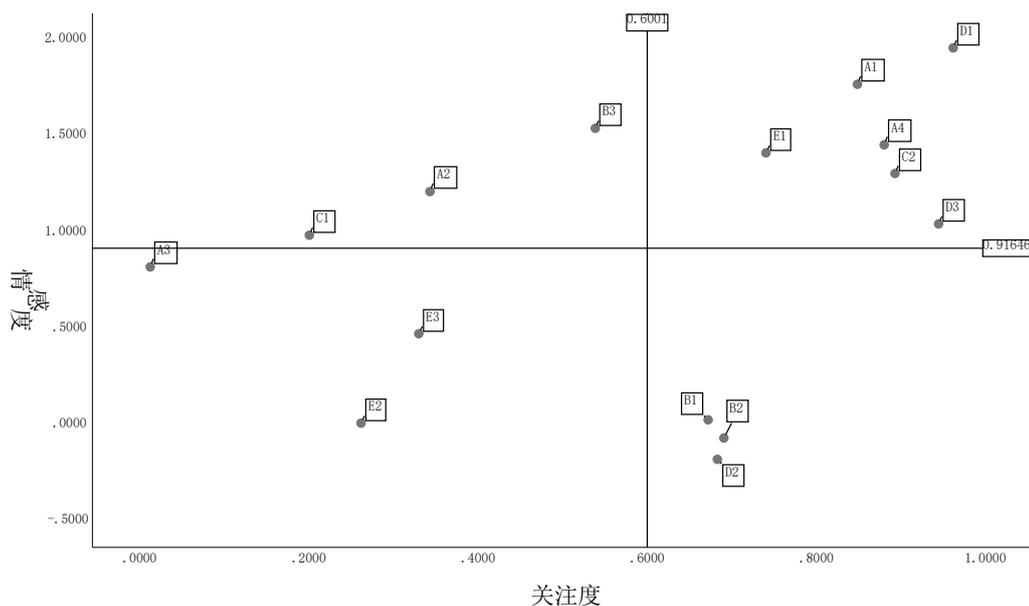


Figure 4. IPA-KANO analysis

图 4. IPA-KANO 分析

图 4 中位于第一象限的特征属性包括：客服服务 D1、适口性 A1、软硬程度 A4、性价比 E1、产品外包 C2 和物流服务 D3。这类指标显性隐性重要性都很高，因此此类指标对该产品消费者来说非常重要，且影响这顾客对产品的整体评价。其中 D1、A3 和 D3 重要性相对最大，表明良好的服务和口感对产品提升顾客的满意度效果最佳，由于这 6 类指标顾客对它们的总体口碑和满意度都很高，所以此类指标需继续保持目前的质量与品质，后续可以高性价比为切入点对产品进行宣传。

位于第二象限的特征属性包括：新鲜度 B3、颗粒大小 A2 和产品内包 C1。该类特征属性的好坏一般不会影响顾客的购买意愿，但提升此类指标的满意度仍有促进销售的巨大价值空间。此类这些用户需求要素是区别于其他同质产品提升用户的满意度的重要突破点，属于顾客潜在需求，产品生产企业若想进一步扩大猫粮生产规模，提高客户粘性，就必须改进位于该象限内的特征，以求在竞争中脱颖而出。在进行产品设计时，企业在第一象限指标均满足的基础上，优先提供高魅力型需求，在资源充足的情况下再考虑低魅力型需求是否提供。例如：在新鲜度方面，国产猫粮可以把猫粮保质新鲜作为吸引顾客的魅力型营销点。同时，针对产品内包等魅力型顾客需求点，商家企业可以加大研发投入，通过改善产品过

度包装、重复包装等问题，使用绿色环保包装等为营销推广的亮点，吸引用户关注。

第三象限的特征属性包括：香味 A3、品牌效应 E3 和赠品与折扣 E2。一般情况下，这类指标隐性和显性重要性均比其他指标低，能产生这种情况的原因有很多，比如猫粮香味对人类来说可能是个重要考虑因素，但对于真正的产品受众群体猫咪来说，香味的刺激并不大；再比如品牌效应指标，由于目前国内猫粮销售市场仍处于国外品牌垄断阶段，能选购国内产品的顾客相对于其他考虑因素，品牌对购买意愿的影响程度就较小等。总体说明上述三类特征属性对改进猫粮产品性能，提高顾客口碑和满意度的作用不大，在企业资源有限的情况下可以先不对这类指标进行改善，但仍应持续关注这些属性，在资源充足的情况下有序改进，以提升市场竞争力。

第四象限的特征属性包括：毛发情况 B1、肠胃情况 B2 和售后服务 D2。该类特征属性在模型矩阵中属于基本要素类型，它们的好坏往往与顾客的购买意愿直接相关，能否进一步提升产品销售额的关键就在其中，需要企业重点关注该类指标的负面评价。而位于该象限的特征属性往往存在顾客总体满意度较高，而部分口碑较低的特点，造成如此不匹配的原因，一方面可能存在商家刷分的线性，另一方面就是购买猫粮的某位消费者对于产品该特征属性的评价及其不满，从而拉低总体情感度得分，尤其是 B2 和 D2，甚至出现了负分，这更加说明该类特征属性的重要性，企业需要对其进行具有针对性的着重改进，争取消灭负面评价，使特征属性向第一象限靠拢。

Table 8. IPA-KANO classification results

表 8. IPA-KANO 分类结果

KANO 要素	特征属性	提升改进	继续保持	KANO 要素	特征属性	提升改进	继续保持	
期望要素	客服服务	-	2	无差异要素	产品内包	2	-	
	适口性	-	6		香味	1	-	
	软硬程度	-	5		品牌效应	1	-	
	性价比	-	4		赠品与折扣	1	-	
	产品外包	-	1		基本要素	毛发情况	6	-
	物流服务	-	3			肠胃情况	7	-
魅力要素	新鲜度	3	-	售后服务	5	-		
	颗粒大小	4	-					

综合上述问题，如表 8 所示，在制定猫粮产品改进策略时，应倡导以提升产品整体质量为主导，以保证那些具有基础性特征属性的主题如口感和食用反馈能满足消费者的需求为主，提升功能性主题质量为辅的产品优化模式。第一，口感方面，维持住原有的猫粮适口性和软硬程度，再根据顾客反馈适当调整猫粮颗粒大小，减少猫粮刺激性香味；第二，食用反馈方面，率先保证产品新鲜程度，减少猫咪食用后出现不良肠胃反映及掉毛的情况，若部分顾客反映还是出现此类情况，需积极做出响应，跟进相关售后服务；第三，包装和服务方面，内包和外包同样重要，前者减少猫粮腐败情况，后者减少货物运损情况，同时，售前和售后也同样重要，是企业获得猫粮产品反馈体验的重要来源，需大幅提升售后服务质量满足顾客需求；第四，产品附加值方面，需制定相应的宣传策略，打造企业品牌，可适当多增加些猫粮产品的折扣活动。

4. 结论

本文以猫粮为例，通过机器学习算法对产品评论进行主题提取及情感分析挖掘需求偏好，再结合联

合分析提出基于在线评论的 IPA-KANO 模型对偏好进行分类, 将消费者对猫粮产品 15 种属性划分为期望、魅力、基本和无差异要素并得到相应改进优先级, 其中适口性、性价比等特征属性, 企业应采取“继续保持”的策略; 对于香味、品牌效应等特征属性, 企业应采取“提升改进”策略。

传统需求分析中的需求分类依赖于 KANO 问卷, 而本文设计了一种文本挖掘在线评论取代问卷调查的方法, 与以往产品需求分析研究相比, 解决了需求确定的主观性及不确定性问题; 引入在线评论消费者对需求的关注度和情感度, 提出一种适用在线评论的 IPA-KANO 模型, 解决了传统需求分析中满意度和重要度定义重复模糊的不足; 以往研究大多止于对产品属性的需求类型的判断[17], 对判断结果如何使用未做进一步分析, 而本文在此基础上, 结合 IPA 模型给出改进方法: 电商企业可以依据评论的关注度能对产品需求实现细分, 资源有限条件下, 利用直观的消费者情感度能获取产品改进最优路径, 对企业利益与消费者满意度协同最优有积极的现实意义。

同时, 本文也存在以下不足: 第一, 由于网络评价的特殊性, 爬取的产品评论中可能混有部分虚假评论未筛除。第二, 在线评论除纯文本外还包含表情包等多形式评论, 目前本文没有涉及。后续研究中需要继续改进机器识别算法, 并尝试纳入其他形式的数据作为研究主体, 为多形态数据下需求研究提供新方向。

参考文献

- [1] 中国互联网络信息中心. 第 52 次《中国互联网络发展状况统计报告》[EB/OL]. <https://www.cnnic.cn/n4/2023/0828/c88-10829.html>, 2023-08-28.
- [2] 邬适融, 陈洁, 曾艺生, 等. 消费者持续满意度研究——基于快乐适应视角[J]. 南开管理评论, 2011, 14(1): 130-137, 156.
- [3] Liang, S.W., Lu, H.P. and Kuo, T.K. (2014) A Study on Using the Kano Two-Dimensional Quality Model to Evaluate the Service Quality of Government Websites. *Journal of Internet Technology*, **15**, 149-162.
- [4] Xu, Q.L., et al. (2019) An Analytical Kano Model for Customer Need Analysis. *Design Studies*, **30**, 87-110. <https://doi.org/10.1016/j.destud.2008.07.001>
- [5] 张振刚, 罗泰晔. 基于在线评论数据挖掘和 Kano 模型的产品需求分析[J]. 管理评论, 2022, 34(11): 109-117.
- [6] 孟庆良, 何林, 朱慧明, 等. 基于 GM(1,1)模型的 KANO 质量要素分类动态预测方法[J]. 中国管理科学, 2015, 23(9): 139-145.
- [7] Kano, N. (1984) Attractive Quality and Must-Be Quality. *Hinshitsu*, **31**, 147-156.
- [8] 肖泉, 李顺, 张星, 等. 特性视域下在线评论系统设计的用户感知非对称性研究[J]. 情报理论与实践, 2022, 45(1): 159-167.
- [9] Martilla, J.A. and James, J.C. (1977) Importance-Performance Analysis. *Journal of Marketing*, **41**, 77-79. <https://doi.org/10.1177/002224297704100112>
- [10] Geng, X. and Chu, X. (2022) A New Importance-Performance Analysis Approach for Customer Satisfaction Evaluation Supporting PSS Design. *Expert Systems with Applications*, **39**, 1492-1502. <https://doi.org/10.1016/j.eswa.2011.08.038>
- [11] Matzler, K., Bailom, F., Hinterhuber, H.H., et al. (2014) The Asymmetric Relationship between Attribute-Level Performance and Overall Customer Satisfaction: A Reconsideration of the Importance Performance Analysis. *Industrial Marketing Management*, **33**, 271-277. [https://doi.org/10.1016/S0019-8501\(03\)00055-5](https://doi.org/10.1016/S0019-8501(03)00055-5)
- [12] Chen, Y., Zhong, Y., Yu, S., et al. (2022) Exploring Bidirectional Performance of Hotel Attributes through Online Reviews Based on Sentiment Analysis and Kano-IPA Model. *Applied Sciences*, **12**, Article 692. <https://doi.org/10.3390/app12020692>
- [13] 孟庆良, 卞玲玲, 何林, 等. 整合 Kano 模型与 IPA 分析的快递服务质量探测方法[J]. 工业工程与管理, 2014, 19(2): 75-80, 88.
- [14] Kuo, Y.F., Chen, J.Y. and Deng, W.J. (2012) IPA-Kano Model: A New Tool for Categorising and Diagnosing Service Quality Attributes. *Total Quality Management & Business Excellence*, **23**, 731-748. <https://doi.org/10.1080/14783363.2011.637811>
- [15] 李宗伟, 张艳辉, 栾东庆. 哪些因素影响消费者的在线购买决策?——顾客感知价值的驱动作用[J]. 管理评论, 2014, 16(1): 109-117.

- 2017, 29(8): 136-146.
- [16] Han, Y. and Moghaddam, M. (2021) Eliciting Attribute-Level User Needs from Online Reviews with Deep Language Models and Information Extraction. *Journal of Mechanical Design*, **143**, Article ID: 061403. <https://doi.org/10.1115/1.4048819>
- [17] 刘岩, 张艳荣, 赵志杰, 等. 基于情感分析的酒店顾客满意度评估[J]. 计算机应用与软件, 2020, 37(2): 54-60, 67.
- [18] Kang, Y. and Zhou, L. (2019) RubE: Rule-Based Methods for Extracting Product Features from Online Consume Reviews. *Information & Management*, **54**, 166-176. <https://doi.org/10.1016/j.im.2016.05.007>
- [19] 贾丹萍, 靳健, 耿骞, 等. 感性工学视角下的用户需求挖掘研究[J]. 情报学报, 2020, 39(3): 308-316.
- [20] 杜学美, 丁璟好, 谢志鸿, 等. 在线评论对消费者购买意愿的影响研究[J]. 管理评论, 2016, 28(3): 173-183
- [21] Hu, Y. and Chen, K.C. (2022) Protecting Hotel Review Helpfulness: The Impact of Review Visibility and Interaction between Hotel Stars and Review Ratings. *Information Journal of Information Management*, **36**, 929-944. <https://doi.org/10.1016/j.ijinfomgt.2016.06.003>
- [22] 赵宇晴, 阮平南, 刘晓燕, 等. 基于在线评论的用户满意度评价研究[J]. 管理评论, 2020, 32(3): 179-189.
- [23] 李贺, 曹阳, 沈旺, 等. 基于 LDA 主题识别与 Kano 模型分析的用户需求研究[J]. 情报科学, 2021, 39(8): 3-11, 36.
- [24] Bi, J.W., Liu, Y., Fan, Z.P. and Cambria, E. (2019) Modelling Customer Satisfaction from Online Reviews Using Ensemble Neural Network and Effect-Based Kano Model. *International Journal of Production Research*, **57**, 7068-7088. <https://doi.org/10.1080/00207543.2019.1574989>
- [25] Xiao, S., Wei, C.P. and Dong, M. (2016) Crowd Intelligence: Analyzing Online Product Reviews for Preference Measurement. *Information & Management*, **53**, 169-182. <https://doi.org/10.1016/j.im.2015.09.010>
- [26] 孙玲玲, 胡彦蓉, 刘洪久. 基于 LSTM-LDA 算法和 IPA 分析的在线品牌社群用户关注热点研究[J]. 情报杂志, 2021, 40(9): 178-186.
- [27] Bi, J.W., Liu, Y., Fan, Z.P. and Zhang, J. (2019) Wisdom of Crowds: Conducting Importance-Performance Analysis (IPA) through Online Reviews. *Tourism Management*, **70**, 460-478. <https://doi.org/10.1016/j.tourman.2018.09.010>
- [28] 张桂宾. 相对程度副词与绝对程度副词[J]. 华东师范大学学报(哲学社会科学版), 1997(2): 92-96.
- [29] 蔺璜, 郭妹慧. 程度副词的特点范围与分类[J]. 山西大学学报(哲学社会科学版), 2003, 26(2): 71-74.
- [30] 徐琳宏, 林鸿飞, 潘宇, 等. 情感词汇本体的构造[J]. 情报学报, 2008, 27(2): 180-185.
- [31] Shi, W. (2013) Sentiment Analysis of Chinese Microblogging Text Based on Sentiment Ontology. Ph.D. Thesis, Tongji University, Shanghai.
- [32] 周知, 方正东. 融合依存句法与产品特征库的用户观点识别研究[J]. 情报理论与实践, 2021, 44(7): 111-117.
- [33] 李慧, 柴亚青. 基于属性特征的评论文本情感极性量化分析[J]. 数据分析与知识发现, 2017, 1(10): 1-11.
- [34] 王雪, 董庆兴, 张斌. 面向在线评论的用户需求分析框架与实证研究: 基于 KANO 模型[J]. 情报理论与实践, 2022, 45(2): 160-167.