

电商平台用户评论的观点挖掘与情感分析 ——以爽肤水产品为例

郑兴涛, 徐德华

同济大学经济与管理学院, 上海

收稿日期: 2025年5月23日; 录用日期: 2025年6月6日; 发布日期: 2025年7月14日

摘要

在线评论作为消费者对产品和服务的主观表达, 具有重要的研究价值。本文以电商平台“天猫”的爽肤水商品的初评与追评为研究对象, 基于UIE模型对评论进行观点提取, 计算语义相似度并构建商品特征体系, 最后根据数据探究商品评论的关注度与满意度的分布规律。结果表明: 用户在初评时对使用体验最为关注, 在追评中用户对适用肤质和服务与物流的关注度提升较大; 追评的满意度普遍低于初评, 其中包装与设计特征的满意度降幅最大, 达到12.3%。本文为商家和平台提供了产品的优化建议, 为电商行业的健康发展提供一定参考。

关键词

在线评论, 细粒度情感分析, 预训练语言模型

Opinion Mining and Sentiment Analysis of User Comments on E-Commerce Platforms —Taking Toner Products as an Example

Xingtao Zheng, Dehua Xu

School of Economics and Management, Tongji University, Shanghai

Received: May 23rd, 2025; accepted: Jun. 6th, 2025; published: Jul. 14th, 2025

Abstract

Online comments, as subjective expressions of consumers towards products and services, have important research value. This article takes the initial and follow-up reviews of toner products on the e-commerce platform “Tmall” as the research object. Based on the UIE model, opinions are extracted

from the reviews, semantic similarity is calculated, and a product feature system is constructed. Finally, the distribution rules of attention and satisfaction of product reviews are explored based on the data. The results showed that users were most concerned about the user experience during the initial review, and in the follow-up review, users showed a significant increase in their attention to suitable skin types, services, and logistics; The satisfaction with the follow-up evaluation is generally lower than that of the initial evaluation, with the packaging and design features showing the largest decrease in satisfaction, reaching 12.3%. This article provides product optimization suggestions for merchants and platforms, providing a certain reference for the healthy development of the e-commerce industry.

Keywords

Online Comments, Fine Grained Sentiment Analysis, Pre Trained Language Model

Copyright © 2025 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

在线评论作为消费者对产品和服务的主观表达,不仅反映了用户的实际体验,还承载着消费偏好和市场反馈,具有重要的商业和研究价值。在电子商务领域中,评论包含对商品质量、服务态度等多维度的反馈信息,其观点结论与情感倾向会直接影响其他消费者的购买决策,因此,如何高效挖掘评论中潜在的高价值信息已成为业界共同关注的课题。

当前,评论观点的挖掘主要依赖基于规则的方法和机器学习技术。规则模板方法主要是通过手动编写规则或模板来匹配评论中的观点,依赖人工,因此泛化能力有限,难以适应多样化的表达方式;基于机器学习的方法,需要大量标注数据,模型在不同领域间的迁移性较差,因此该方法并不易于使用。同时,由于电商评论涉及多品类商品,再加上用户表达习惯差异显著、短文本信息的杂乱无章,使得传统的数据分析方法在处理大规模文本数据时更难准确识别出用户评论中的关键信息,在需要理解文本情感和细节时更是如此。

相较于其他技术,基于方面的情感分析(Aspect-Based Sentiment Analysis, ABSA)在理解用户表达的情感时会高效许多。ABSA不仅可以识别文本中的具体方面或属性,还能分析与每个方面相关的情感倾向,有效地从数据中提炼出观点主体和情感倾向,为商家优化产品与服务提供关键依据。许多学者在处理该领域的任务时采用了大型预训练语言模型,预训练语言模型实质性地改进了各种ABSA任务,并提升了ABSA模型的泛化能力和鲁棒性。

在当前电商领域的相关研究中,不少学者挖掘商品评论所隐藏的信息以及评论对消费者决策的影响[1][2],但少有学者关注初评与追评之间的差异。因此,本文主要研究两个问题:①大型预训练语言模型该如何更好地作用于电子商务领域的商品评论?②追评和初评之间的用户观点分布是否存在差异?为什么会存在这些差异?本文以国内典型的电商平台为研究对象,采集电商平台上的用户对于商品的初评与追评数据,运用预训练语言模型技术,构建商品特征库,并借助FastText模型来自动化地处理和分析大量用户评论,对评论进行细粒度观点挖掘,从文字数据中提取用户的真实体验和反馈。这样的分析不仅能够帮助商家以及平台更好地理解用户的期望,还能够指导他们在实践中做出更加精准和有针对性的改进措施,最终实现服务的持续优化和创新。

2. 国内外研究现状

在当前时代, 用户常常通过在线评论的方式表达自己对商品的观点, 这导致互联网上积累大量包含主观情绪的文本数据。针对电子商务平台上的评论进行分析, 能够有效地指导产品优化和服务的改进。然而, 由于文本内容的海量性质, 以人工的方式提取、处理其中蕴含的信息会显得较为困难。情感分析 (Sentiment Analysis) 旨在通过分析自然语言文本, 深入了解人们表达的情感、观点[3]。由于现实需求, 情感分析技术受到了广泛的学术关注, 在各个领域都有着重要的应用。

方面级情感分析, 作为一种更细粒度的情感分析方法, 关注于文本中特定实体、对象所携带的情感信息, 可以深度探索文本中各个维度的情感倾向, 从而更为精确地揭示产品的特性和与之相关的情感倾向。鉴于目前国内电子商务平台的快速增长, 再加上传统情感分析方法在处理短文本时效果较差, 方面级情感分析在商业领域具有巨大的价值和发展潜力, 成为了目前学者们研究的重点。为了更好地进行研究探讨, 申影利将方面级情感分析的主要任务界定为方面提取和情感分类两大方面[4]。

在方面词提取方面, 国外学者 Singh 在探究在线产品评论时, 借助机器学习算法, 成功将评论细化为形式、功能、行为、服务和其他五大特征类别, 并通过 k 折交叉验证验证了该算法的有效性[5]。Lee 则进一步利用卷积神经网络与迁移学习技术来精准提取方面词, 并在情感预测模型中充分考虑产品属性的相对重要性[6]。为解决方面提取过渡依赖人工标注, 覆盖领域少的问题, Georgios 提出一种大规模、多领域的反馈评论数据集, 使方面词的提取更加自动化[7]。

在情感分析领域, 方面级情感分析是为了从方面层面分析和理解人们的观点[8]。Jiang 等人为简化情感分析的语法联系, 设计了三条规则来确定文本在不同方面的“情感支持词”, 并验证其能显著提高情感分析能力。当前, 许多前沿的文本情感分析技术都已经采用了预训练模型[9]。Zhang 等人为解决符合任务中获得更加完整的方面级情感信息问题, 提出了一种新的分类法, 将预训练语言模型应用到其中[8]。其中, 一种备受瞩目的预训练模型是由 Google 提出的 BERT [10]。该算法使用深度学习技术和神经网络方法进行特征抽取和分类识别, 具有很好的扩展性、灵活性与高效性。BERT 采用双向 Transformer 技术创建了一个多功能模型, 该模型适用于处理多种自然语言任务, 而对于某些特定任务, 只需增加额外的神经网络层即可。该方法具有较强的普适性。在进行情感分析任务时, 众多学者采用了微调 BERT 模型, 并因此获得了明显的研究成果。Sun 及其团队将方面级情感分析任务巧妙地转换为句子分类任务, 并对 BERT 预训练模型进行了精准调整, 从而在 SentiHood 和 SemEval2014 数据集上都取得了当时的最佳表现[11]。

观点泛化是指将具体、个别的观点扩大为一般的观点, 目前常用的方法主要有基于 LDA 的主题挖掘模型以及基于语义相似度的分类模型。LDA 模型可以根据评论文本, 提取出评论的主题以及对应的关键词, 实现关键词的聚类; 基于语义相似度的方法主要采用 Word2Vec 和 FastText 等模型, 将抽取出来的特征词转换为多维度的词向量, 将观点词进行词向量表示, 并根据特征词库以及词向量距离进行分类和聚类。

在观点提取方面, 本文选取实验效果较好的 UIE 模型提取观点二元组。UIE 模型是基于 ERNIE 3.0 开发的预训练语言模型, 来自 ERNIE 3.0 的强大性能使得 UIE 在实体抽取、情感分析、关系抽取等方面具有良好表现。由于 LDA 模型只能提取主题词, 依赖词共现模式, 对短文本和稀疏数据效果差, 并不适合本文的研究内容, 因此本文选择使用 FastText 模型来对词语进行向量化表示并用于观点泛化。最后, 本文基于观点泛化的结果, 探究用户评论的观点的分布规律和初评与追评之间的关注度和满意度差异及其原因。

3. 研究框架与方法

基于预训练语言模型对电子商务平台用户评论数据进行二元观点提取以及观点泛化的研究框架如图 1 所示, 主要包括 5 个部分。

(1) 实验数据获取。本研究选择电子商务平台“天猫”作为数据采集来源, 该平台拥有庞大的用户基数, 可以提供充足的实验数据。在使用 Python 语言编写爬虫语言并获得所需的实验数据后, 对数据进行预处理等工作, 以获得可使用的数据集。

(2) 观点词提取。观点词的提取需要选择与研究任务相匹配的生成式预训练模型, 本文选择使用 PaddleNLP 开源的 UIE 模型, 利用数据集对 UIE 模型进行训练, 从用户的评论数据中提取出二元观点组<评论观点词; 情感倾向>。

(3) 商品特征库构建。商品特征库包括粗粒度特征和细粒度特征, 要构建一个合理的商品特征库, 需要参考主流电子商务平台中相关商品的参数设置。统计来自 UIE 模型的二元观点组中的评论观点词, 对词频进行逆序排名, 选择性地将排名前列的评论观点词作为商品特征库的细粒度特征种子词汇, 并参考拼多多、天猫等电商的商品参数, 总结出适合商品的粗粒度特征, 两者结合, 得到适用于产品的商品特征库。

(4) 商品特征体系构建。使用 FastText 模型将观点词进行词向量表示, 计算评论观点词与商品特征库中各词的语义相似度, 依据语义相似度进行特征词的分类, 将二元组泛化为<粗粒度特征; 细粒度特征; 中心观点词; 情感倾向>的观点四元组, 最终形成商品特征体系。定义词 W_a 与 W_b 之间的语义相似度如公式所示, 其中 K 表示词向量的维度, V_{ai} 为 V_a 词向量的第 i 个向量值。

$$Sim(W_a, W_b) = \frac{\sum_i^K V_{ai} \times V_{bi}}{\sqrt{\sum_i^K V_{ai} \times V_{ai}} \times \sqrt{\sum_i^K V_{bi} \times V_{bi}}}$$

(5) 观点挖掘与分析。分析二元观点组的泛化结果, 提取电商用户的关注度、满意度, 并探索其在初评与追评视角下的应用。本文将基于用户关注度和满意度数据, 探寻商品评论中潜在的高价值信息, 为商家和平台提供产品的优化建议。

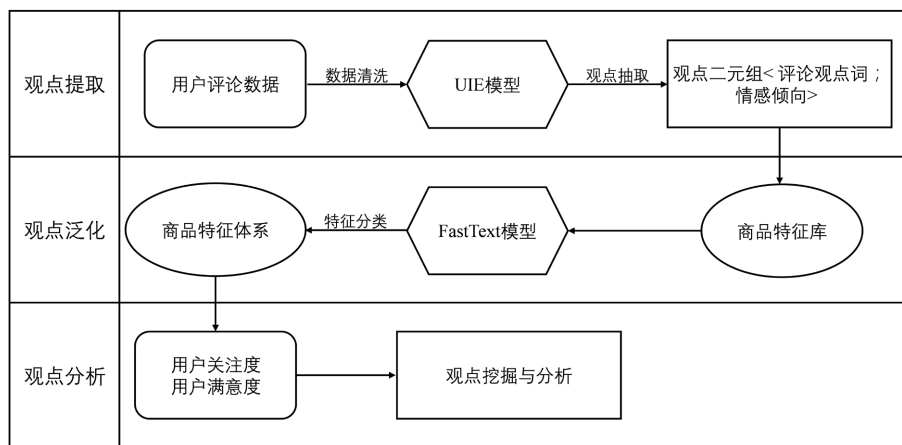


Figure 1. Research framework diagram

图 1. 研究框架图

4. 观点提取与观点泛化

本文选取天猫平台作为研究对象, 并对其“爽肤水”关键词下销量排名前 180 名的商品数据进行收

集, 包括商品详情信息以及用户的评论数据, 用户评论数据包括用户的初评、追评、评论间隔天数等字段。经过对数据的清理(例如删除重复评论或系统默认的评论), 收集到来自于天猫平台的共 2 万条初评与追评数据。

4.1. 建立商品特征库

首先, 利用 UIE 模型对经过预处理的数据进行二元观点提取, 为了对评论观点进行分类和过滤, 需要构建对应的商品特征库。参考拼多多、天猫等电商平台中爽肤水相关参数的分类与设置, 并结合高频评论特征词, 初步构建商品特征库。

通过分词和词频统计, 可以发现少量观点词只出现过一次, 高频词汇占比非常高, 同时低频词汇往往是无法代表评论的词汇, 如“小孩”“企业”“一瓶”等, 因此对于低频观点词汇进行了筛选、去除, 最终得到 12,796 个观点词以及对应的情感倾向。高频词语涵盖了多种词性, 其中一些并不能代表商品本身的特征, 在将这些词汇纳入初始特征词集进行分析时, 可能会遭遇众多的干扰元素, 为了减少影响, 也为了更好地了解用户对产品和服务的看法和态度, 有必要对初始特征词集进一步筛选。通过筛选得出词频统计排名前 40 的词, 并对结果进行了人工审核, 以选择适合的词语作为特征种子词, 并依据筛选结果初步构建包含粗粒度特征、细粒度特征 2 个维度的商品特征库, 如表 1 所示, 粗粒度特征主要划分为核心功效、使用体验、包装与设计、性价比、气味、适用肤质、服务与物流 7 个维度。

Table 1. Product feature library
表 1. 商品特征库

粗粒度特征	细粒度特征
核心功效	保湿效果, 补水效果, 控油效果, 美白效果, 祛痘效果
使用体验	肤感, 质地, 滋润, 清爽, 温和性, 刺激性
包装与设计	喷头, 包装, 瓶身, 容量
性价比	价格, 性价比, 赠品
气味	香味, 味道
适用肤质	油皮, 敏感肌, 干皮, 混合肌
服务与物流	物流速度, 客服态度

4.2. 建立商品特征体系

基于爽肤水商品特征库, 使用 FastText 模型将其余评论特征词向量化, 计算评论特征词与特征库中各词汇的向量的余弦距离作为语义相似度, 并将评论特征词分类到语义相似度高的细粒度特征。比如对于“优惠”一词, 模型会将该词向量化表示, 并将向量化表示后的结果与“价格”“香味”“物流速度”等细粒度特征的词向量进行语义相似度的计算, 最终得到该词与“价格”的语义最为相似, 并将该词划为“价格”这一细粒度特征的范围中。最终获得由 7 个粗粒度特征、26 个细粒度特征以及多个评论特征词构成的三维商品特征体系, 如表 2 所示。

Table 2. Product feature system
表 2. 商品特征体系

粗粒度特征	细粒度特征	评论特征
核心功效	保湿效果, 补水效果, 控油效果, 美白效果, 祛痘效果	产品效果, 使用效果, 保湿能力, 美白, 吸收速度, 保湿程度, 肤色效果……

续表

使用体验	肤感，质地，滋润，清爽，温和性，刺激性	体验感，使用感，皮肤状态，质感，作用，修复效果，滋润度，触感……
包装与设计	喷头，包装，瓶身，容量	用量，瓶装，空瓶，设计，包装，泵头，水泵，喷嘴，份量……
性价比	价格，性价比，赠品	价格，价钱，优惠，活动，平价，性价比，小样……
气味	香味，味道	香精，味道，成分，酒精味，清香……
适用肤质	油皮，敏感肌，干皮，混合肌	油皮，皮肤，干裂，痘痘，粉刺，过敏……
服务与物流	物流速度，客服态度	服务，速度，物流速度，客服态度，服务……

5. 观点挖掘与分析

5.1. 观点整体分布

通过计算观点二元观点组集合中各个维度、特征的正向、负向观点数以及不同情感倾向的观点占比，可以看出用户对于爽肤水这款产品的关注度以及满意度的分布，结果见图 2，可以看出：① 最受用户关注的特征是使用体验，占比 26.6%，相比其他特征高出许多。爽肤水是一款典型的体验型产品，用户需要先使用才能得到对产品的评价，因此用户在初评时最为看重产品的使用体验；② 用户对核心功效关注度最低，仅占 7%，这可能是因为爽肤水在护肤流程中被默认为“辅助步骤”，用户对其核心功效(如美白、抗老)的预期远低于精华或面霜，因此对爽肤水的核心功效心理阈值低，只要满足基础功能(保湿、清爽)即可达标。

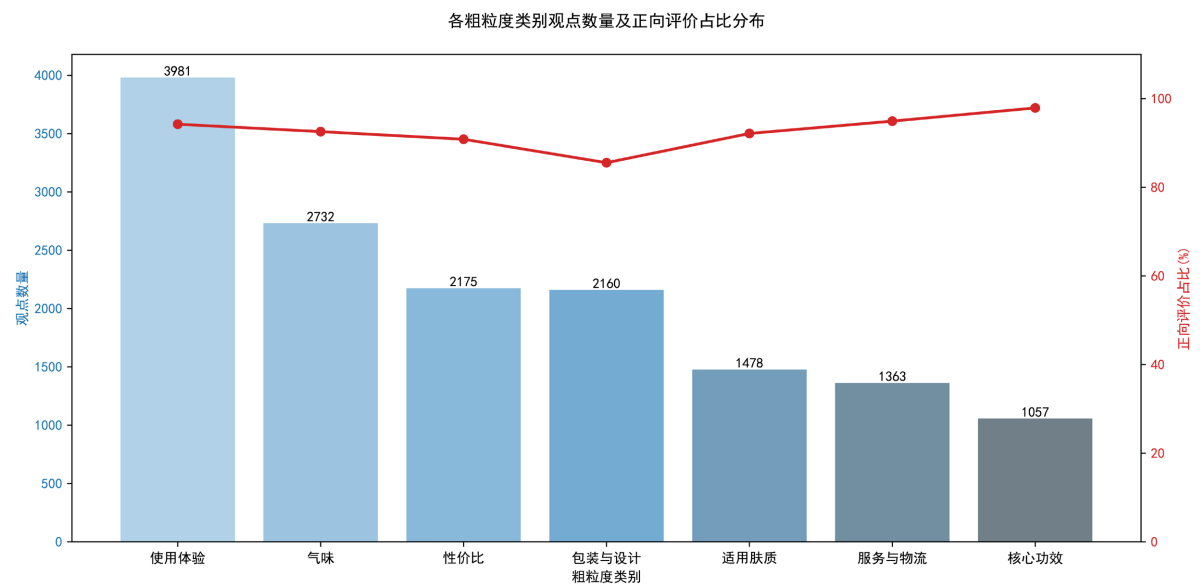


Figure 2. The attention to each feature of the product and the proportion of positive reviews
图 2. 商品各特征关注度以及好评率占比

在用户满意度方面，可以看出，用户对于大部分特征的满意度都在 90%以上，其中核心功效的满意度最高，达到了 97%，这可能是因为品牌深谙用户对爽肤水的低功效期待，因此主打“安全牌”，宣传产品无酒精或者温和配方，精准满足基础需求，反而减少用户的差评。对满意度最低的包装与设计特征

下的对应观点词进行聚类, 得到其负面观点主要是对瓶身尤其是瓶盖的差评, 具体体现在“瓶盖难打开”“瓶子品控差”等, 这表明未来商家需更加注重爽肤水的包装与设计, 让用户易于使用。

5.2. 基于初评与追评的观点分布对比

对观点泛化后的数据进行统计分析, 得到用户从初评到追评的关注度占比以及满意度变化, 其结果分别见图3、图4, 其中商品各特征关注度与满意度数据变化的原因分析如下: ① 关注度变化分析。从图3中看出, 在初评中, 用户关注度最高的特征为使用体验, 占比为28.9%; 而在追评中, 使用体验的关注度大幅下降, 气味、包装与设计、使用功效的关注度小幅下降, 性价比、服务与物流、使用肤质的关注度都有所提升。这可能是因为爽肤水的肤感、吸收速度、是否黏腻等体验是用户首次使用最直接的感受, 并且能快速感知、适合分享, 容易成为初评焦点。在长期使用后, 用户评价更偏向实际价值、长期适用性, 因此关注点发生变化。② 满意度变化分析。从图4中看出, 除了服务与物流, 其余特征在追评中的用户满意度均低于初评, 其中降幅最大的特征是包装与设计, 好评率下降了12.5%, 降幅最小的是核心功效, 仅有1.3%。将初评与追评间隔天数作为使用时长, 并从使用时长维度分析满意度的变化, 可以发现, 有三分之一的人在产品到货的当天就发表追评, 此时仅有4%的差评, 有三分之一的人在使用产品的9天内发表追评, 此时差评率有5.7%, 但随着使用时间的提升, 差评率会急速增加, 使用时长超过9天的用户对于产品的差评率达到14.6%, 这种现象在包装与设计这一特征上表现的尤为明显, 使用时长低于9天的用户的差评率仅有9%, 但时长超过9天的用户差评率达到24.6%。这可能是因为, 用户刚收到产品时, 情绪较为积极, 倾向于宽容评价。随着产品的深入使用, 颜值红利褪去, 用户对于产品感知逐渐偏向实际使用, 此时产品的实用性缺陷会暴露并被放大, 如滴管难吸、泵头卡顿、玻璃瓶太重等, 严重影响用户的体验与感受, 这提醒商家应更加注重品控, 优化实物的瓶身设计, 保证实物和后续使用体验与描述一致。

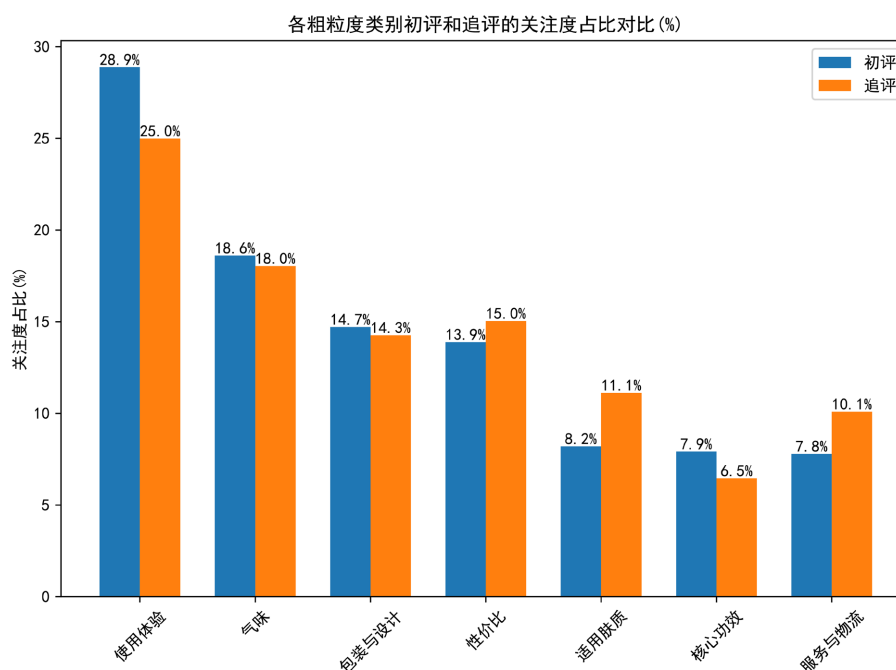


Figure 3. Proportion of initial and follow-up product reviews, attention, and positive reviews
图3. 商品初评与追评关注度以及正面评价占比

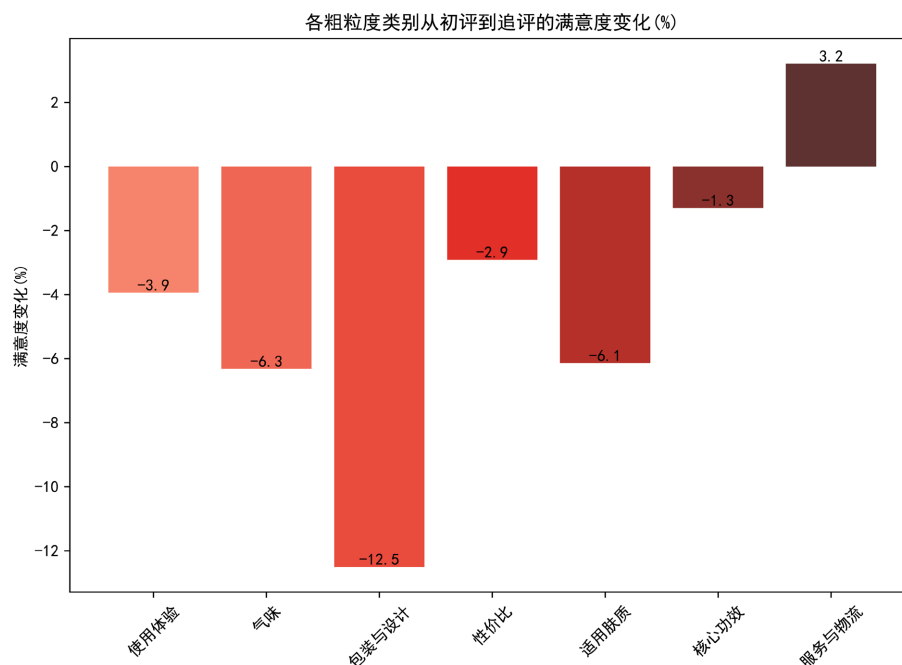


Figure 4. Comparison of user satisfaction from initial review to follow up review
图 4. 从初评到追评用户满意度对比

6. 结论与展望

本文提出了一种通过预训练语言模型并结合商品特征库来提取用户细粒度观点的方法, 并对电商平台上的用户评论进行了细粒度情感分析, 此外, 本文依据评论数据以及构建出来的商品特征体系, 对用户观点进行规律和差异分析, 得出如下结论: 从初评到追评, 用户对使用体验的关注度显著下降, 性价比、服务与物流、使用肤质的关注度有所提升。这表明随着使用时间的增加, 用户更倾向于评价产品的实际价值、长期适用性; 追评中的用户满意度普遍低于初评, 尤其是包装与设计特征的好评率下降明显。这一结果提示商家需重视产品的长期使用体验和品控管理, 避免此类现象的出现。

本研究仍然存在一定不足, 后续的研究可以从以下方向开展: ① 评论信息并不只有文本, 还包括图片、表情等非结构化数据, 本文只针对评论信息中的文本信息进行了分析与探索, 后续研究可以考虑将其他非结构数据纳入研究范围, 建立表情包对应的情感参数字典, 获得更完善、准确的评论观点与情感信息, 从而获得更准确的结论; ② 本文所构建的商品特征体系依赖人工, 需要人为筛选特征与特征种子词, 因此具备一定的主观性, 后续研究可以考虑依据数据自动构建商品特征体系的方法。

参考文献

- [1] Zhang, X., Zhang, C. and Wang, J. (2020) Research on the Influence of Online-Shopping Additional Reviews on Product Sales of Search Commodity and Experience Commodity. *Journal of Physics: Conference Series*, **1693**, Article 012105. <https://doi.org/10.1088/1742-6596/1693/1/012105>
- [2] Pan, Y. and Zhang, J.Q. (2011) Born Unequal: A Study of the Helpfulness of User-Generated Product Reviews. *Journal of Retailing*, **87**, 598-612. <https://doi.org/10.1016/j.jretai.2011.05.002>
- [3] 杨立公, 朱俭, 汤世平. 文本情感分析综述[J]. 计算机应用, 2013, 33(6): 1574-1607.
- [4] 申影利, 赵小兵. 基于深度学习的方面级情感分析综述[J]. 信息技术与标准化, 2020(Z1): 50-53+58.
- [5] Singh, A. and Tucker, C.S. (2017) A Machine Learning Approach to Product Review Disambiguation Based on Function, Form and Behavior Classification. *Decision Support Systems*, **97**, 81-91. <https://doi.org/10.1016/j.dss.2017.03.007>

- [6] Lee, Y., Park, J. and Cho, S. (2020) Extraction and Prioritization of Product Attributes Using an Explainable Neural Network. *Pattern Analysis and Applications*, **23**, 1767-1777. <https://doi.org/10.1007/s10044-020-00878-5>
- [7] Kontonatsios, G., Clive, J., Harrison, G., Metcalfe, T., Sliwiak, P., Tahir, H., *et al.* (2023) FABSA: An Aspect-Based Sentiment Analysis Dataset of User Reviews. *Neurocomputing*, **562**, Article 126867. <https://doi.org/10.1016/j.neucom.2023.126867>
- [8] Zhang, W., Li, X., Deng, Y., Bing, L. and Lam, W. (2023) A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges. *IEEE Transactions on Knowledge and Data Engineering*, **35**, 11019-11038. <https://doi.org/10.1109/tkde.2022.3230975>
- [9] Jiang, L., Zou, Z., Liao, J. and Li, Y. (2023) A Study on the Application of Sentiment-Support Words on Aspect-Based Sentiment Analysis. *International Journal of Pattern Recognition and Artificial Intelligence*, **37**, 1-23. <https://doi.org/10.1142/s0218001423570045>
- [10] Liang, Y., Meng, F., Zhang, J., Chen, Y., Xu, J. and Zhou, J. (2021) A Dependency Syntactic Knowledge Augmented Interactive Architecture for End-to-End Aspect-Based Sentiment Analysis. *Neurocomputing*, **454**, 291-302. <https://doi.org/10.1016/j.neucom.2021.05.028>
- [11] Sun, C., Huang, L. and Qiu, X. (2019) Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence. *Proceedings of NAACL-HLT*, **19**, 380-385.