

评论数据驱动的产品情感满意度预测模型构建

孙利, 张硕, 覃忠志, 吴俭涛, 李江南

(燕山大学, 河北 秦皇岛 066004)

摘要: **目的** 针对康养辅具产品多维设计要素与用户情感满意度的复杂映射问题, 提出了一种基于评论数据的产品情感满意度预测模型。**方法** 首先爬取样本评论数据, 并利用 LDA 模型构建产品多维设计要素空间, 同时对样本进行要素编码; 其次对各样本对应的评论数据进行情感满意度分析, 以获得情感满意度分值; 最后通过 BP 神经网络构建情感满意度预测模型, 结合 K-fold 交叉验证法测试模型可靠性。**结果** 以康养辅具电动轮椅为例, 通过 30 个样本数据集交叉验证后得到预测值与期望值的均方误差值为 0.044 1, 表明该预测模型精度良好, 具有较高的可靠性。最后利用该预测模型对 104 976 个随机组合方案进行情感满意度计算, 得出电动轮椅最佳组合方案。**结论** 基于用户评论数据构建的产品情感满意度预测模型, 能够有效建立产品多维设计要素与情感满意度间的匹配关系, 辅助设计人员快速识别出用户情感满意度较高的多维设计要素组合, 提高了设计方案决策的科学性。

关键词: 产品设计; 康养辅具产品; 多维设计要素; 情感满意度; 用户评论数据

中图分类号: TB472 **文献标识码:** A **文章编号:** 1001-3563(2023)06-0042-10

DOI: 10.19554/j.cnki.1001-3563.2023.06.005

Construction of Product Emotional Satisfaction Prediction Model Driven by Comment Data

SUN Li, ZHANG Shuo, QIN Zhong-zhi, WU Jian-tao, LI Jiang-nan

(Yanshan University, Hebei Qinhuangdao 066004, China)

ABSTRACT: The work aims to propose a prediction model of product emotional satisfaction based on comment data aiming at the complex mapping problem between multi-dimensional design elements and user emotional satisfaction of health care assistive devices. Firstly, the comment data of samples were crawled, the multi-dimensional element space of products was constructed by LDA model and the elements of the samples were coded. Secondly, the emotional satisfaction of the comment data corresponding to each sample was analyzed to obtain the emotional satisfaction score. Finally, the emotional satisfaction prediction model was constructed by BP neural network, and the reliability of the model was tested by K-fold cross-validation method. With the electric wheelchair as an example, the mean square error between the predicted value and the expected value was 0.044 1 after cross-validation of 30 sample data sets, which showed that the prediction model had good accuracy and high reliability. Next, the emotional satisfaction of 104 976 random combination schemes was calculated by the prediction model, and the best combination scheme of electric wheelchair was obtained. The prediction model of product emotional satisfaction based on user comment data can effectively establish the matching relationship between multi-dimensional design elements and emotional satisfaction, assist designers to quickly identify the combination of multi-dimensional design elements with high user emotional satisfaction, and improve the scientific decision-making of design schemes.

KEY WORDS: product design; health care assistive device; multidimensional design elements; emotional satisfaction; user comment data

收稿日期: 2022-10-22

基金项目: 国家社科基金艺术学项目 (22BG125)

作者简介: 孙利 (1973—), 男, 教授, 主要研究方向为机械装备及康养辅具创新设计、智能数字化设计理论与方法研究。

通信作者: 张硕 (1992—), 女, 博士生, 主要研究方向为康复辅具产品设计, 产品智能设计方法。

产品情感满意度是用户对产品整体评价的真实反馈,是衡量产品投放市场后销售优劣的关键指标^[1]。尤其在康复辅具设计领域,仍存在设计师主观性强、用户评价收集数据少,且反馈到产品设计流程时效性差等问题,继而影响用户对产品的情感满意度。因此,在设计阶段构建产品设计要素与用户情感满意度的映射模型,并进行用户情感满意度预测,能够帮助企业精准筛选投放市场的产品设计要素,降低成本与风险。近年来,如何准确构建用户情感满意度模型也成为众多学者的研究热点。BP 神经网络作为一种智能信息处理系统,具有较强的非线性映射能力和柔性的网络结构,在模型预测领域应用广泛^[2]。

郭惠萍等^[3]将感性工学与 BP 神经网络相结合,建立了轮式拖拉机多维设计要素与用户感性评价关系的满意度模型;蔡婉欣等^[4]结合 AHP 与 BP 神经网络建立了产品意象多维设计要素组合推导方法,并以此预测了四旋翼无人机设计方案;窦金花等^[5]基于深度学习构建了产品外观意象与用户情感满意度需求的映射模型,实现产品设计方案与用户意象需求的有效匹配;朱思羽等^[6]提出多因素智能化融合设计,从多个维度建立产品综合评价模型,解决设计因素单一的问题。

上述文献大多仅从产品造型出发研究产品情感满意度,忽略其他要素对产品满意度的影响,缺乏了测量的多维性;另一部分虽考虑了产品多维设计要素的测量,但设计要素大多采用人工选择或问卷调查的方式,无法保证产品要素与情感满意度关联的必要性、准确性与实时性。因此,文中采用 LDA 模型与 BP 神经网络相结合的方式构建产品情感满意度模型,通过爬取产品在线评论大数据,以海量的用户真实反馈数据精准确定影响产品情感满意度的关键要素,并通过 BP 神经网络模型构建产品多维设计要素与情感满意度的映射模型,为后续待开发产品的情感满意度评价与市场预测提供参考。

1 理论基础

1.1 LDA 模型

LDA 模型是 Blei 等^[7]学者首次提出的一种无监督形式的机器学习方法,可应用于识别大量文本隐含的主题研究。LDA 主题模型假设文档集由一些隐含主题通过一定概率进行的随机组合^[8],其结构为三层贝叶斯模型,主要由文本、主题和词三部分构成,且各隐含主题彼此独立^[9],其模型见图 1。

其中 α 、 β 为狄利克雷函数的先验参数, α 代表各主题间的强弱关系, β 代表主题的概率分布。 θ 为该主题在整个文档中的多项分布参数, ϕ 为该词在整个主题中的多项分布参数。 z 为隐含主题, w 为文本中的特征词^[10]。

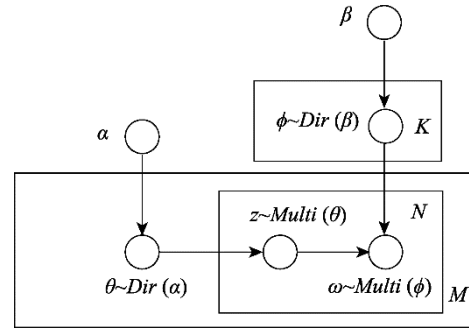


图 1 LDA 贝叶斯结构模型

Fig.1 Bayesian network structure of LDA model

第 i 个特征词 w_i 的生成概率 $P(w_i)$ 的公式如下:

$$P(w_i) = \sum_{k=1}^T P(w_i | z_i = k) P(z_i = k) \quad (1)$$

式中: $P(w_i | z = k)$ 为词 w_i 在第 k 个主题中的概率, $P(z_i = k)$ 表示第 k 个主题在评论文本中的概率。通过 LDA 模型可将研究数据挖掘出隐含主题,尤其针对评论数据文本,能够实现用户核心观点挖掘,达到快速准确的文本聚类目的^[11]。

1.2 情感满意度分析

情感满意度分析又称倾向性分析,是对带有情感满意度色彩的主观性文本进行分析、处理、归纳和推理的过程^[12]。针对电商平台用户对产品的评论数据而言,情感满意度分析能够有效反映用户对产品的真实情感倾向,帮助企业快速、客观、准确地了解消费者对产品的偏好程度^[13]。此外,将大量的用户真实的评论数据进行分析,可及时判断已上市产品改进要点的优先级排列,并在此基础上构建产品情感满意度评价模型,从而实现未投入市场产品的情感满意度分值预测,帮助企业或设计师快速调整产品设计战略。

1.3 BP 神经网络

BP 神经网络是由 Rumelhart 等提出的一种前向型神经网络模型,具有高度非线性映射能力和较好的容错能力,既适用于线性模式的数据集,也适用于非线性的数据集^[14]。通过 BP 神经网络可以真实反映输入变量和输出变量之间的映射关系,实现对于未知数据的有效预测,可以帮助企业市场人员和设计人员快速决策出满足用户偏好的产品方案^[15]。

2 评论数据驱动的产品情感满意度模型构建

本文拟结合 LDA 模型、情感满意度分析与 BP 神经网络构建产品情感满意度模型,通过 LDA 模型对产品评论数据进行聚类分析,确定产品多维设计要素。同时结合评论数据的情感分析结果,构建以产品多维设计要素与情感满意度为变量的 BP 神经网络模型,用以研究两者之间的隐性关联,并提出以评论数据为基础、结合 LDA 模型与 BP 神经网络的产品情

感满意度预测流程，见图 2。

1) 评论数据获取与分类。采用电商评论大数据对产品进行分析，首先要确定研究样本，并采用爬虫工具获取各样本图及其对应的评论数据。鉴于网络商务平台产品评论数据质量参差不齐，需对收集到的评论进行数据清洗等操作，从而降低噪声影响^[16]。数据清洗完毕后，采用 Jieba 分词处理并得到评论数据词汇集合，为进一步采用 LDA 模型进行文本聚类奠定基础。

2) 产品多维设计要素空间建立。分词后的词汇集合采用 LDA 模型进行聚类分析，通过在 Python 中运行 pyLDAvis 可视化工具确定聚类个数，聚类结果即为产品宏观设计要素。在此基础上，结合研究对象进行多维设计要素拆解，确定研究对象具体多维设计要素，从而完成产品多维要素空间的建立。随后将多个研究样本进行编码处理，作为后续 BP 神经网络

测模型的输入层数据。

3) 评论数据情感满意度量化。通过 ROST CM6 对各样本对应的评论数据进行情感满意度分析，可得到各评论数据的情感满意度分值。为保障数据准确性，采用莱依达准则去除极端数据，将处理后的数据进行均值处理，即为各样本的最终情感满意度分值，并作为后续 BP 神经网络预测模型的输出层数据。

4) 情感满意度预测模型建立。将各样本对应多维设计要素空间，进行元素编码并作为输入数据集，各样本情感满意度分值作为输出数据集，运用 BP 神经网络构建并训练多维设计要素与情感满意度映射模型，同时采用 K-fold 交叉训练方法对预测模型进行检测，验证该预测模型的精确度。后通过构建的情感满意度模型对随机组合的方案进行预测，输出各方案的情感满意度分值并进行排序，从而获得最佳组合方案。

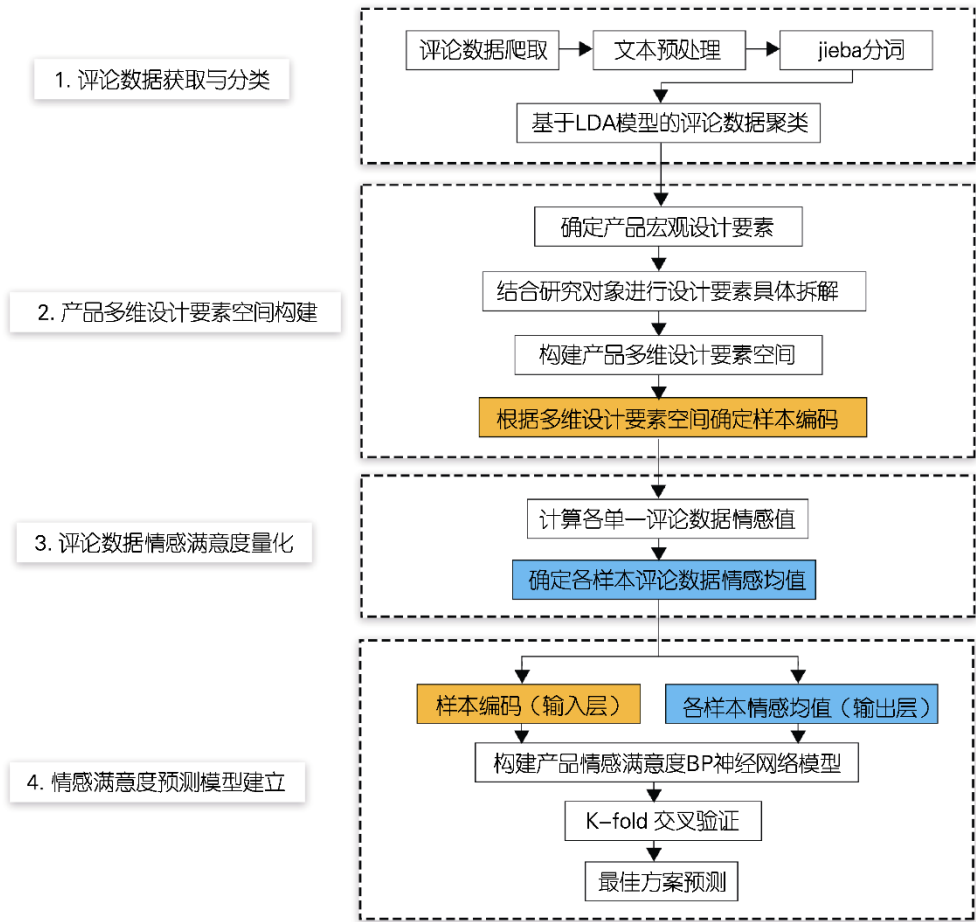


图 2 产品情感满意度预测模型研究流程
Fig.2 Research process of product emotional satisfaction prediction model

3 设计分析与模型构建

人口老龄化提升了用户对康复辅具的设计要求，现有康复辅具设计存在设计师主观性强、忽视用户真实反馈或用户评价不及时反馈到设计流程等诸多问题，降低了用户对产品的情感满意度。本文以康复辅

具电动轮椅为例，通过 LDA 模型对样本进行评论数据挖掘和情感满意度计算，从而构建产品情感满意度 BP 神经网络预测模型，为康复领域产品的设计定位提供理论依据和指导。

3.1 数据获取与预处理

实验环境基于 Windows 10 的 64 位操作系统，

通过利用 Jupyter Notebook 运行 Python 语言搭建实验平台。通过对比各大电商平台的电动康复轮椅销售情况,发现京东平台的用户数据的海量性和真实性更佳,更能真实反映用户对产品的全面评价,故选择京东平台作为数据获取源。以京东销量前 30 名的电动康复轮椅产品作为研究样本,见图 3。

采用“八爪鱼采集器”爬虫工具获取 30 个样本的评论数据,并将其作为初始语料库,最终收集到 8 213 条初始数据。鉴于网络商务平台的在线评论质量参差不齐,通过网络爬虫的形式爬取到的数据存在异常,易影响数据分析的准确性。因此,需要对收集到的数据进行数据清洗,通过删除默认评论、垃圾评

论和重复评论来减少噪音干扰值。通过数据预处理,最终将初始语料库数据清洗至 7 432 条可用数据。

清洗后的语料库需进行分词处理,鉴于数据来源为电商平台的在线评论,其具有用词的多样性和口语化等特点,因此采用更适合于点评类数据的 Jieba 分词工具。为进一步提升文本质量并降低文本特征的维度,需在分词前去除停用词。此处采用常用的百度停用词表对电动轮椅文本数据进行停用词过滤,通过在 Jupyter Notebook 中运行 Python 去除停用词代码,最终得到分词后的电动康复轮椅文本集合。分词结果示例如图 4 所示,经过分词后的数据集即可进行主题聚类等分析。



图 3 30 个电动轮椅样本
Fig.3 30 samples of electric wheelchairs



图 4 电动轮椅样本评论数据分词
Fig.4 Tokenization of comment data for electric wheelchair samples

3.2 LDA 主题聚类分析

为了使构建的产品情感满意度预测模型更为准确,其 BP 神经网络的输入层需摒弃以往的人工界定,转而采用通过 LDA 模型挖掘评论数据的方法确定产

品多维设计要素。在 Python 中运行 LDA 模型,对电动轮椅分词后的评论数据进行聚类分析,同时采用 pyLDAvis 进行主题可视化操作,根据主题分布形态最终确定主题数为 4 时分类效果最佳,电动轮椅的主题聚类结果见图 5。

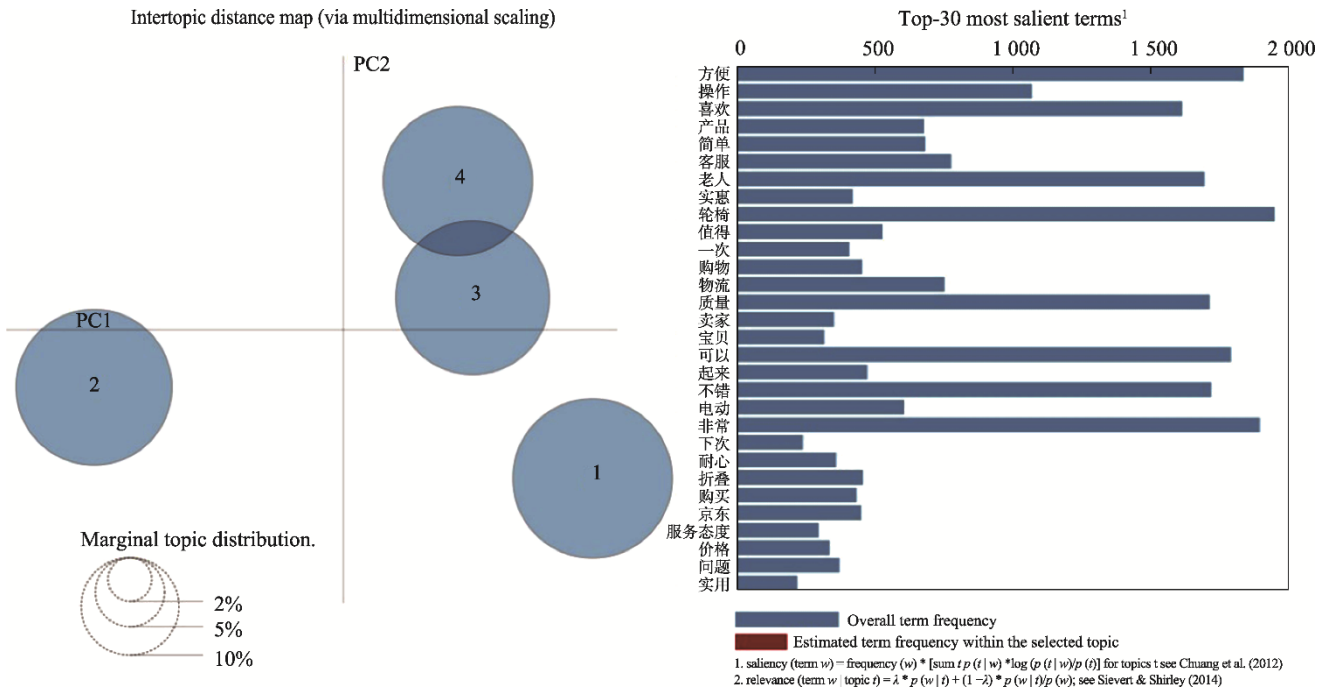


图 5 电动轮椅样本评论数据 LDA 模型可视化
Fig.5 LDA model visualization of comment data for electric wheelchair samples

图 5 左侧的每个圆圈代表一个主题,圆圈的大小代表主题的频率高低,圆圈越大则代表该主题重要度越高。主题的聚类优劣可以根据代表各主题的圆圈分布来分析,主题圆圈之间的间隙越小,代表两主题内容越接近。选定特定主题时,右侧会展示该主题对应的特征词,默认为紧密度最高的前 30 个词。

通过对每个主题及其对应的特征词进行详细分析,最终确定 4 个主题名称及其主题特征词,每个主题选取 10 个特征词,见表 1。由表可知,用户对电动轮椅的形态、参数、价格和服务几个宏观设计要素

较为关注,而以往利用 BP 神经网络构建的预测模型,大多数只关注形态设计要素,由此可见,产品多维设计要素更贴合用户的真实需求,为后续精准构建用户情感满意度 BP 神经网络模型夯实基础。

3.3 产品多维设计要素空间构建

为了构建产品情感满意度预测 BP 神经网络模型,需确定具体产品多维设计要素并作为 BP 神经网络的输入层。上文已确定 4 个宏观设计要素即形态、参数、价格和服务。分析可知,服务属于可以主观动态调控的要素,与产品本身的客观要素无关。因此,此处只考虑形态、参数、价格三个方向。根据表 1 对各主题及其特征词进行分析可知,形态方面用户对产品的造型、灵活性等有一定需求,后续采用形态拆解法对电动轮椅形态进行分析,从而确定产品形态要素。参数方面用户主要关注操作的轻便性、行驶里程、安全性等特征,价格方面则可对具体价格区间作后续分析。通过形态拆解法对电动康复轮椅进行分析,由于把手处遥控装置属于通用型装置,各样本此处差异性很小,且根据评论数据可知用户对遥控部分鲜少提及。故将电动轮椅形态构成分为靠背、脚踏、扶手、侧支撑及底座支架 5 个部分,见图 6。

利用形态拆解法对电动康复轮椅分析后,对比

表 1 电动轮椅样本聚类主题及其特征词
Tab.1 Clustering topics and characteristic words of electric wheelchair samples

序号	主题名称	主题特征词
1	形态	方便 喜欢 结实 灵活 小巧 稳当 款式 好看 把手 设计
2	参数	方便 操作 简单 电动 折叠 轻便 电池 安全 做工 上坡
3	价格	值得 满意 实惠 一般 划算 不贵 性价比 便宜 降价 赠品
4	服务	物流 满意 客服 快递 送货 京东 及时 收到 发货 服务



图 6 电动轮椅形态拆解
Fig.6 Dismantling of electric wheelchair form

30 个研究样本分析 5 部分形态类型并汇总形态特征元素集合。对形态设计要素分析后，进而对参数与价

格两个要素进行分析。鉴于参数要素在 LDA 模型分类后的主题特征词中有显著体现，因此着重参考表 1 中“参数”主题对应的特征词，分别为方便、操作、简单、电动、折叠、轻便、电池、安全、做工、上坡等。通过联合对比产品说明书进行分析，最终将参数要素确定为续航、重量、平躺、护腿板、爬坡角度、电磁刹车，每项参数的具体区间也严格遵照产品说明书的规格设定。对价格要素主要参考产品数据爬取时的电商售价，由于产品价格数目不均等，故此处也以区间形式划分，最终确定 12 项产品多维设计要素，见图 7。此部分设计产品要素即作为产品情感满意度评价 BP 神经网络模型输入层的编码参考项。

形态	1. 靠背			
	2. 脚踏			
	3. 扶手			
	4. 侧支撑			
	5. 底座支架			
参数	6. 续航/km	8~15	16~25	26~35
	7. 重量/kg	20~30	31~40	41~45
	8. 平躺	可平躺	不可平躺	
	9. 护腿板	有	无	
	10. 爬坡角度/(°)	6~15	16~25	26~35
	11. 电磁刹车	有	无	
价格	12. 价格/元	1 000~3 000	3 001~5 000	5 001~7 000

图 7 电动轮椅 12 项产品多维设计要素
Fig.7 12 multi-dimensional design elements of electric wheelchair

3.4 评论数据情感满意度分值量化

产品情感满意度 BP 神经网络模型输出层，即为产品的情感满意度评价价值，需对 30 个样本的评论数据进行情感满意度分析，进而求得各样本评论数据的满意度分值。此处情感满意度分析工具采用武汉大学沈阳教授研发的 ROST CM6，该软件为国内目前唯一以辅助人文社会科学研究的大型免费社会计算平台。

通过将评论数据导入求解，可直接得出各评论文本的情感满意度数值，数值的正负代表用户对该产品的情感满意度倾向即积极或消极，数值的大小代表用户的情感倾向程度。由于通过情感满意度分析得到的情感满意度分值存在极端评价，需要对情感满意度分析后的数据采用莱依达准则即 3σ 准则做剔除处理。莱依达准则对数据进行计算处理得到标准偏差，并通过公

式 $|vb|=|xb-x|>3\sigma$ 确定一个区间, 凡超过这个区间的误差, 就不属于随机误差而是粗大误差, 含有该误差的数据应予以剔除。将剔除粗大误差数据后的各样本评论数据进行情感满意度均值处理, 以用于构建样本输出数据集, 并将其作为产品情感满意度 BP 神经网络模型的输出层数据, 各样本情感满意度均值见表 2。

表 2 30 个电动轮椅样本情感满意值
Tab.2 Emotional satisfaction value of 30 electric wheelchair samples

序号	样本	情感满意值	序号	样本	情感满意值
1	鱼跃	1.35	16	仁和 1	1.69
2	祖医堂	1.33	17	互邦	1.44
3	凤凰	1.45	18	护卫神	1.15
4	九圆	1.71	19	仁和 2	1.71
5	联想	1.12	20	金百合	1.35
6	迈德斯特	1.37	21	英洛华 2	1.26
7	英洛华	1.22	22	英洛华 3	1.47
8	Ainsnbot	1.49	23	迈德斯特 2	1.98
9	Konyee	1.73	24	凯莱宝	1.32
10	好哥	1.72	25	左点	1.71
11	福宏	1.13	26	欧航 2	1.14
12	YATLL	1.32	27	永久	1.69
13	欧航	1.42	28	斯维驰	1.08
14	佳康顺	1.43	29	三贵	1.48
15	小飞哥	1.76	30	ichigo	1.43

3.5 产品情感满意度预测模型构建

3.5.1 样本特征编码

将 30 个电动轮椅样本对照图 7 进行多维设计要

素判定, 得到各样本设计要素分布, 其中所属类别 1、2、3 分别对应图 7 中产品多维设计要素的展开项目的序号, 如图 8 所示, 30 个样本各项要素无完全重复项且分布良好, 故 30 个样本数据均作为构建 BP 神经网络的训练集。

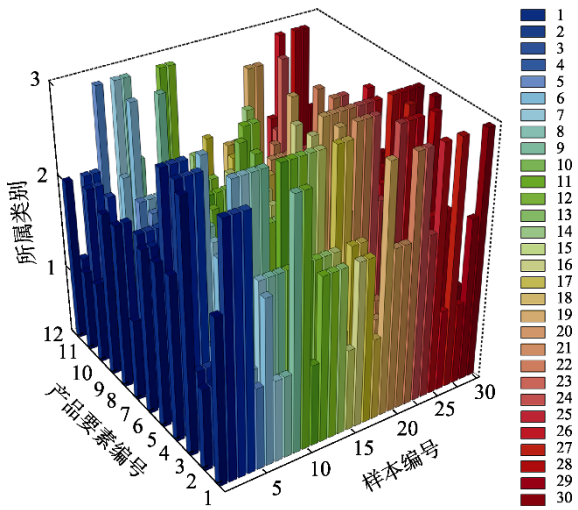


图 8 30 个电动轮椅样本多维设计要素分布
Fig.8 Multi-dimensional design element distribution of 30 electric wheelchair samples

结合图 8 可将各样本进行特征编码以构建样本输入数据集, 以样本 1 为例, 编号 1 的产品设计要素所属类别为 2, 该项设计要素共 3 个类别, 则样本 1 在第一项产品设计要素即靠背要素下的编码为 010, 该样本其他项设计要素对应编码以此类推。鉴于篇幅所限, 各样本产品要素编码做部分展示, 见表 3。

表 3 30 个电动轮椅样本产品要素编码
Tab.3 Product element code of 30 electric wheelchair samples

序号	样本	形态	参数	价格	产品要素编码
1	鱼跃	010	010	10	01010001010010010100010110010010
2	祖医堂	001	010	01	00110100001001010010011010001100
3	凤凰	001	100	01	00110100001001100010101001001010
4	九圆	001	010	10	0011010000100101001010101010010010
...	...	...	...	...	...
30	ichigo	001	010	01	0011010000100101001010101010001010

3.5.2 神经网络模型参数设置

常用的 BP 神经网络往往由输入层、隐含层、输出层三级构成, 每一层包含若干的神经元节点, 各层的神经元数量视实际情况而定。本文通过 Matlab 应用 newff 创建 BP 神经网络, 将 30 个样本设计元素编码作为神经网络的输入层, 将样本情感满意度分值作为输出层, 其结构见图 9。

由表 3 可知, 产品 12 项多维设计要素对应 32 项具体要素, 则输入层节点数量为 32; 情感满意度分值作为输出层, 则输出层节点数量为 1。然而隐含层

神经元节点数量的取值则无固定要求, 往往通过经验公式确定其取值范围, 隐含层神经元数量的经验公式如下:

$$p \leq \sqrt{n \times (q + 3)}$$
 (2)

式中: n 代表输入层的神经元数量; p 代表隐含层的神经元数量; q 代表输出层的神经元数量。根据公式 (2) 计算可得隐含层神经元数量最大值为 11, 为进一步确定隐含层神经元数量, 通过试凑法以确定最佳隐含层神经元节点数量, 发现当 p 值为 3 时, 该 BP 神经网络的输出误差相对最小, 最终确定神经网络

络预测模型各层级神经元的个数分别为 32、3、1, 具体网络结构见图 10。

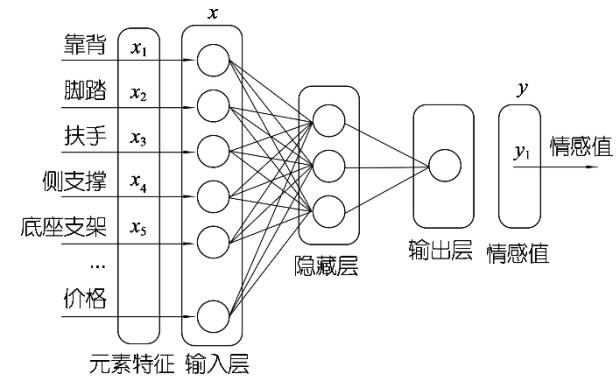


图 9 产品情感满意度 BP 神经网络结构
Fig.9 BP neural network structure of product emotional satisfaction

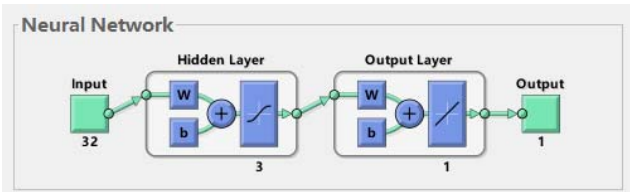


图 10 产品情感满意度 BP 神经网络结构
Fig.10 BP neural network structure of product emotional satisfaction

3.5.3 神经网络模型测试验证

本文通过 Matlab 应用 newff 创建 BP 神经网络, 采用 trainlm 算法进行训练, tan-gent sigmoid 作为隐含层传递函数, purelin 作为输出层应用函数, 学习次数设置为 1 000 次, 误差值设置为 0.01, 通过样本输入和输出数据集进行训练计算得到 BP 神经网络各层级之间的权重连接值。由于样本数据有限, 为避免无法拟合或过拟合需提升模型泛化能力, 采用 K-fold 进行交叉验证。将样本数据集随机划分为 K 组, 其中 K-1 组作为训练数据, 剩余 1 组作为测试数据。经过 K 轮循环, 使每组数据都能充当训练集和测试集, 较大提升样本数据的使用率, 同时有效避免将数据集简单划分为训练集、测试集所带来的局限性。

将样本数据集输入到预测模型之中, 将 K 值设置为 30, 即将样本数据集中随机划分为 30 组, 通过对预测模型进行 30 次交叉验证, 得到各样本情感满意度预测值, 见图 11。

由图 11 可知情感满意度预测值和期望值的数据误差较小, 随后通过计算各组样本情感满意度预测值与期望值相对误差得到平均均方误差 MSE 值为 0.044 1, 结果表明该神经网络模型符合预测精度要求, 可靠性较高。

3.6 最优设计方案输出

由图 7 可知, 产品多维设计要素共有 12 类, 包

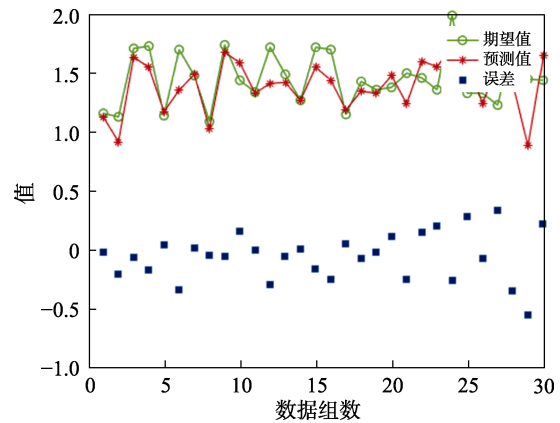


图 11 测试样本预测值和期望值误差对比
Fig.11 Deviation comparison between the predicted value and the expected value of the test sample

含 32 个产品要素, 将产品要素随机组合可产生 $3 \times 2 \times 3 \times 3 \times 3 \times 3 \times 2 \times 2 \times 3 \times 2 \times 3$, 共 104 976 个设计方案。将方案进行元素编码并作为神经网络预测模型的输入层, 利用上述已搭建的神经网络预测模型, 进行情感满意度分值预测, 见表 4。

表 4 各方案情感满意度预测值
Tab.4 Predicted emotional satisfaction value of each scheme

方案序号	方案组合元素编码	情感满意度预测值
1	01001010100010010100101000101001	2.186 1
2	01001010100001100100101000101001	2.182 4
3	010010101000010011100101000101001	2.181 5
4	01001010100001010100101000101010	2.180 4
5	010010101000010011100101000101010	2.179 9
6	01001010010001010100101000101001	2.179 3
7	01001010010001010100101000101010	2.173 7
8	010010100100010011100101000101001	2.173 2
9	01001010100001100100100100101010	2.173 0
10	010010100100010011100101000101010	2.171 9
...	...	...
821	100010011100001100100101000101001	1.999 4
...	...	...
104 976	10001001100100001010010101001100	0.858 7

鉴于方案数量庞大且相邻方案间情感满意度分值差距甚微, 故此处按情感满意度分值降序展示部分方案预测结果, 且预测值保留小数点后 4 位。由表 4 可知, 方案 1 的情感满意度预测值最高为 2.186 1, 其对应的组合方案元素编码 01001010100010010100101000101001, 故将其作为最优方案, 其具体产品要素见图 12。

根据图 12 中的设计要素, 可进一步确定产品的外观形态绘制方向、电池大小、材质及相应的结构等

	1. 靠背	2. 脚踏	3. 扶手	4. 侧支撑	5. 底座支架	
形态						
参数	6. 续航/km	7. 质量/kg	8. 平躺	9. 护腿板	10. 爬坡角度/(°)	11. 电磁刹车
	16~25	20~30	可平躺	有	26~35	无
价格/元	5 001~7 000					

图 12 电动轮椅最优方案产品多维设计要素
Fig.12 Product multi-dimensional design elements in the optimal scheme for electric wheelchair

关键性多维设计要素。以续航为例，用户满意度最佳的方案中对续航为 16~25 km，对应研究样本可知，需选择 20 A 的电池才能保证续航要求，重量约为 4 kg。车身整体重量需控制在 20~30 kg，除 4 kg 的锂电池外，车架重量最大上限约为 26 kg。若车架整体均采用碳钢材质则车重超标，故车身材质以碳钢和铝合金结合的形式进行设计，且在价格方面能够保证在 5 001~7 000 元，其他要素均以图 12 所列完成最终设计，电动轮椅效果图见图 13—14。



图 13 电动轮椅最优方案左前视图和左后视图
Fig.13 Left-front view and left-rear view of electric wheelchair under the optimal scheme



图 14 电动轮椅最优方案前视图和侧视图
Fig.14 Front view and side view of the electric wheelchair under the optimal scheme

4 结语

本文以康复辅具电动轮椅为例，爬取用户真实评论数据并通过 LDA 模型对评论数据进行聚类分析，

从而构建产品多维设计要素空间，并作为产品情感满意度 BP 神经网络预测模型的输入层。通过 ROST CM6 对样本评论数据进行情感满意度分值计算，求得各样本的情感满意度均值并作为 BP 神经网络的输出层，随后构建多维设计要素与情感满意度的映射模型，并利用 K-fold 交叉验证证实了该预测模型的可行性。

研究表明：（1）通过应用 LDA 模型对产品评论数据进行分析，可将大量文本数据进行降维并有效归纳影响产品情感满意度的多维设计要素；（2）利用 ROST CM6 对用户评论数据进行情感分析，将用户对产品的真实评价及时、准确地量化，并应用到预测模型中，能够有效提升方案的信度与效度；（3）以用户评论数据为基础构建的产品情感满意度模型，能够有效预测用户对产品的偏好程度，对设计人员快速筛选最佳方案提供切实可行的参考价值。

本文以京东平台销量较高的电动康复轮椅用户评论作为数据来源进行分析，在后续研究中将扩充多个电商平台或论坛等数据渠道，进而优化数据的多样性，并降低单一渠道数据集中分布的影响，从而提升数据挖掘的精准性。此外，未来将尝试对用户的隐式情感如人机体验等角度进行挖掘，更精准计算用户对产品的情感满意值，进而在康复产品垂类上提升情感值数据化的准确率，构建更加完善的产品情感满意度预测模型。

参考文献：

[1] 王安宁，张强，彭张林，等. 基于在线评论的区域需求偏好识别方法[J]. 中国管理科学，2019，27(7): 167-176.
WANG An-ning, ZHANG Qiang, PENG Zhang-lin, et al. Identifying Regional Demand Preferences from Online Reviews[J]. Chinese Journal of Management Science, 2019, 27(7): 167-176.

[2] 邵宏宇，孟琦，赵楠，等. 基于 BP 神经网络的产品性能满意度预测分析[J]. 天津大学学报(自然科学与工程技术版)，2019，52(9): 932-940.
SHAO Hong-yu, MENG Qi, ZHAO Nan, et al.

- Prediction and Analysis of Product Performance Satisfaction Based on Backpropagation Neural Network[J]. Journal of Tianjin University (Science and Technology), 2019, 52(9): 932-940.
- [3] 郭惠萍. 基于前脸特征的轮式拖拉机形态设计与感性评价方法研究[D]. 杨凌: 西北农林科技大学, 2018.
GUO Hui-ping. The Form Design and Kansei Evaluation of Wheeled Tractors Based on the Front Face Characteristic[D]. Yangling: Northwest A & F University, 2018.
- [4] 蔡婉欣, 林丽, 郭主恩, 等. 基于 KAB 模型的设计要素组合推导[J]. 包装工程. 2022(18): 48-56.
CAI Wan-xin, LIN Li, GUO Zhu-en, et al. Deduction of Design Elements Combination Based on KAB Model[J]. Packaging Engineering, 2022(18): 48-56.
- [5] 窦金花, 覃京燕. 基于深度学习的产品外观意象情感计算服务平台研究[J]. 包装工程, 2020, 41(6): 20-25, 31.
DOU Jin-hua, QIN Jing-yan. Affective Computing Service Platform of Product Appearance Image Based on Deep Learning[J]. Packaging Engineering, 2020, 41(6): 20-25.
- [6] 朱思羽, 胡洁, 戚进. 基于美学、人体工学及性能的多因素融合设计[J]. 机械工程学报, 2020, 56(21): 199-207.
ZHU Si-yu, HU Jie, QI Jin. Multi-Factor Coupling Design Based on Aesthetic Measure, Ergonomics, and Performance[J]. Journal of Mechanical Engineering, 2020, 56(21): 199-207.
- [7] BLEI M D, NG A, JORDAN M. Latent Dirichlet Allocation[J]. Journal of Machine Learning Research, 2003(3): 601-608.
- [8] 吴江, 侯绍新, 靳萌萌, 等. 基于 LDA 模型特征选择的在线医疗社区文本分类及用户聚类研究[J]. 情报学报, 2017, 36(11): 1183-1191.
WU Jiang, HOU Shao-xin, JIN Meng-meng, et al. LDA Feature Selection Based Text Classification and User Clustering in Chinese Online Health Community[J]. Journal of the China Society for Scientific and Technical Information, 2017, 36(11): 1183-1191.
- [9] 王春龙, 张敬旭. 基于 LDA 的改进 K-means 算法在文本聚类中的应用[J]. 计算机应用, 2014, 34(1): 249-254.
WANG Chun-long, ZHANG Jing-xu. Improved K-Means Algorithm Based on Latent Dirichlet Allocation for Text Clustering[J]. Journal of Computer Applications, 2014, 34(1): 249-254.
- [10] 廖列法, 勒孚刚, 朱亚兰. LDA 模型在专利文本分类中的应用[J]. 现代情报, 2017, 37(3): 35-39.
LIAO Lie-fa, LE Fu-gang, ZHU Ya-lan. The Application of LDA Model in Patent Text Classification[J]. Journal of Modern Information, 2017, 37(3): 35-39.
- [11] 王少鹏, 彭岩, 王洁. 基于 LDA 的文本聚类在网络舆情分析中的应用研究[J]. 山东大学学报(理学版), 2014, 49(9): 129-134.
WANG Shao-peng, PENG Yan, WANG Jie. Research of the Text Clustering Based on LDA Using in Network Public Opinion Analysis[J]. Journal of Shandong University (Natural Science), 2014, 49(9): 129-134.
- [12] 贾闻俊, 张晖, 杨春明, 等. 面向产品属性的用户情感模型[J]. 计算机应用, 2016, 36(1): 175-180.
JIA Wen-jun, ZHANG Hui, YANG Chun-ming, et al. User Sentiment Model Oriented to Product Attribute[J]. Journal of Computer Applications, 2016, 36(1): 175-180.
- [13] 蒋旒, 李然, 刘春尧, 等. PAD 情感模型在用户情感体验评估中的应用[J]. 包装工程, 2021, 42(22): 413-420.
JIANG Ni, LI Ran, LIU Chun-yao, et al. Application of PAD Emotion Model in User Emotional Experience Evaluation[J]. Packaging Engineering, 2021, 42(22): 413-420.
- [14] 李静, 李方义, 周丽蓉, 等. 基于 BP 神经网络的产品生命周期评价敏感性分析[J]. 计算机集成制造系统, 2016, 22(3): 666-671.
LI Jing, LI Fang-yi, ZHOU Li-rong, et al. Sensitivity Analysis for Life Cycle Assessment of Product Based on back Propagation Neural Network[J]. Computer Integrated Manufacturing Systems, 2016, 22(3): 666-671.
- [15] 程永胜, 徐骁琪, 陈国强, 等. 基于神经网络的电动汽车造型意象预测模型[J]. 计算机集成制造系统, 2021, 27(4): 1135-1145.
CHENG Yong-sheng, XU Xiao-qi, CHEN Guo-qiang, et al. Image Prediction Model of Electric Vehicle Based on Neural Network[J]. Computer Integrated Manufacturing Systems, 2021, 27(4): 1135-1145.
- [16] 聂爽爽, 王维红. 基于文本挖掘的速卖通平台女性消费者连衣裙偏好分析[J]. 现代纺织技术, 2021, 29(6): 113-121.
NIE Shuang-shuang, WANG Wei-hong. Research on Female Consumers' Dress Preference on AliExpress Based on Text Mining[J]. Advanced Textile Technology, 2021, 29(6): 113-121.

责任编辑: 陈作