

# 智慧健康监测设备市场的情感驱动力：基于 BERT 文本分析与 LDA 模型的整合研究

武赛 何炳升 岳梦雅

广东工业大学管理学院，广东广州，510520；

**摘要：**随着智慧健康监测设备的普及，电商平台用户评论成为优化产品设计与服务的重要依据。本研究以京东、天猫、淘宝平台的动态血糖仪用户评论为数据源，结合 BERT 模型进行情感分类，并利用 LDA 模型提取高频主题。结果表明，用户对动态血糖仪整体满意度较高，负向反馈集中于测量精度、客服质量与使用体验。LDA 主题模型进一步揭示用户对老年适用性、性价比及技术创新的关注。研究验证了多模型融合方法的有效性，为企业优化核心功能、提升用户体验提供了可操作的改进方向。

**关键词：**BERT；情感分类；LDA 主题模型；用户评论挖掘；产品优化策略

**DOI：**10.69979/3029-2700.25.07.099

## 引言

近年来，智慧健康监测设备（如动态血糖仪、智能手环等）在慢性病管理及健康监测领域广泛应用，其用户规模与市场价值持续增长。电商平台用户评论作为消费者体验的直接反馈，蕴含丰富的产品改进线索与市场需求信息。然而，传统分析方法（如单一情感分类或人工主题归纳）难以从海量文本中系统挖掘用户需求，且易忽略文本语义的深层关联。

现有研究多聚焦于单一技术手段的应用，例如基于机器学习的情感分析或基于词频统计的主题挖掘，缺乏对多模型协同分析的探索。此外，针对老年健康护理设备的研究，尤其是结合用户行为与情感反馈的综合分析仍较为匮乏。

本文以动态血糖仪为例，提出一种融合 BERT 情感分析与 LDA 主题模型的用户评论分析方法。通过 BERT 模型捕捉评论的细粒度情感倾向，结合 LDA 模型提取用户关注的核心主题，实现从“情感-主题”双维度洞察用户需求。研究不仅为动态血糖仪的产品优化提供实证依据，也为智慧健康设备的用户反馈分析提供了可扩展的方法框架。

## 1 理论基础与实验设计

### 1.1 理论基础

随着智能穿戴设备普及，智慧健康监测设备市场的用户评论与舆情分析成为研究热点。传统情感分析方法存在语义理解不足、主题泛化性弱等问题，现有研究表明，BERT 模型在捕捉文本深层语义特征方面表现优异，而 LDA 模型则擅长主题聚类，二者的整合研究为突破这一瓶颈提供了新的技术路径。

在技术整合方面，Venugopalan 等<sup>[1]</sup>提出的增强引导 LDA 模型通过结合 BERT 的语义相似性计算，显著提升了多词情感特征的提取能力。曾子明等<sup>[2]</sup>也将 BERT 与 BiLSTM-Attention 结合，应用在对突发公共卫生事件舆情分析中。在应用层面，BERT 生成的语义向量能有效提升情感分类的准确率<sup>[3]</sup>，而 LDA 聚类结果则揭示了客户最关注的问题<sup>[4-5]</sup>。这一模式同样适用于健康监测设备评论分析，例如识别用户对设备准确性、易用性等情感倾向。

因此，BERT 与 LDA 的整合能从语义理解、主题挖掘和趋势分析三个层面深化情感分析效果，为智慧健康监测设备市场的产品优化与用户需求洞察提供理论支持。

### 1.2 实验设计

#### 1.2.1 样本选取及数据来源

本研究以智慧健康监测设备中的动态血糖仪为研究对象，数据来源于京东、天猫、淘宝三大电商平台的用户评论。通过 Python 网络爬虫技术采集 2019–2023 年的用户评论文本，初始获取数据总量为 2,580 条。经过清洗后，共获得有效评论 1,276 条。数据预处理包括分词、去除停用词及文本向量化，最终构建标准化语料库。数据处理与分析通过 Python 3.8 与 PyTorch 框架实现。

#### 1.2.2 研究模型

##### BERT 情感分析模型：

采用预训练的 BERT-base-Chinese 模型，调整输出层为三分类结构（全连接层+Softmax）；参数设置：学习率设为  $2e-5$ ，批次大小为 32，训练轮次为 10，使用交叉熵损失函数优化；

##### LDA 主题模型：

主题数确定：通过计算主题一致性分数 (Coherence Score)，遍历候选主题数范围，选择一致性分数最优的配置；参数配置：Dirichlet 先验参数  $\alpha$ 、 $\beta$  设为自动优化，采用 Gibbs 采样算法进行迭代计算；主题解释：结合高频词分布与人工归纳，定义各主题的核心语义。

## 2 数据处理与分析

### 2.1 基于 BERT 的产品评论情感分析

预处理后的评论文本输入 BERT 模型进行情感分类训练；全部评论输入 LDA 模型进行主题建模，生成主题-词分布矩阵。对原 BERT 模型进行了参数调整，改变其输出层以实现三分类效果。在预训练阶段，该模型准确率达到 90.92%，展现出良好的情感识别和分类效果。

将预处理好的评论数据全部导入到预先训练好的 BERT 模型中，输出结果为：在 1276 条用户评论数据中，正向评价占 92.08% (1175 条)，负向评价占 4.78% (61 条)，中性评价约占 3.14% (40 条)。这说明用户对动态血糖仪产品整体上持看好的态度。

由于中性评论没有任何情感倾向，也不会对最终分析结果造成影响，因此，本研究针对负向评论进行深度剖析动态血糖仪产品存在的不足。通过提取负向评论，进行分词去停词等预处理，提取前 20 个高频词，具体如图 1 所示。

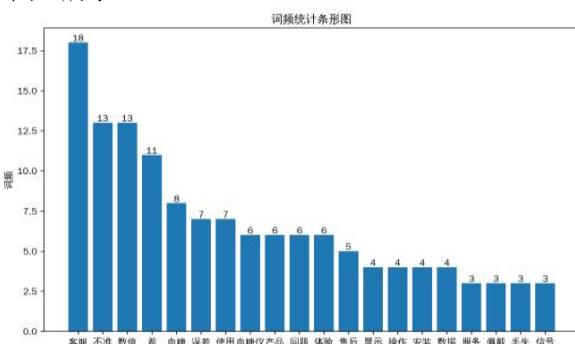


图 1 词频统计条形图

结果显示，用户对动态血糖仪的负向反馈主要集中在测量准确性、客服服务质量、使用体验三个方面。企业应优先解决测量准确性问题，减少产品的测量误差，确保产品核心功能的可靠性，同时优化客服服务，提高响应速度和解决问题的能力，并持续改进产品设计和使用体验，以满足用户需求，提升用户满意度和忠诚度。

### 2.2 基于 LDA 模型的产品评论分析

LDA 是基于 PLSA 的三层贝叶斯概率主题模型，它将文本视为词频向量，假定文档是多种主题的混合。LDA 学习算法以设定参数开始，在本研究中， $\alpha$  和  $\beta$  设置为自动调整 (auto)，主题数 K 设定为 4。根据分析，得

出聚类的 4 个主题的词云图，如下图 2-图 5 所示：



图 2 主题 0 词云图



图 3 主题 1 词云图



图 4 主题 2 词云图



图 5 主题 3 词云图

由图 2 可见,主题 0 聚焦于用户对产品整体体验的积极评价,关键词如“好”“不错”“喜欢”等直接体现用户满意情感,“物流”“发货”“客服”反映购买流程中的关键服务环节,“三诺”“东西”关联品牌及产品实体属性,整体表现出用户从购买到收货的正面感受及对品牌的印象。

由图 3 可见,主题 1 围绕血糖监测核心功能展开,涉及“老人”“安装”“佩戴”“监测”“数据”“简单”等词,体现产品针对老年用户群体的适用性,用户关注监测的准确性、数据的呈现与管理方式,以及操作的简便程度,反映动态血糖仪产品在特定用户群体中的功能表现和用户体验。

由图 4 可见,主题 2 深入探讨血糖仪的性能细节与使用感受,“比较”“动态”“误差”体现用户对测量精度和稳定性要求,“便宜”“喜欢”“佩戴”关联产品性价比、用户喜好及佩戴舒适度,展现用户在使用过程中的综合评价和期望改进方向。

由图 5 可见,主题 3 突出产品的技术创新点与差异化优势,“不用扎”“手指”“动态”“监测”等词表明用户认可产品减少采血痛苦、实时动态监测等新特性,“一代”“二代”等词反映用户对技术升级的关注和产品迭代的认可。

## 2.3 基于情感分析与 LDA 模型的用户反馈综合洞察

### (1) 产品优势的全面认可

LDA 分析的主题 0 与情感分析的正向评价相呼应,表明用户对产品的整体体验和购买流程中的服务环节感到满意,这与高比例的正向评价一致,说明动态血糖仪产品在市场上具有良好的品牌形象和用户体验。

### (2) 核心功能的适用性

主题 1 深入揭示了产品在老年用户群体中的功能表现,这与情感分析中用户对测量准确性和使用体验的反馈相契合。尽管正向评价占主导,但负向评价中的测量准确性问题提示我们在这一核心功能上仍有提升空间。

### (3) 性能细节与用户体验

主题 2 中的关键词如“误差”“便宜”“佩戴”等,与情感分析中用户对使用体验的反馈相呼应。这表明用户在追求性价比的同时,也对产品的准确性和舒适度有较高期望。

### (4) 技术创新的积极反馈

主题 3 中的关键词反映了用户对产品技术创新的认可,这与正向评价中的用户满意度一致,说明产品的差

异化优势是其市场竞争力的重要组成部分。

## 3 结论与建议

本研究基于电商平台动态血糖仪用户评论的多模型分析表明,用户对产品整体满意度较高,但测量精度偏差、客服响应滞后及操作复杂性问题构成主要负向反馈,而老年群体适配性、性价比优化与技术创新需求成为潜在市场增长点。通过 BERT-LDA 融合模型验证,有效识别隐性需求,为企业精准优化提供了技术支撑,为智慧健康设备优化提供了可复用的方法论框架。

## 参考文献

- [1] Manju Venugopalan, Deepa Gupta. An enhanced guided LDA model augmented with BERT based semantic strength for aspect term extraction in sentiment analysis[J]. Knowledge-Based Systems, 2022, 246: 108668.
- [2] 曾子明, 陈思语. 基于 LDA 与 BERT-BiLSTM-Attention 模型的突发公共卫生事件网络舆情演化分析[J]. 情报理论与实践, 2023, 46(09): 158-166.
- [3] 禹方中, 田宇阳. 基于 BERT 和 LDA 模型的酒店评论文本挖掘[J]. 计算机应用与软件, 2023, 40(07): 71-76+90.
- [4] 海骏林峰, 严素梅, 陈荣, 等. 基于 LDA-BERT 相似性测度模型的文本主题演化研究[J]. 图书馆工作与研究, 2024, 4(01): 72-79.
- [5] 胡泽文, 王梦雅, 韩雅蓉. 基于 LDA2Vec-BERT 的新兴技术主题多维指标识别与演化分析研究——以颠覆性技术领域: 区块链为例[J]. 现代情报, 2024, 44(09): 42-58.

作者简介: 武赛, 女 (1977.12-), 汉族, 籍贯江苏省徐州市, 博士, 讲师, 研究方向: 商业模式创新, 战略管理与战略联盟等。

何炳升, 男 (2003.12-), 汉族, 籍贯广东省揭阳市, 本科生。

岳梦雅, 女 (2004.01-), 汉族, 籍贯广东省东莞市, 本科生。

基金项目: 2025 年大学生创新创业训练计划项目 (xj2025118450402): 新能源汽车市场的情感驱动力: 基于 BERT 的文本分析与 TPB-NAM 模型的整合研究。