

考虑消费者体验型产品特征偏好的评论排序研究

吕钰菁, 陈磊

东华大学旭日工商管理学院, 上海

收稿日期: 2023年10月29日; 录用日期: 2024年1月4日; 发布日期: 2024年1月15日

摘要

在线评论能够降低消费者在体验型产品购买决策中的感知风险, 但评论信息过载对高效的搜索构成了挑战。由于体验型产品具有个性化的特点, 消费者对评论信息的需求因产品特征偏好而异, 这一问题在现有研究中尚未得到充分阐述。研究通过LDA2Vec、LCR等方法, 构建消费者类模型及基于类的评论有用性预测模型, 从而实现评论个性化排序。实验结果证实了个性化排序模型在提高评论感知有用性上的有效性, 其在评论相关性、完整性、信息诊断性、消费者满意度等方面都显著优于有用性投票排序机制。本研究为如何识别体验型产品特征, 并通过消费者体验型产品特征偏好解决体验型产品评论排序提供了具体的方法, 更为通过评论个性化排序缓解评论信息过载问题提供了理论见解、模型成果和经验证据。

关键词

评论排序, 体验型产品特征, LDA2Vec, LCR

Review Ranking Considering Consumers' Preferences for Experiential Product Features

Yujing Lv, Lei Chen

Glorious Sun School of Business and Management, Donghua University, Shanghai

Received: Oct. 29th, 2023; accepted: Jan. 4th, 2024; published: Jan. 15th, 2024

Abstract

Online reviews serve as a potent tool in diminishing the perceived risk consumers' face when

making purchasing decisions regarding experiential products. However, the challenge lies in efficiently navigating through the deluge of review information, a task that becomes increasingly complex due to the personalized nature of experiential products. Consumer demand for review information fluctuates based on their preferences for specific product features, a facet that existing research has yet to fully explore. This study seeks to bridge this gap by constructing a consumer class model and a class-based review helpfulness prediction model, leveraging the capabilities of LDA2Vec and LCR to achieve a personalized review ranking. The empirical results underscore the efficacy of the personalized ranking model in enhancing the perceived helpfulness of reviews. It significantly surpasses the helpfulness voting ranking mechanism in terms of review relevance, completeness, information diagnosticity, and consumer satisfaction. This research offers a robust methodology to identify the features of experiential products and to tailor the ranking of their reviews based on consumer preferences for these features. Additionally, it provides theoretical insights, modelling outcomes, and empirical evidence to alleviate the issue of review information overload through personalized review ranking.

Keywords

Review Ranking, Experiential Product Features, LDA2Vec, LCR

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

在线评论是消费者关于产品或服务的意见表达,是潜在消费者的重要信息来源,也是平台提高用户粘性并实现可持续发展的关键。现有研究表明,在线评论的作用在体验型产品的购买决策过程中格外显著[1]。根据 Power Reviews 的一项研究,超过 85%的消费者在购买体验型产品之前会查看在线评论[2]。感知风险理论指出,由于体验型产品的质量在消费之前难以准确评估,因此人们对其感到的不确定性比搜索型产品更高。而在线评论能通过消费者的意见表达,有效降低消费者的感知风险[3]。然而,随着互联网的发展,体验型产品在线评论的数量飞速增长,海量的评论给消费者带来严重的信息过载问题。全球知名旅游分享网站 TripAdvisor 累计评论数超 8 亿条[4],携程的评论量累计已超 1.2 亿条[5]。因此,如何帮助消费者在大量评论中快速找到有用评论已成为有价值的研究问题。

解决评论信息过载的方法之一是评论排序。亚马逊最早推出“有用性投票”排序机制,平台根据有用性投票数对评论进行降序排列。虽然此举为亚马逊带来超过 27 亿美元的额外收入[6],但是有用性投票排序机制存在马太效应的问题——有用性投票数多的评论会得到更高的排名,而更高的排名也容易被消费者看见而带来更多的投票数,相反,较少的投票数评论则可能被埋没或忽略。实际上,获票数少的评论也可能对消费者“有用”。除了马太效应等问题,现有的评论排序还存在同质化的问题,即不同消费者看到的评论排序是一样的。这忽略了消费者对体验型产品的个性化偏好。事实上,不同的消费者对于同一评论的感知有用性是存在差异的,因此相同的评论排序并不能满足所有消费者的需求。

基于上述内容,本研究提出以下 3 个研究问题:

- 1) 如何全面地概括体验型产品的产品特征;
- 2) 如何根据消费者的体验型产品特征偏好构建个性化评论排序方法;
- 3) 个性化评论排序方法能否提高消费者的评论感知有用性。

为了回答这些问题, 我们采用了以下方法:

首先, 本研究利用 LDA2Vec 分析体验型产品的评论, 通过主题一致性确定最佳主题数量, 解决产品特征受限于研究人员先验知识的问题。同时, 利用 Word2Vec 实现产品特征向量化, 并利用 LDA2Vec 实现评论主题分配, 避免 LDA 忽略上下文语义的缺陷, 全面、有效地概括体验型产品的产品特征。其次, 引入 LCR 模型, 基于消费者体验型产品特征偏好实现消费者类建模, 并构建每个消费者类别的评论有用性预测模型, 基于差异化的有用性得分实现消费者个性化评论排序。最后, 通过 Hamming 距离和 Levenshtein 距离验证个性化评论排序方法下消费者评论感知有用性的提高, 为个性化评论排序提供有力证据, 为精准营销领域的研究提供新的视角和思路。

2. 文献综述

2.1. 在线评论有用性排序

评论有用性是一个备受关注的话题, 因为它反映了消费者对评论的价值感知程度。然而, 当前大多数关于评论有用性的研究主要集中在研究其影响因素。先前的影响因素研究主要集中在评论特征和评论发布者特征上, 如评论长度[7]、评论星级[8]、评论时效性[9]、评论者信息披露[10]以及是否为专家[11]等。近年来, 一些新的预测变量被引入到评论有用性的研究中, 如评论情感[12] [13]、图片[14]和评论可读性[15]。

随着评论信息过载问题日渐加剧, 评论有用性排序引起大家的研究兴趣。现有的评论排序研究大多针对评论者特性和评论内容特性提出综合排序方法[16] [17]。还有部分研究基于机器学习通过训练大量的评论数据集, 结合文本特征、用户属性和社交网络等信息来提高排序的准确性和效率[18] [19]。在研究内容方面, 不少研究的重心在于关键指标的选取和权重的分配上。郭顺利等在餐饮评论的排序上, 通过文献调研和专家咨询选取在线评论有用性的 8 项指标, 采用模糊层次分析法和加权灰色关联分析进行指标赋权和排序计算[20]。张艳丰等则以亚马逊手机评论为例, 结合并改进 TOPSIS 分析法构建评论有用性排序模型[21]。

综合来看, 现有评论排序研究依然存在不足。在评论有用性影响因素模型上, 大部分模型只考虑评论特征和评论者特征, 少有研究考虑到消费者特征, 仅有的研究也只是根据先验知识将消费者分为“专家”或“新手”[22], 而未充分考虑到消费者多样化的分类。换言之, 现有的评论排序模型大多是静态的, 无法适应消费者的动态需求, 即消费者的个性化需求。在研究内容上, 不少研究使用了机器学习算法来预测评论有用性排序, 但目前大多数模型都是监督学习模型, 这意味着它们需要大量的标注数据。这对于少量产品评论排序甚至单一产品的评论排序而言适用性较差, 且基于机器学习或深度学习的模型的可解释性也较差, 人们难以从中得到实践启发。因此, 真正有用的评论排序应该考虑到消费者的个性化需求和偏好, 以提供更加个性化和精准的评论推荐, 且无论是单一产品或多产品均能适用。

2.2. 体验型产品特征

体验型产品和搜索型产品的区别在于消费者在购买前能否获得完全的产品信息。搜索型产品的特征可以在购买前完全了解, 而体验型产品的特征则需要通过实际使用或体验后才能完全了解。在线评论作为消费者决策的重要参考依据, 在这两种产品类型中表现出明显的差异。

对于搜索型产品, 消费者可以根据产品广告等客观描述来获取信息, 评论更多作为一种补充内容辅助消费者决策。因为搜索型产品的特征则更为明确(如电脑的产品特征——显示、存储等), 消费者更容易比较和分析。相比之下, 体验型产品的特征通常比较抽象和主观, 例如美容产品的效果、餐厅的服务质量等。消费者需要通过其他消费者的评价和反馈来了解产品的真实情况。因此, 消费者更注重对产品使

用场景和目的的评价[23]。对于体验型产品, 消费者的使用经验和情感分享变得尤为重要, 这有助于消费者对其他消费者的评价产生信任感。因此, 体验型产品的个性化评论率远高于搜索型产品。

李宗伟和张艳辉以淘宝的在线评论作为实验对象[24], 通过实证分析来探讨不同类型商品在线评论的差异性。实验结果表明, 体验型产品的个性化评论率远高于搜索型产品。对于搜索型产品, 消费者的购买决策主要依赖于客观信息, 而评论的作用更多是作为一种补充内容; 而对于体验型产品, 消费者的购买决策主要依赖于其他消费者的评价和反馈, 评论的有用性主要体现在帮助消费者做出决策。

然而, 现有的大部分研究聚焦于不同产品类型的作用, 而以体验型产品偏好为立足点的研究较少。仅有的研究常常根据先验知识将评论阅读者分为“专家”和“新手”, 但这样的分类未考虑到消费者的体验型产品特征偏好。此外, 学者们通常采用 LDA 方法进行体验型产品特征提取, 但这种方法也存在一些问题。首先, LDA 主题模型没有充分利用词间语义关系和上下文语境, 因此会导致主题可解释性差的问题; 其次, 大部分学者在 LDA 模型的使用中提前设定了主题数量, 但这对于体验型产品特征来说并不十分适用。鉴于体验型产品特征个性化较强的特点, 学者难以预先设定其主题数量, 换言之, 预先设定的主题数量往往融入了较强的研究者个人的主观因素, 无法客观、全面地反映消费者的个性化偏好。

3. 数据和模型

为了回答研究问题 1, 本文基于携程酒店数据, 采用 LDA2Vec 技术对酒店评论进行分析, 通过主题一致性确定最佳主题数量, 并结合 Word2Vec 使特征向量化, 从而全面、客观地获取体验型产品特征, 同步获取评论情感极性、可读性等特征; 为了回答研究问题 2, 本文将 LCR 引入体验型产品评论领域, 实现基于消费者体验型产品特征偏好的消费者类构建, 并构建各类别消费者的有用性预测模型, 而后将评论根据有用性得分排序。总的来说, 本文通过以下步骤实现考虑消费者体验型产品特征偏好的评论个性化排序: 体验型产品特征提取、消费者类建模、评论个性化(如图 1)。

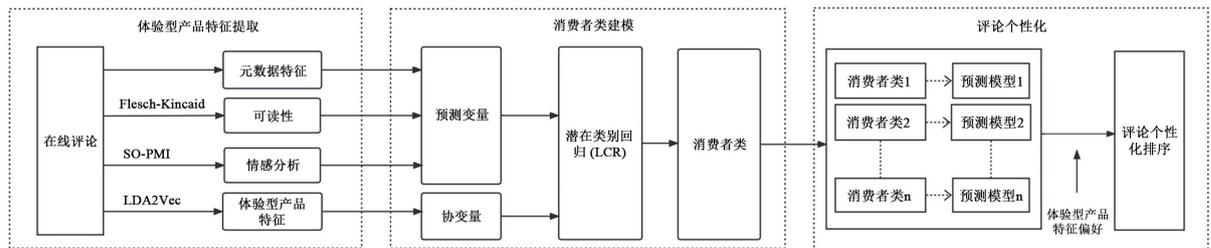


Figure 1. Personalized review ranking considering consumers' experiential product feature preferences

图 1. 考虑消费者体验型产品特征偏好的评论排序

3.1. 体验型产品特征提取

本研究以酒店评论为例, 参考《HOTELS》2022 年酒店排行及各酒店品牌在中国的受众情况, 选取豪华型酒店。平台选择方面, 本研究最终以携程为数据平台, 因为携程作为全球第三、亚洲最大的 OTA 平台, 涵盖的酒店数据丰富, 拥有完善的在线评论系统和海量的评论数据, 常作为国内学者的研究对象[25]。最终, 本文通过 Python 爬虫获取上海外滩某著名酒店的评论, 共计获得 2507 条评论。对评论进行清洗和预处理后, 共计获得 2275 条有效评论。

在体验型产品特征提取上, 本文首先基于 LDA 模型揭示酒店评论的潜在主题, 即酒店评论的体验型产品特征。相较于现有的预先设定主题数量的研究, 本研究结合主题一致性, 通过相似度计算确定最佳主题个数, 从而确定最佳体验型产品特征, 这可以避免在确定主题数量时受限于研究人员先验知识的问

题。再通过 Word2Vec 实现主题和评论的向量化, 最后计算评论在各体验型产品特征下的分布。

最终, 本研究确定了 6 个主题并确定了每条评论在各产品特征下的分布情况。每个主题涵盖了一些高频词汇, 这些词汇对应了不同的产品特征。各变量及其所涵盖的高频词汇如表 1 所示。

Table 1. Product feature and words

表 1. 产品特征及高频词汇

主题	产品特征	词汇
1	酒店管理	专属、经理、管家、保安、管理、培训、卫生
2	酒店服务	服务、前台、热情、态度、工作人员、专业、体验
3	总体 & 性价比	性价比、总体、五星、品牌、水准、整体、高级
4	环境 & 设施 & 位置	外滩、环境、设施、建筑、交通、地理位置、步行
5	景观 & 餐饮	早餐、江景、风景、景观、味道、单点、口味
6	房间内部	房间、升级、套房、客房、电视、浴室、双床

本研究以大连理工大学汉语情感词典为基础情感词典, 整合程度副词和否定词等, 并结合点互信息 (SO-PMI) 对基础词典进行情感极性判别, 人工整理后得到酒店领域情感词典, 从而准确判断情感极性。该词典包含 2177 个情感词, 其中褒义词 1341 个, 贬义词 836 个。除了情感极性, 评论的可读性也是预测模型的重要自变量之一。评论的可读性反映了评论在多大程度上能被消费者理解。除此之外, 文本长度、评论星级、是否含图片、评论发布天数、评论者的历史评论数等也被证明在评论有用性预测中有显著作用, 因此本文也将这些变量作为评论有用性预测模型的自变量。最终, 本研究的预测模型包括了评论基本信息、情感极性和可读性等多类自变量, 具体如下表 2 所示。

Table 2. Parameters of prediction model

表 2. 预测变量

变量	均值	变量描述
Length	64.90	文本长度
History_review	11.45	评论发布者的历史评论数
Star_rating	4.58	评论星级
Image	0.33	评论是否包含图片
Day_elapse	568.51	评论发布天数
Sentiment_score	3.43	评论情感
Readability	15.06	评论可读性

3.2. 消费者类建模

传统的评论有用性预测研究通常将消费者分为预先定义的类别, 例如专家和新手消费者。本研究采用 LCR, 根据消费者在体验型产品特征上的偏好, 将消费者分为不同的类别。这种方法可以更精准地反映消费者之间的差异, 从而提高评论有用性预测模型的准确性。具体来说, LCR [26] 具备的优势如下: 第一, 该模型可以识别观测变量下的隐藏的异质性, 而传统的线性回归和判别分析等模型仅描述观测变量之间的关系。第二, LCR 模型还能同步地为每一类别的消费者生成基于类的评论有用性预测模型,

它的优势是无需执行通常的第二阶段的判别分析以将识别的类别与协变量相关联, 同样也不需要传统聚类方法(如 K-Means)来确定聚类数目, 毕竟这一方法可能会受到启发式先验的影响。第三, LCR 模型可以处理非平衡数据集, 由于大部分在线评论的有用性投票数为 0, 不同类别的样本数量差异很大, 传统的分类模型在处理非平衡数据集时容易出现偏差和误差。最后, LCR 模型的可解释性强, 即通过分析模型的参数来解释模型的预测结果, 这有助于更好地理解模型的内在机制。

具体来说, 假设消费者被归类于一组类别 C 中的一个, 其概率分布表示为 $\{\pi_c | c=1 \dots C, 0 \leq \pi_c \leq 1\}$, 不同类别中的成员资格未知但已知 $\sum_{c=1}^C \pi_c = 1$ 。每个类别具有特定的回归模型来预测观察向量 Y_i , 即评论有用性预测得分。类成员模型依据消费者关心的不同体验型产品特征而将消费者归为不同的类别, 并为不同类别的消费者构建基于类别的评论有用性预测模型。本研究采用的类成员模型是基于多项式逻辑回归的模型, 写作:

$$P(c|Z_n) = \exp(\theta_c Z_n) / \sum_{k=1}^C \exp(\theta_k Z_n) \tag{1}$$

其中, Z_n 是消费者 n 的体验型产品特征向量, 这决定了消费者 n 归属于哪个类别, θ_c 为类别 c 的系数向量, C 为类别总数。

基于类别的观测变量预测模型如下:

$$P(Y_k | X_{k,c}) = \exp(\beta_c X_k) / \sum_{k=1}^K \exp(\beta_c X_k) \tag{2}$$

其中 $k=1 \dots K$, K 代表评论总数, Y_k 是关于评论 k 的评论有用性预测得分, $P(Y_k | X_{k,c})$ 表示基于各消费者类别预测所得的评论有用性得分的概率。因此, 一条评论的有用性等于消费者 n 属于 c 类的概率乘以 c 类成员条件下该评论的有用性概率, 再求和。

本文借助 LatentGold [27] 软件, 分别构建 1~5 个类的模型, 结果如表 3 所示。结果显示, Model 3 的 BIC 最小, 并且该模型的 R^2 值达到 0.9412, P 值 $2.2e-7 (< 0.001)$, 显著性强。因此, 我们选择 Model 3。

Table 3. Model fit statistics

表 3. LCR 模型选择

	Number of class	LL	BIC (LL)	p-value	R ²
Model 1	1-class regression	-1724.1727	3510.1833	2.0e-52***	0.5969
Model 2	2-class regression	-1368.9529	2915.6897	5.7e-12***	0.9026
Model 3	3-class regression	-1295.6196	2884.9691	2.2e-7***	0.9412
Model 4	4-class regression	-1265.6902	2941.0565	4.3e-6***	0.9566
Model 5	5-class regression	-1253.5291	3032.6802	7.2e-6***	0.9569

Notes: *** $p < 0.001$ 。

同步地, LCR 模型在迭代中也生成了类成员模型, 如表 4 所示。结果显示, 景观 & 餐饮、环境 & 设施 & 位置、酒店管理这 3 个特征在分类模型中的贡献较大, 这 3 个变量对消费者分类有显著影响。结合公式(1), θ_c 代表不同类别的系数, 由于分母均相同, 可简化理解为 θ_c 越大, 属于该类的概率也越大。举个例子, 当消费者只关注景观 & 餐饮模块(即, 给定特征权重关注矩阵为(0, 0, 0, 0, 1, 0), 则他被归为 class 1 的概率也越大, 因为 class 1 景观 & 餐饮的系数最大。值得注意的是, 成员类判别模型给出的是类别概率, 这并不意味着消费者一定归属于某类。

Table 4. Coefficients for the class membership model
表 4. 类成员模型系数

	Class 1	Class 2	Class 3	Wald	p-value
Intercept	-0.0231	0.9495	-0.9264	23.0895	9.7e-6***
酒店管理	1.1519	-0.5120	-0.6399	8.5540	0.014*
酒店服务	0.5858	-0.2812	-0.3047	3.0349	0.22
总体 & 性价比	0.6267	1.4176	-2.0442	1.0805	0.58
环境 & 设施 & 位置	0.9652	-1.2302	0.2650	10.4076	0.0055**
景观 & 餐饮	1.6343	-0.5228	-1.1115	17.3510	0.00017***
房间内部	-0.1043	-0.1423	0.2466	0.3354	0.85

Notes: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

基于类的评论有用性预测模型则如表 5 所示。结果显示, 历史评论数、评论星级、是否含有图片、文本长度、评论发布天数、情感极性、评论可读性均对评论有用性预测有显著影响。

Table 5. Coefficients of the review helpfulness prediction model
表 5. 评论有用性预测模型系数

	Class 1	Class 2	Class 3	Wald	p-value
Intercept	2.1939	0.5510	-10.1811	144.7279	3.6e-31***
History_review	-0.0085	0.0051	0.0292	129.2253	7.9e-28***
Star_rating	-0.8576	-1.1759	0.5357	900.4561	7.0e-195***
Image	1.5356	1.5098	5.6904	345.0955	1.7e-74***
Length	0.0049	0.0028	0.0082	371.5328	3.2e-80***
Day_elapse	-0.0019	0.0027	0.0027	321.3388	2.4e-69***
Sentiment_score	0.0304	-0.0198	-0.1030	31.1732	7.8e-7***
Readability	-0.0013	0.0043	0.0086	50.2403	7.1e-11***
R ²	0.9304	0.9057	0.9691		

Notes: *** $p < 0.001$.

3.3. 评论个性化

评论个性化的实现依赖于消费者类别 C 的分配以及基于该类别建立的回归模型。

首先, 根据消费者的产品特征偏好, 将消费者分配到消费者类别 C 中的某一类。然后, 基于类的评论有用性预测模型确定每条在线评论对该消费者的有用性得分, 即该评论对消费者的帮助程度。最终, 所有的评论将按照其帮助程度分数的降序进行排序, 参考公式(4)。这意味着对消费者帮助程度越高的评论将被排在前面, 而帮助程度较低的评论则会被排在后面。这样的排名方式能够更好地满足消费者的个性化需求, 提高评论的有用性和参考价值。

$$\text{Helpfulness}_c = \text{Constant}_c (\text{intercept}) + \beta 1_c * \text{History_review} + \beta 2_c * \text{Star_rating} + \beta 3_c * \text{Photo} + \beta 4_c * \text{Length} + \beta 5_c * \text{Day_elapse} + \beta 6_c * \text{Sentiment_score} + \beta 7_c * \text{Readability} + \varepsilon \quad (4)$$

4. 假设和检验

4.1. 实验设计

为了检验模型效果, 本文举行了一场实验。本实验的目的在于获取用户预期排序, 被试只需给出他们认为的最佳结果即可, 因此模拟实验的方法在不会带来高额的投入下就能很好地满足实验目的。在模拟酒店预定环境下, 被试被要求从评论相关性、完整性、诊断性和消费者满意度方面[28]对给定评论进行排序。通过比较“考虑消费者体验型产品特征偏好的评论排序结果和用户预期评论排序结果之间的距离”及“基于有用性投票的评论排序结果和用户预期评论排序结果之间的距离”, 证明考虑消费者体验型产品特征偏好的评论排序机制是否能提高消费者的评论感知有用性, 从而回答 RQ 3。

我们招募了 93 名被试, 每位被试仅能参与一次实验, 实验后给与一定报酬。排除 19 份不合格样本后, 最终得到 74 份有效样本。有效样本中, 46% 的被试每次预定酒店前都会看评论, 27% 的人大部分时候会看评论。

为了尽可能控制变量, 我们选择了评论长度在 90~110 之间的评论。评论阅读数量方面, 暂无研究给出明确阅读数量偏好, 本研究综合了各研究成果, 取相对折中的 13 作为评论数量[29] [30] REF_Ref141623953 \r \h [31]。在评论所属的酒店特征方面, 我们确保 6 个酒店特征下各涵盖 2~3 条评论, 且控制评论情感为褒贬均有。此外, 为了更直观得到结果, 我们选择的 13 条评论中大约二分之一是投票数多但有用性得分低, 另外二分之一则是高有用性得分但投票数少。考虑消费者体验型产品特征偏好的排序模型的优势在于, 投票数少但有用性较强的评论往往能被发现, 这些评论在按照有用性投票的排序中往往是被忽视的。

4.2. 实验过程

在实验中, 被试置身于模拟酒店预定环境下。以上海某高档酒店为例, 假设被试正通过在线预定的方式了解酒店信息并计划通过查看评论决定是否预定该酒店, 实验中明确说明假设排除价格、品牌、预定平台等偏好。我们要求被试对酒店的 6 个特征(酒店管理、酒店服务、总体 & 性价比、环境 & 设施 & 地理位置、景观 & 餐饮、房间内部)给出权重矩阵, 权重之和为 10, 权重越大代表该被试在查看评论时越关注该特征。

被试被要求对给定的 13 条评论分别从相关性、完整性、诊断性、满意度 4 个方面进行排序。为了避免评论展示顺序的影响, 所有评论均以随机顺序呈现。在评论排序上, 为了方便被试得出最终排序结果, 实验要求被试在系统性地看完所有评论后先将评论依照相关性高低(以相关性为例)分为高相关性、中相关性、低相关性 3 大类, 而后再对各个大类中的评论进行相关性由高到低排序, 最终综合得到最终 13 条评论的排序。这种做法有两个好处: 一方面, 它可以避免被试直接对 13 条评论进行排序时难以抉择, 因为通常情况下, 人们对第 8 名和第 9 名的敏感性并不高。将评论分成大类后, 被试只需要对每个大类中的 3~5 条评论进行排序, 这样可以减轻被试的负担, 避免长时间集中注意力。另一方面, 被试在对各个大类进行排序时, 相当于对评论排序进行了二次检查, 这有助于确保评论排序结果的准确性。

4.3. 结果讨论

我们将序列间距离视为评论结果之间的距离。为了得到更科学、权威的结果, 本文采用了两种序列距离计算方式计算评论排序的距离: Hamming 距离和 Levenshtein 距离。

Hamming 距离是最常用的计算等长序列间距离的计算方式[32]。它通过计算两个序列在相同位置上不同元素的个数来衡量它们的差异。该计算方式简单且直观, 计算的复杂度也较低, 计算速度快, 加之鲁棒性较强, 对噪声和缺失数据不敏感, 常被用于文本相似度计算。本文首先采用 Hamming 距离计算不

同排序结果(归一化后)之间的距离, 结果如下表 6 所示。其中, “个性化排序距离”代表考虑消费者体验型产品特征偏好的评论排序结果和用户预期评论排序结果之间的距离, “有用性投票排序距离”代表基于有用性投票的评论排序结果和用户预期评论排序结果之间的距离。

Table 6. Distance comparison based on Hamming
表 6. 基于 Hamming 距离的排序方法距离比较

	个性化排序 距离均值	有用性投票排序 距离均值	p 值
相关性	0.735	0.870	0.000***
完整性	0.719	0.854	0.000***
诊断性	0.714	0.822	0.002**
满意度	0.719	0.865	0.000***

Notes: *** $p < 0.001$; ** $p < 0.005$; * $p < 0.05$ 。

结果显示, 在评论相关性上, 考虑消费者体验型产品特征偏好的评论排序结果距离均值为 0.735, 小于基于有用性投票模型的距离(Mean = 0.870), p 值 0.000, 十分显著, 这表明个性化排序结果距用户预期的排序结果更近。同理, 在完整性、诊断性、满意度方面, 个性化排序结果也优于基于投票数排序的模型。

Hamming 距离虽然计算方式简单, 但其忽略了元素的顺序信息。与 Hamming 距离不同, Levenshtein 距离在计算过程中考虑了操作的顺序, 这种特性在考虑序列元素排列关系的评论序列距离计算中尤为有用。此外, Levenshtein 距离在考虑了序列中的插入、删除和替换操作时, 将三种操作看作不同的操作, 并分配不同的代价。这使得该算法能更准确地描述序列间的相似性和差异性, 在自然语言处理中的应用尤为广泛[33]。因此, 本文采用 Levenshtein 距离二次验证, 结果如下表 7 所示。

Table 7. Distance comparison based on Levenshtein
表 7. 基于 Levenshtein 距离的排序方法距离比较

	个性化排序 距离均值	有用性投票排序 距离均值	p 值
相关性	0.649	0.739	0.003**
完整性	0.667	0.743	0.005**
诊断性	0.613	0.721	0.001***
满意度	0.635	0.725	0.003**

Notes: *** $p < 0.001$; ** $p < 0.005$; * $p < 0.05$ 。

结果表明, 个性化排序结果在评论相关性、完整性、诊断性、消费者满意度上与用户预期结果的距离均小于基于投票的模型, 且 p 值均小于 0.05 或 0.001, 假设 1~4 得到验证。

综上, 无论是 Hamming 距离还是 Levenshtein 距离, 考虑消费者体验型产品特征偏好的评论排序在评论相关性、完整性、诊断性、满意度上均显著优于基于有用性投票的排序方式。

5. 研究贡献

5.1. 理论贡献

本研究主要有 4 个方面的研究贡献。首先, 在体验型产品评论领域引入了 LCR 模型, 为体验型产品

的评论排序研究丰富了新的思路并提供了新的理论依据。与以往基于 TOPSIS、AHP 的评论排序研究不同,本研究首次尝试将潜在类别回归(LCR)引入体验型产品评论领域,并通过实证检验证明了方法的可行性。这一研究不仅拓宽了 LCR 模型的应用领域,也丰富了体验型产品评论排序的研究思路。其次,在研究过程中,派生了消费者类,有助于识别不同消费者群体,拓展了精准营销领域的研究视角。传统的市场营销方法往往根据基本的人口统计学特征来划分消费者群体,本文通过消费者关注的不同产品特征,将消费者分类,细化了消费者群体的划分。此外,为酒店评论排序问题中如何识别产品特征并构建消费者产品特征偏好提供了具体的方法。本文尝试用 LDA 方法解决体验型产品难以精准描述产品特征的问题,并用 LCR 回归解决消费者细分的问题。这为酒店评论排序问题中的产品特征识别和消费者产品特征偏好构建提供了具体的方法和技术支持,为相关领域的研究和实践提供了新的思路和理论依据。最后,通过实证检验,揭示了产品特征和评论特征对评论感知有用性的显著影响。研究结论与先前的评论有用性影响因素研究是一致的,进一步证明评论长度、评论情感、是否有图片、历史评论数、评论可读性等因素对评论感知有用性的显著影响,这一发现为评论有用性研究提供了更充足的证据和支持,丰富了相关领域的研究内容。同时,也证明了体验型产品特征对评论感知有用性的显著影响。传统的评论有用性研究往往忽略了体验型产品的特殊性质,无法充分反映消费者对体验型产品的真实感受。本研究通过对酒店评论数据的分析,成功揭示了体验型产品特征对评论感知有用性的显著影响,为体验型产品评论研究提供了新的理论依据。

5.2. 管理启示

本研究在评论的分类、排名、呈现上具有重要意义。在消费者层面,本文基于消费者的个性化偏好进行评论排序,有助于帮助消费者快速找到有用评论,缓解信息过载问题,提高信息检索效率,同时也能协助消费者更为科学地进行酒店预订决策。在平台层面,本文的研究为在线评论平台提供了更加准确和个性化的评论管理服务。个性化评论排序有助于提高评论的质量和可信度,从而提高平台的用户粘性和满意度。在酒店层面,本研究的发现为企业提供了更加准确的评论管理和营销策略制定的依据。通过消费者个性化偏好识别,企业可以更加准确地把握消费者的需求和偏好,并依此进行个性化评论排序,从而制定更加精准的评论管理和营销策略,提高销售额和市场份额。

6. 局限性和未来展望

本研究在探究如何缓解消费者购买体验型产品时的信息过载问题取得了一定的研究进展。然而,本研究仍存在一些不足之处。首先,在消费者个性化偏好层面,本研究着重考虑了消费者产品特征偏好,而未考虑其他相对重要的特征变量;其次,在评论有用性预测模型方面,本研究采用的预测变量大多为评论特征,但在评论发布者特征方面相对欠缺;第三,本研究在实验过程中由被试主动提供产品特征偏好而缺乏进一步的自动化方法。

针对以上局限性,未来的研究可以从以下几个方面进行改进:首先,更加关注消费者的个性化需求和偏好,将更多的消费者特征纳入研究中;其次,采用更加全面和准确的预测模型,结合更多的技术手段和数据源;最后,可以结合用户行为日志分析等技术,自动识别消费者偏好,从而更好地缓解消费者购买体验型产品时的信息过载问题。

参考文献

- [1] Racherla, P.J. and Sajjan, S. (2012) A Comparison of Online Customer Reviews of Experiential and Search Goods. *Journal of Business Research*, **65**, 1249-1254.
- [2] Power Reviews (2021) The Power Reviews Consumer Study.

- <https://www.powerreviews.com/>
- [3] Li, X., Wang, D. and Liang, J. (2019) Why Do Consumers Read Online Reviews before Making Purchase Decisions? An Expectancy Theory Perspective. *Journal of Business Research*, **100**, 411-420.
- [4] TripAdvisor (2023) About TripAdvisor. <https://www.tripadvisor.com/TripAdvisorInsights/n2706/about-tripadvisor/>
- [5] Ctrip (2023) About Ctrip. <https://pages.ctrip.com/public/ctripab/abctrip.htm>
- [6] Spool, J.M. (2009) The Magic behind Amazon's 2.7 Billion Dollar Question. <http://www.uie.com/articles/magicbehindamazon>
- [7] 彭丽徽, 李贺, 张艳丰, 陈远方. 基于品牌声誉感知差异的在线评论有用性影响因素实证研究[J]. 情报科学, 2017, 35(9): 159-164.
- [8] 苗蕊, 徐健. 评分不一致性对在线评论有用性的影响——归因理论的视角[J]. 中国管理科学, 2018, 26(5): 178-186.
- [9] Ye, Q., Law, R. and Gu, B. (2019) The Impact of Review Recency, Frequency and Variance on Hotel Booking: A Mixed-Methods Approach. *Tourism Management*, **73**, 1-12.
- [10] Karimi, S. and Wang, F. (2017) Online Review Helpfulness: Impact of Reviewer Profile Image. *Decision Support Systems*, **96**, 39-48. <https://doi.org/10.1016/j.dss.2017.02.001>
- [11] Liu, X., Li, X. and Law, R. (2019) Expertise Matters: The Role of Expertise in Online Travel Review Evaluation. *Journal of Travel Research*, **58**, 1385-1399.
- [12] Wang, X., Tang, L. and Kim, E. (2019) More than Words: Do Emotional Content and Linguistic Style Matching Matter on Restaurant Review Helpfulness? *International Journal of Hospitality Management*, **77**, 438-447. <https://doi.org/10.1016/j.ijhm.2018.08.007>
- [13] 张紫琼, 叶强, 李一军. 互联网商品评论情感分析研究综述[J]. 管理科学学报, 2010, 13(6): 84-96.
- [14] Cheng, Y.H. and Ho, H.Y. (2016) Social Influence's Impact on Reader Perceptions of Online Reviews. *Journal of Business Research*, **68**, 883-887. <https://doi.org/10.1016/j.jbusres.2014.11.046>
- [15] 王丽, 徐红. 评论文本可读性对评论有用性的影响研究[J]. 情报理论与实践, 2019, 42(3): 90-96.
- [16] 熊涛, 王锦, 刘扬. 基于多特征融合的在线评论排序算法研究[J]. 数据分析与知识发现, 2021, 5(7): 95-103.
- [17] Wang, H., Li, Y. and Zhang, Y. (2021) Comment Helpfulness Prediction Using a Combination of Text Features and User Reputation. *Journal of the Association for Information Science and Technology*, **72**, 1066-1079.
- [18] Cheng, H., Liu, T. and Li, Y. (2019) Comment Helpfulness Prediction Based on a Deep Learning Model with Social Network Analysis. *Journal of the Association for Information Science and Technology*, **70**, 536-549.
- [19] Min, et al. (2017) Deep Learning for Automatic Sentiment Classification. *IEEE Transactions on Knowledge and Data Engineering*, **31**, 937-948.
- [20] 郭顺利, 张向先, 李中梅. 面向用户信息需求的移动 O2O 在线评论有用性排序模型研究——以美团为例[J]. 图书情报工作, 2015, 59(23): 85-93.
- [21] 张艳丰, 李贺, 翟倩, 等. 基于模糊 TOPSIS 分析的在线评论有用性排序过滤模型研究——以亚马逊手机评论为例[J]. 图书情报工作, 2016, 60(13): 109.
- [22] Zhang, R. and Tran, T. (2010) Helpful or Unhelpful: A Linear Approach for Ranking Product Reviews. *Journal of Electronic Commerce Research*, **11**, 220.
- [23] Liu, Y. and Li, H. (2014) Can Search and Experience Product Reviews Be Useful? An Information Processing Perspective. *Journal of Business Research*, **67**, 882-888.
- [24] 张艳辉, 李宗伟. 在线评论有用性的影响因素研究: 基于产品类型的调节效应[J]. 管理评论, 2016, 28(10): 123-132.
- [25] 单晓红, 张晓月, 刘晓燕. 基于在线评论的用户画像研究——以携程酒店为例[J]. 情报理论与实践, 2018, 41(4): 99.
- [26] Newman, M. and Dredze, M. (2015) Flexible Text Classification with Imprecise Labels. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Lisbon, 17-21 September 2015, 167-177.
- [27] McCutcheon, A.L. (1987) Latent Class Regression. *Journal of the American Statistical Association*, **82**, 338-347.
- [28] Mudambi, S.M. and Schuff, D. (2010) What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Quarterly*, **34**, 185-200. <https://doi.org/10.2307/20721420>
- [29] Vermeulen, I.E. and Seegers, D. (2009) Tried and Tested: The Impact of Online Hotel Reviews on Consumer Consideration. *Tourism Management*, **30**, 123-127. <https://doi.org/10.1016/j.tourman.2008.04.008>

- [30] Xiang, Z., Du, Q., Ma, Y. and Fan, W. (2017) A Comparative Analysis of Major Online Review Platforms: Implications for Social Media Analytics in Hospitality and Tourism. *Tourism Management*, **58**, 51-65.
<https://doi.org/10.1016/j.tourman.2016.10.001>
- [31] BrightLocal (2019) Local Consumer Review Survey.
<https://www.brightlocal.com/research/local-consumer-review-survey/>
- [32] 张焕炯, 王国胜, 钟义信. 基于汉明距离的文本相似度计算[J]. 计算机工程与应用, 2001, 37(19): 21-22.
- [33] Zhang, Y. and Li, X. (2020) A Text Sentiment Analysis Algorithm Based on Levenshtein Distance and Machine Learning. *Journal of Physics: Conference Series*, **1662**, Article ID: 012015.