

基于在线评论挖掘的宠物家具产品设计改进方法

黄文倩, 李雪莲*

浙江理工大学, 艺术与设计学院, 浙江 杭州

收稿日期: 2023年3月14日; 录用日期: 2023年5月31日; 发布日期: 2023年6月7日

摘要

为了获取用户对产品的满意度并用于产品的改进设计, 提出了一种基于在线评论的设计改进方法。本文以宠物家具为研究对象, 采用爬虫技术从电商平台获取了大量的在线评论; 通过分词技术和词向量技术对在线评论进行了预处理; 接着构建用户极性词典并借助狄利克雷主题模型确定了产品特征; 最后分析挖掘了产品各个特征需改进程度从而建立了宠物家具评价指标体系, 并找到亟需改进的产品特征, 从而制定产品的改进策略。该方法将数据挖掘与情感分析结合, 量化了产品需改进的程度, 为设计师提供了有针对性的家具产品改进方案。

关键词

在线评论, 文本挖掘, 情感分析, 产品特征, 产品改进, 宠物家具

A Design Improvement Method for Pet Furniture Products Based on Online Comment Mining

Wenqian Huang, Xuelian Li*

School of Art and Design, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: Mar. 14th, 2023; accepted: May 31st, 2023; published: Jun. 7th, 2023

Abstract

In order to obtain user satisfaction with the product and use it for product improvement design, a design improvement method based on online review is proposed. Taking pet furniture as the re-

*通讯作者。

search object and using crawler technology to obtain a large number of online comments from e-commerce platforms, the online comments are preprocessed by word segmentation technology and word vector technology; the user polarity dictionary is constructed and the product characteristics are determined with the help of Dirichlet theme model; the user satisfaction and the degree of improvement of each product feature are analyzed and excavated to establish the pet furniture evaluation index system, and find the product features that need to be improved, so as to formulate the product improvement strategy. This method combines data mining with emotion analysis, quantifies user satisfaction, and provides targeted product improvement direction for designers.

Keywords

Online Review, Text Mining, Emotional Analysis, Product Characteristics, Product Improvement, Pet Furniture

Copyright © 2023 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

由于互联网的发展,传统家具制造业的生存环境发生了改变,在新的消费升级背景下,企业想要快速地对市场做出反应,必须加快产品设计更新的速度。

传统的调研模式,用户的参与度低、调研范围小,反馈的信息不具有代表性,容易使产品的设计方向与用户真实需求偏移;当产品设计制造完成后投放市场,还需要用户测试、用户使用反馈,从而不断完善产品。若采用传统的调研方式,必然会降低产品设计节奏、增加更新周期,加大开发成本。因此,准确快速地获取用户需求信息,跟踪用户的反馈,对产品设计改进具有重要的意义。

随着互联网的普及,用户在电商平台消费后,可以在平台内发表自己对产品和购物过程的真实体验。相比于实验环境下得到的用户反馈信息更加的全面而真实。制造厂商也可以根据网络舆情快速做出有利决策,获得竞争优势。因此,以海量在线评论为信息来源对用户满意度信息进行挖掘和分析,是产品改进设计研究的一个有潜力的方向。

已有学者利用大量在线评论挖掘用户满意度,且发现比其他来源(如查阅文献、问卷调查、专家咨询、用户访谈等)获取用户数据更加可靠。高新勤[1]等人通过网络爬虫技术获取用户在线评论原始语料,从用户评论中挖掘有用信息,通过获取用户极性词典和产品特征构建了设计决策辅助分析模型。成方敏[2]等人通过对在线评论的采集与分析,有效地提取了在线评论中的用户体验要素,并识别出对用户体验影响较大的产品特征和使用环境。张振华[3]通过文本数据挖掘分析消费端的评论数据,构建了木家具顾客满意度测量方法,分析识别出了高度影响用户满意度的产品特征。杨程[4]等人通过爬取手机买家的在线评论,利用 LDA 主题模型确定了手机的产品属性,并通过计算量化了用户对各产品属性的满意度,从而为手机的改进提供可靠的方向。卢春燕[5]通过 LDA 主题模型对通过爬取到的在线评论数据进行聚类,构建关键识别模型,对用户需求进行量化,最后用 Kano 模型识别用户关键需求。Wang W M 等[6]利用 WordNet 词表中的不同词的词性关系,生成了用户情感词典,并以玩具产品为例验证了方法的可行性。贾丹萍[7]等人以产品评论为原始语料利用 word2vec 模型生成用户感性情感词典和产品特征词表,并在此基础上提出特征-感性情感模型。JIAO 等人[8]通过用户在线评论提取用户情感词,并构建了产品属性和用户感知关系的提取方法。刘名成[9]通过 Text Rank 算法及规则构建了产品设计要素池,并以用户满意度和关注度作为产品评价指标。任海兵[10]等人通过关键事件法提取影响用户体验的产品因素,并基于马

田系统对手机进行综合评价。综合上述相关研究可以看出, 大多数学者主要是从在线评论中获取用户对产品的情感态度, 并没有结合在线评论获取用户视角下的产品特征。若将用户视角下的产品特征与设计师视角下的产品特征相结合, 则更能准确地获取用户对于各产品特征的满意度。

本文在前人研究的基础上, 面对宠物家具产品设计, 提出一种基于评论大数据的用户对产品满意度的挖掘方法。通过网络爬虫技术从电商平台上获取用户评论数据源, 利用 word2vec 模型生成用户极性词典, 狄利克雷主题模型确定产品特征, 以此为基础分析产品的需改进程度, 并提出产品设计改进的方向。相比于传统的用户需求挖掘方法, 本文通过评论大数据获取的用户反馈信息, 数据样本大、覆盖范围广、真实性强、实时性强, 可以增加产品开发过程中的用户参与度, 缩短开发周期, 准确而迅速地理解用户需求并找到迎合市场的设计方向。

2. 研究方法

本研究通过将用户评论数据挖掘和用户情感分析结合构建了产品设计改进方法, 主要包括在线评论原始语料构建、数据预处理、有用评论获取、用户评价分析。方法框架图如图 1 所示。

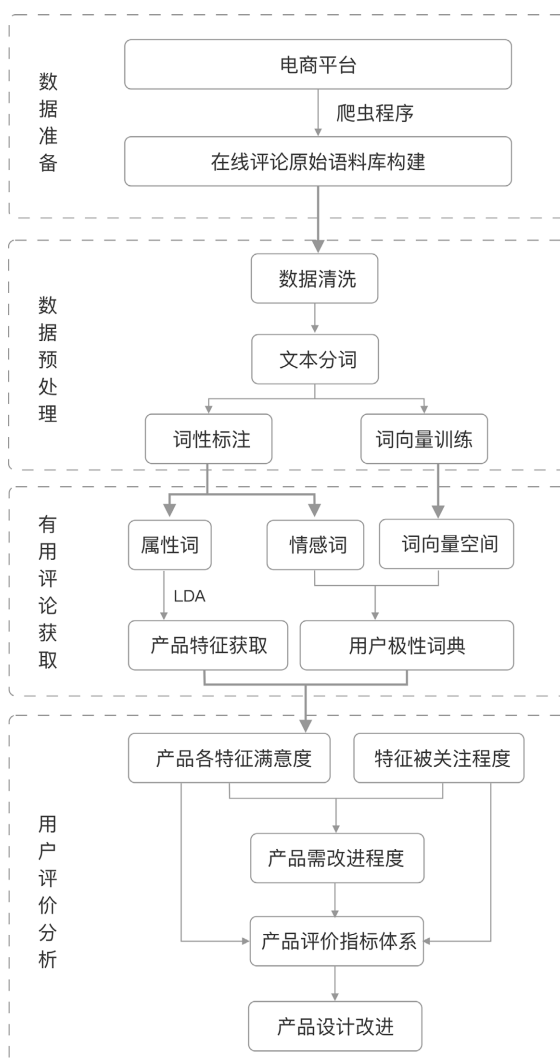


Figure 1. Method framework diagram
图 1. 方法框架图

2.1. 数据准备

各大电商平台是用户反馈产品体验的主要入口, 领先的电商平台具有完善的评价机制, 其评论实时性好, 样本量大, 真实性强, 具有较高的质量。因此可以通过电商平台采用网络爬虫技术获取大量的在线评论作为原始语料库。网络爬虫是一种自动化程序, 它们会按照指定的规则遍历互联网上的各种网页或其他网络资源, 并从中抽取有用的数据。

2.2. 数据预处理

2.2.1. 分词技术

通过网络爬虫技术抓取的数据量庞大, 往往存在很多噪声数据和大量错误, 无法使初始数据具备机器可读性, 同时也会增加数据处理的难度, 影响其结果的准确性和可靠性。因此, 需要对数据先进行筛选、去重、纠错等清洗工作, 再进行中文分词、去除停用词, 词性标注等预处理。常用的文本分词技术有 NLPIR、HanLP 和结巴分词等。本文构建的在线评论原始语料库, 往往存在表述不规范、有错别字等情况, 且文本表达的场景较局限, 都是用户对同一产品的反馈意见。而结巴分词它采用了无向图最大匹配算法, 能够较好地处理新词、歧义等问题, 在中文分词领域具有较高的准确率, 在中小型应用场景下较为适用。对于多数中文评论内容, 分词结果还能够比较好地反映其语义和情感属性。

2.2.2. 词向量技术

在线评论通常是没有固定格式的数据属于为非结构化数据, 词向量技术可以使文本数据可视化和可分析化, 从而更好地理解 and 表示文本中的信息。文本数据向量化就是将文本中的单词或短语映射到一个固定长度的向量空间中, 用向量表示文本中的词语, 将语义上相似的词映射为相似的向量, 进而来实现词与词之间的相互比较与替换和词语间的相似度的计算。常见的词向量模型有(Continuous Bag of Words, CBOW)和跳字模型(Skip-gram), 其中 Skip-Gram 模型能够捕捉到更精细的语义信息, 例如词汇之间的关系, 适合处理词汇库较大的场景, 但是其只考虑单词之间的语义关系。腾讯 AI Lab 在 Skip-gram 模型基础上自研出另一种结构 Directional Skip-Gram (DSG), 这种模型可以同时捕捉单词之间的语义关系和句法关系, 具有更高的综合表达能力。因此, 本文采用 DSG 模型生成在线评论数据的词向量空间。

2.3. 有用评论获取

2.3.1. 产品特征获取

获取特征词的一种常用方法是通过词频统计选取高频词作为特征词, 这种方式选取出来的特征词虽然一定程度上能代表产品特征, 但却会遗漏一些重要的低频词[11], 且在线评论数据量大, 从海量的评论中直接获取有用的信息比较困难, 且不能准确反映产品特征。因此, 需要结合 LDA 模型对在线评论进行产品属性词的抽取, 获得用户关注的产品特征。

LDA (Latent Dirichlet Allocation)是一种文档主题生成模型, 它是一种无监督学习的算法, 是在未知具体类别的属性的情况下, 对所有文档进行聚类的方法[5]。它将文本看成是由不同的主题组成, 每个主题又由一系列相关的单词组成, 通过将文本数据转换成不同主题和单词联系起来的概率分布, 经过大量的文本数据训练, 最终主题相关的文本聚类在一起即可获得其潜在的主题特征。

2.3.2. 用户极性词典构建

极性词指的是文本中具有情感色彩的词汇, 它们可以表示某种情绪、态度、好恶等情感因素, 是用户表达态度和观点的词汇, 通常为形容词。构建极性词典最重要的是对极性词进行情感分析, 具体步骤如下(图 2):

- 1) 从候选情感词汇中筛选同义词、近义词并合并, 得到形容词组。
- 2) 选出高频、有情感倾向的词组作为基准词, 并使用知网(How Net)情感词词库来定义这些基准词的情感倾向和强度。
- 3) 使用 DSG 模型来训练词向量空间, 并计算非基准词和基准词之间的关联程度。
- 4) 用基准词的情感强度作为标准, 计算非基准词的情感强度 E_b 即:

$$E_b = E_a \times E_{ab} \tag{1}$$

式中: E_a 为基准词的情感强度; E_{ab} 为非基准词和基准词的关联度。

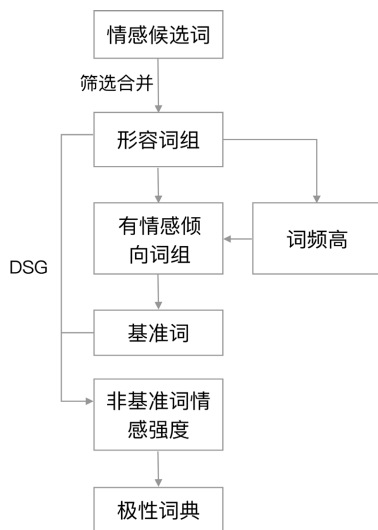


Figure 2. User polarity dictionary construction process
图 2. 用户极性词典构建流程

2.4. 用户评价分析

产品评价指标就是对产品的评判与界定。目的就是在产品开发过程中最大程度的降低风险、避免设计质量的问题和提高产品开发的效率。通过对用户满意度的量化, 明确各产品特征的价值, 根据产品的需改进程度选取符合市场发展和用户需求的设计方案。通过获取到的产品特征和用户极性词典, 计算用户对产品属性的满意度、关注度和改进程度, 建立产品评价指标体系, 分析指标得分, 确定产品改进的优先级和方向。具体步骤如下(图 3)。

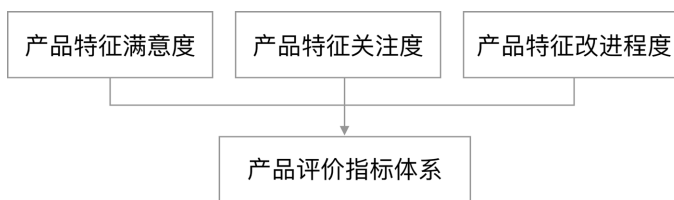


Figure 3. Product evaluation index system
图 3. 产品评价指标体系

2.4.1. 满意度的获取

为了更好的把握对用户的满意度认知, 需将用户对产品特征的需求量化, 通过 python 语言得到与其对应的产品特征满意度评价值。操作步骤具体如下:

- 1) 将产品属性词典和用户极性词典导入 python 中, 并读取评论。
- 2) 将评论拆解成最小句, 找到属性词, 若属性词存在找到与其对应的情感词, 更新属性词的情感强度, 并更新属性词所涉及的评论数。
- 3) 在此基础上可计算用户对该产品属性的满意度 Q 定义如下:

$$Q = \frac{\sum_{i=1}^m (E_i \times P_i)}{\sum_{i=1}^m P_i} + \frac{\sum_{j=1}^n (E_j \times P_j)}{\sum_{j=1}^n P_j} \quad (2)$$

式中: m 为该产品特征正向极性词, n 为负向极性词的数量; E_i 和 P_i 分别为该产品特征的第 i 个正向极性词的情感强度和词频; E_j 和 P_j 分别为该产品特征的第 j 个负向极性词的情感强度和词频。当感性评价值为正值($Q \geq 0$)时, 表明用户对该产品特征持肯定态度, 反之则持否定态度。用户对产品的满意程度取决于 Q 值的大小。

2.4.2. 产品特征关注度

用户对于产品某个特征的满意度很低, 不代表某个特征就最需要改进, 还要结合用户对这个特征的关注程度考虑。因此, 评论中某属性出现的频次, 反映了用户对产品属性的关注程度, 用 T_n 表示用户对第 n 个产品属性的关注:

$$T_n = \frac{t_n}{n} \quad (3)$$

式中: t_n 为第 n 个产品属性的评论个数; n 为总的有用评论数。 T_n 值越大用户的关注度越高。

2.4.3. 产品特征改进程度

用户对产品总的评价度为 $Z = \sum_{i=1}^N (Q_N \times T_N)$ 式中 Q_N 为用户对第 N 个产品属性的满意度, 为了让 Z 值最大化, 每个产品属性的评价度 $Q_N \times T_N$ 都应该达到最大值 1, 因此单个产品属性的评价度可提升空间为 $(1 - Q_N) \times T_N$ 。由于企业资源有限, 为了使利益最大化, 我们应该优先选择评价度可提升空间高的几个产品属性进行改进。即产品属性亟需改进的程度 S_N :

$$S_N = (1 - Q_N) \times T_N \quad (4)$$

3. 应用案例

因为狗、猫等宠物在人类的生活中扮演着越来越重要的角色, 所以宠物家具的市场需求也在逐渐增加。从用户评论中挖掘有用信息, 对改进宠物家具设计提升用户满意度有着重要意义。

3.1. 原始语料库构建

为了获取高质量的用户评论数据, 本研究选择 C2C 为主的淘宝商城作为数据来源平台, 研究宠物家具产品的销售情况。选取产品价位在 150~700 元之间销量较高的猫爬架、猫屋、猫笼产品。使用 Python 语言编写在线爬虫程序, 从淘宝商城在线评论页面获取数据。共采集了网络口碑 26,070 条, 涉及品牌 9 种。并删除无用、重复的评论, 得到有效评论 24,399 条。

3.2. 在线评论预处理

对清洗后的宠物家具用户语料库通过结巴分词工具将连续的句子切分为灵活的词并进行词性的标注, 得到情感候选词和特征候选词, 分词的结果如图 4 所示。使用 DSG 模型训练在线评论的词向量空间。词向量的维度为 100, 迭代训练 50 次, 词向量的上下文最大距离为 3。

情感候选词	特征候选词
('好', 3276)	('空间', 991)
('不错', 1498)	('物流', 878)
('大', 883)	('客服', 804)
('方便', 719)	('结实', 695)
('很大', 477)	('品质', 567)
('小', 372)	('评论', 484)
('合适', 268)	('尺寸', 434)
('简单', 262)	('猫笼', 408)
('高', 248)	('别墅', 396)
('快', 235)	('发货', 357)
('耐心', 221)	('卖家', 350)
('差', 207)	('外观设计', 281)
('宽敞', 199)	('视频', 229)
('舒服', 198)	('商家', 202)
('牢固', 189)	('态度', 190)
('容易', 176)	('店家', 189)
('完全', 162)	('服务态度', 183)
('很', 161)	('购物', 169)
('漂亮', 123)	('材质', 169)
('很漂亮', 116)	('轮子', 161)
('一般', 117)	('货', 144)
('完美', 112)	('盆', 141)

Figure 4. Online comment segmentation results

图 4. 在线评论分词结果

3.3. 产品特征的获取

经过预处理后的文档都是由词组成。本文利用 LDA 主题模型在 python 环境中对文本进行聚类, 辅助挖掘用户需求。LDA 模型的核心是发现潜在群体(例如主题)下的词语[12], 基于 LDA 模型对产品特征后选词进行产品特征词抽取, 结合家具设计评价相关文献、市场调研和专家意见, 得出市场上家具产品特征评价指标, 如表 1 所示, 可知家具的产品特征有 10 项。因此将产品特征个数 K 值设为 10, 得到 10 个特征词集合。在这些分类结果中, 有意思相近的产品特征, 例如: 物流和态度他们都代表着产品的服务, 需要合并, 最终确定了 7 个属性词集合, 确定了每个产品属性对应的属性词数量, 具体信息见表 1。

Table 1. Product evaluation indicators

表 1. 产品评价指标

序号	产品评价指标	评价子指标
1	外观款式	造型、色彩、风格、装饰
2	功能实用	凭椅、储存、支持功能、方便操作、舒适度、社会价值、地位
3	环保	有害物质限量
4	用料	材料成本符合目标定位、耐用性、表现力
5	做工工艺结构	各位构件外观质量、强度、耐久性、稳定性、安装方便
6	警示标示(儿童)	适用年龄段标示、安装、折叠、调整
7	价格定位	价格、性价比、价格变动

Continued

8	交易与支持服务	支持顾客交易与支付服务
9	售后服务	售后维修、退换货等服务
10	物流服务	提供物流运输、安装的服务

见表 2 中, 质量方面涉及到家具做工中的稳定性、工艺质量, 产品是否有毛刺、有缝隙等瑕疵, 以及产品是否容易变形等; 结构涉及到产品中的平台, 板子、阁楼、框架、骨架、梯子、跳板、垫子、吊床、门锁、天窗等组成产品的结构; 用料方面涉及产品的材质、产品是否有异味; 功能实用涉及产品内部空间、打扫卫生时产品是否容易清理、产品底部是否带有轮子可以移动、内部是否有灯光等。

Table 2. Product characteristics

表 2. 产品特征

产品特征	质量	结构	用料	功能实用	产品服务	外观设计	安装
特征词个数	58	34	21	48	22	28	18

服务方面涉及产品发货的速度、快递小哥、产品售前和售后的态度的态度等; 外观款式方面涉及产品整体的布局, 产品是否时尚、大方、高端等。颜色是否好看, 表面是否有装饰等; 安装方面涉及用户在自行安装产品时需要用到的说明书、教程、视频、图纸和需要使用到的螺丝、螺丝刀、部件、工具等。

3.4. 用户极性词典的构建

用户在发表评论时会对产品有不同的情感倾向, 即用户对于产品的情感态度。在本研究中将用户的情感倾向分为两类: 正向、负向, 在相同的情感倾向下用户表达情感的程度也不相同。因此, 只有确定了正向、负向的情感词, 即基准词, 才能根据关联程度得到其它词的情感程度。将获得所有情感候选词中无效或意思相近的形容词进行剔除、归类, 从中选取基准词候选词时需满足两个条件: 1) 形容词成对出现且词语间的意思相反; 2) 词语有情感倾向, 能够代表用户的情感态度。由此得到有明显情感倾向且可以判断出正负向的情感词语 11 对, 如表 3 所示。需要根据产品特征判断情感倾向的词语 6 对, 表 4 所示。同时定义其正向极性词的情感强度为+1, 负向极性词的情感强度为-1。

Table 3. Positive and negative reference phrases

表 3. 正负基准词组

正向基准词	词频	负向基准词	词频
好	5284	差	335
方便	791	麻烦	103
牢固	575	松动	20
简单	457	难	121
严实	147	松动	20
精细	124	粗糙	33
耐心	221	不耐心	8

Continued

愉快	87	生气	29
光滑	54	粗糙	33
充足	26	少	49
合理	3	不合理	10

Table 4. Reference phrases

表 4. 基准词组

基准词	词频	基准词	词频
大	2310	小	488
高	324	低	19
快	255	慢	34
厚	109	薄	49
粗	30	细	28
软	51	硬	28
自由	88	挤	6

接下来根据 7 类宠物家具的产品特征匹配相应的基准词, 见表 5 所示:

Table 5. Benchmark terms for various characteristics of pet furniture products

表 5. 宠物家具产品各特征基准词

质量		结构	
正向基准词	负向基准词	正向基准词	负向基准词
好	差	好	差
精细	粗糙	厚	薄
牢固	松动	牢固	松动
安装		功能实用	
正向基准词	负向基准词	正向基准词	负向基准词
快	慢	好	差
简单	难	自由	挤
方便	麻烦	方便	麻烦
产品服务		外观款式	
正向基准词	负向基准词	正向基准词	负向基准词
快	慢	好	差
耐心	不耐烦	合理	不合理
严实	松动	精细	粗糙

Continued

用料	
正向基准词	负向基准词
好	差
厚实	薄
很足	少

以图5中7组3对感性评价基准词为中心, 结合训练得到的词向量空间, 获得每一个极性词与产品不同属性下对应的基准词的关联程度, 再结合python语言利用式1, 得出非基准词的情感强度, 即若极性词与正向基准词关联程度大则乘以+1, 若与负向基准词关联程度最大则乘以-1, 由此得到获得7份极性词典, 如图5所示。

结构	用料	安装	外观款式	实用功能	产品服务	质量
好: 1	好: 1	好: 0.851155084439889	一般: 0.5329672196390488	好: 1	好: 0.6191724410879608	好: 1
一般: 0.5329672196390488	一般: 0.5329672196390488	一般: 0.5329672196390488	大: 0.593880952164485	一般: -0.5340208641484547	一般: 0.5329672196390488	一般: 0.5329672196390488
大: 0.593880952164485	大: -1	大: 0.593880952164485	小: 0.5676514401770494	大: 0.4010999238828459	大: 0.593880952164485	大: 0.593880952164485
小: 0.5676514401770494	小: 1	小: 0.5676514401770494	很好: 0.6639934936291016	小: 0.3605070348346416	小: 0.5676514401770494	小: 0.5676514401770494
很好: 0.6639934936291016	很好: 0.6639934936291016	很好: 0.6639934936291016	很大: -0.4941369186696625	很好: 0.6639934936291016	很好: 0.6639934936291016	很好: 0.6639934936291016
很大: -0.4941369186696625	很大: -0.5580346001163462	很大: -0.4941369186696625	最大: 0.38029554285130596	很大: 0.6371230289672437	很大: -0.4941369186696625	很大: -0.4941369186696625
最大: 0.38029554285130596	最大: -0.4614070271771692	最大: 0.38029554285130596	不错: 0.6010999238828459	最大: 0.5685070348346416	最大: 0.38029554285130596	最大: 0.38029554285130596
不错: 0.5397405100419417	不错: 0.6137990565564911	不错: 0.6010999238828459	明显: -0.503975631472317	不错: 0.6010999238828459	不错: 0.6010999238828459	不错: 0.6010999238828459
明显: -0.503975631472317	明显: -0.5432311751748803	明显: -0.503975631472317	容易: 0.5397405100419417	明显: 0.5203991349188365	明显: -0.503975631472317	明显: -0.5065360138499084
容易: 0.5397405100419417	容易: -0.5418256843388176	容易: 0.6689090528525845	快: 0.5574835737389813	容易: -0.5447983216133535	容易: 0.5397405100419417	容易: 0.5397405100419417
快: 0.5574835737389813	快: 0.46237009929171824	快: 1	高: 0.5592172600686266	快: 0.5574835737389813	快: 1	快: 0.5574835737389813
高: 0.5592172600686266	高: 0.5592172600686266	高: -0.5592172600686266	少: -0.5612766281332999	高: 0.6292355818173065	高: 0.5592172600686266	高: 0.5592172600686266
少: -0.5612766281332999	少: -0.5612766281332999	少: 0.5612766281332999	简单: 1	少: 0.6100186459281183	少: -0.5612766281332999	少: -0.5612766281332999
简单: 0.5506439791770206	简单: 0.5506439791770206	简单: 1	严重: -0.5154908400335555	简单: 0.5506439791770206	简单: 0.5506439791770206	简单: 0.5506439791770206
严重: -0.4780234447786318	严重: -0.4332944918115401	严重: -0.4750863863160474	新: 0.4726589182967471	严重: 0.4922046698484104	严重: -0.4332944918115401	严重: -0.4780234447786318
新: 0.4726589182967471	新: -0.5043526187961265	新: 0.4726589182967471	清楚: 0.5586474869914819	新: 0.4726589182967471	新: 0.4726589182967471	新: 0.4726589182967471
清楚: 0.5505886823892379	清楚: 0.5505886823892379	清楚: 0.5505886823892379	迅速: -0.4084930370149711	清楚: 0.5505886823892379	清楚: 0.5505886823892379	清楚: 0.5505886823892379
迅速: -0.4084930370149711	迅速: -0.4264004902345738	迅速: 0.43575061760321643	幸福: 0.4947837151598504	迅速: 0.4174838076592269	迅速: 0.529589229111829	迅速: -0.4084930370149711
幸福: 0.4947837151598504	幸福: 0.4947837151598504	幸福: 0.4947837151598504	最高: 0.4542344173361328	幸福: 0.4947837151598504	幸福: 0.4947837151598504	幸福: 0.4947837151598504
最高: -0.3671810035718531	最高: -0.3982466091360025	最高: -0.3671810035718531	厉害: 0.5306145075467819	最高: 0.40806514156211826	最高: -0.3671810035718531	最高: -0.3671810035718531
厉害: 0.5306145075467819	厉害: 0.5306145075467819	厉害: 0.5306145075467819	安全: 0.5115434300675801	厉害: 0.5361748106508273	厉害: 0.591443664472075	厉害: 0.5306145075467819
安全: 0.4768859039700598	安全: 0.4539685715384565	安全: 0.586209249344931	麻烦: 0.45795795681899776	安全: 0.586209249344931	安全: 0.4539685715384565	安全: 0.4768859039700598
麻烦: -0.45795795681899776	麻烦: -0.5044021416905821	麻烦: -1	方便: -0.505755292708138	麻烦: -1	麻烦: -0.45795795681899776	麻烦: 0.45795795681899776
方便: 0.4933470088919296	方便: 0.4933470088919296	方便: 1	强: 0.6259065158383582	方便: 1	方便: 0.4115196096873833	方便: 0.4933470088919296

Figure 5. User polarity dictionary corresponding to various characteristics of pet furniture products

图5. 宠物家具产品各特征对应的用户极性词典

3.5. 用户评价分析

通过产品特征和用户极性词典, 获得相关特征的极性词对应的情感强度, 根据式(2)计算用户对该产品特征的感性评价价值, 同时根据式(3)计算用户对产品特征的关注程度, 根据式(4)计算产品的需改进程度, 得到的产品评价体系见表6所示。

Table 6. Evaluation indicators of pet furniture products

表6. 宠物家具产品评价指标

一级评价指标(产品属性)	满意度	关注度	需改进程度
质量	-0.057005519	0.475976495	0.503109782
功能实用	0.055061414	0.291393018	0.275348506
产品服务	0.119389163	0.281368821	0.247776433
用料	0.110742933	0.126512271	0.112501931

Continued

结构	0.001478849	0.190217767	0.090084349
外观款式	0.123969277	0.095402696	0.083575693
安装	-0.00250547	0.023850674	0.023910431

通过对宠物家具在线评论文本信息的挖掘, 提炼影响用户感知的各种产品特征, 发掘关键项目可以作为企业决策的依据。通过以上的数据挖掘分析, 得出了宠物家具产品各特征的需改进程度(如图 6), 在产品质量、功能实用、产品服务、用料、结构、外观款式、安装等七个方面各有侧重。

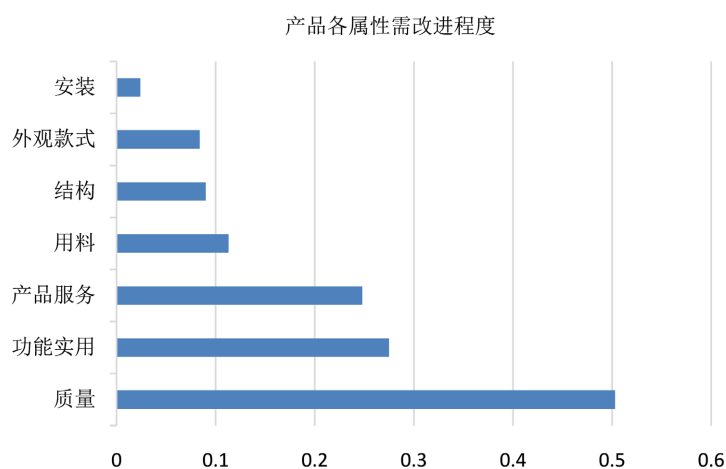


Figure 6. The degree of improvement required for various features of pet furniture products
图 6. 宠物家具产品各特征的需改进程度

对比各产品特征的需改进程度, 可以发现质量的满意度为负值且关注度较大, 说明其是亟需改进的; 其次是功能实用和产品服务, 用户对其满意度相差较大, 但是需改进程度却相差无几, 这是因为用户更在关注产品本身的属性, 而服务只是产品的附加值; 外观款式、结构、用料这三个特征也是产品需要改进的, 但是相对优先级不高, 若想在同类产品中脱颖而出可以在这三个特征上加大开发成本; 最后是产品的安装, 用户满意度低关注度也低, 若开发成本有限可以不予考虑。

3.6. 产品设计改进

在找到宠物家具产品各个亟需改进的用户不满意的特征点后, 基于专家意见提出具体的改进策略。宠物家具产品的具体改进策略如下:

提高质量: 优化设计和工艺, 在设计和制造阶段要考虑质量和耐用性, 使用优质材料和合适的工具和技术。并致力于提高生产过程中的精度、细度, 采取有效的步骤来确保产品的一致性和质量。在实施质量管理计划时, 确保在制造家具的过程中完成测试和检查, 并及时对质量问题进行更正和纠正。

功能实用: 在设计宠物家具时, 要考虑到宠物的不同需求。可以设计出多功能的家具, 例如可以作为床、窝和储物空间的床铺、储物箱等, 还应该考虑宠物的大小和体型, 在设计家具时尺寸应根据此来确定。确保宠物可以完全舒适地休息、玩耍和行动。

产品服务: 售前向客户提供足够的帮助和建议, 详细的产品描述和图片等; 售中对订单的处理迅速响应, 制造和交货过程精益求精, 让客户在规定的期限内收到自己家具; 售后提供产品的保修服务、售后服务热线、快速响应客户的投诉、提供客户满意的售后服务等。

用料: 宠物喜欢咬、抓和嚼东西, 要选择不含有毒有害物质的材料, 同时也需要耐久性好, 容易清理, 如环保、易清洁、抗压抗摔的 EVA 材料。

结构: 宠物家具的结构设计要简单明了, 避免复杂的结构和功能, 以降低生产成本, 提高产品质量; 同时还要便于清洁和维护, 例如, 床垫和床铺底座可以轻松拆卸和重新组装, 提高实用功效。

外观款式: 尽量贴近宠物的兴趣和喜好, 例如某些宠物喜欢爬高, 所以可以设计一些高低不一的垫子和爬架, 让宠物更容易进出; 可分为不同的风格, 例如现代风格、简约风格、奢侈风格等等, 可根据消费群体的需求和喜好做出不同的选择; 也可以打造品牌特点, 例如制作一个固定的标志, 作为宠物家具品牌的代表。

安装: 在包装中提供图文并茂的安装手册, 可以帮助消费者轻松理解, 降低错误安装率; 提供必要的安装工具, 如螺丝刀、棘轮等, 以确保正确的安装; 设计简单的可拆卸式结构, 方便消费者随时拆卸和移动家具; 在产品售后服务中提供技术支持, 如电话或在线视频等服务, 以便消费者随时解决问题。

4. 结语

由于互联网时代传统家具制造业的生存环境发生了改变, 积累并分析数据是宠物家具企业必须要做的, 特别是消费端的数据, 包括用户的偏好和对产品的反馈, 通过这些数据, 企业可以更好地了解和服务用户, 指导家具产品开发与制造。本文主要构建了基于在线文本挖掘的宠物家具产品评价模型, 得到了用户选择宠物家具产品的 7 个准则——质量、结构、安装、实用功能、产品服务、外观款式用料, 以及在产品新产品开发过程中各个特征的需要改进程度。利用大数据获取最新的用户需求动态信息, 以此作为产品开发设计的依据, 找出产品的优势、劣势及用户需求的重要程度的差异, 为帮助家具企业在有限资源下进行产品改进和创新设计做好准备。

相较于传统的感性工学方法, 运用自然语言处理技术研究宠物家具设计理论, 以数据驱动用户需求挖掘, 其数据更加真实可靠, 评价结果更加科学合理。将用户的满意度、产品的需改进程度进行量化以便指导宠物家具企业改进产品设计, 解决了设计问题的模糊性。

本文研究的主要是对产品的一级特征和满意度进行挖掘, 如何进一步通过一级特征得到更加细致的用户需求并提出具体的设计改进方案, 是今后需要进一步研究的方向。

注 释

文中所有图片均为作者自绘。

参考文献

- [1] 高新勤, 金雨昊, 王雪萍, 等. 基于在线评论挖掘的产品感性评价方法研究[J]. 现代制造工程, 2021(12): 13-20. <https://doi.org/10.16731/j.cnki.1671-3133.2021.12.003>
- [2] 成方敏, 余隋怀, 初建杰, 等. 在线产品评论中用户体验信息采集与分析方法[J]. 机械科学与技术, 2022, 41(2): 227-234. <https://doi.org/10.13433/j.cnki.1003-8728.20200408>
- [3] 张振华. “互联网+”时代木家具用户决策机制及新产品设计策划的对应性研究[D]: [博士学位论文]. 南京: 南京林业大学, 2020. <https://doi.org/10.27242/d.cnki.gnjlu.2020.000026>
- [4] 杨程, 谭昆, 俞春阳. 基于评论大数据的手机产品改进[J]. 计算机集成制造系统, 2020, 26(11): 3074-3083. <https://doi.org/10.13196/j.cims.2020.11.018>
- [5] 卢春燕. 基于在线评论文本挖掘的体验型产品用户需求分析[D]: [硕士学位论文]. 太原: 山西财经大学, 2021. <https://doi.org/10.27283/d.cnki.gsxcc.2021.000041>
- [6] Wang, W.M., Wang, J.W., Li, Z., et al. (2019) Multiple Affective At-Tribute Classification of Online Customer Product Reviews: A Heuristic Deep Learning Method for Supporting Kansei Engineering. *Engineering Applications of Artificial Intelligence*, **85**, 33-45.

- [7] 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.
- [8] Jiao, Y.R. and Qu, Q.X. (2019) A Proposal for Kansei Knowledge Extraction Method Based on Natural Language Processing Technology and Online Product Reviews. *Computers in Industry*, **108**, 1-11.
- [9] 刘名成. 基于评论数据的产品设计辅助决策知识服务技术方法研究[D]: [硕士学位论文]. 宁波: 宁波大学, 2020. <https://doi.org/10.27256/d.cnki.gnbou.2020.001315>
- [10] 任海兵, 周晶, 冯万群. 基于马田系统的智能手机用户体验评价研究[J]. 包装工程, 2017, 38(20): 204-210.
- [11] 徐斌. 数据驱动的用户需求挖掘与方案改进知识服务研究[D]: [硕士学位论文]. 宁波: 宁波大学, 2020. <https://doi.org/10.27256/d.cnki.gnbou.2020.001410>
- [12] 张蝶. 在线评论挖掘用户需求导向的拇外翻矫形器改良设计[D]: [硕士学位论文]. 上海: 东华大学, 2022. <https://doi.org/10.27012/d.cnki.gdhuu.2022.000871>