E-Commerce Letters 电子商务评论, 2026, 15(4), 633-641
Published Online April 2026 in Hans. https://www.hanspub.org/journal/ecl
https://doi.org/10.12677/ecl.2026.154439

Hans汉斯

从UGC到精细化运营：基于电商评论挖掘的付费自习室关键属性识别与优化策略

颜志萍

江苏大学管理学院，江苏 镇江

收稿日期：2026年3月5日；录用日期：2026年3月17日；发布日期：2026年4月16日

摘　要

本研究针对付费自习室行业在O2O模式下面临的同质化竞争与效率挑战，构建了从UGC挖掘到运营决策的分析框架。通过采集平台评论数据，综合运用LDA、情感分析与IPA方法，识别并量化了环境、设施、服务、位置与体验五大核心属性。研究发现，“环境”是当前核心优势，“设施”为关键短板与改进优先项。基于此，依据4P理论提出了包含产品优化、分层定价、渠道融合及社群运营的数据驱动策略，为付费自习室从“流量运营”转向“价值运营”提供了实证路径与方法论参考。

关键词

付费自习室，用户生成内容(UGC)，文本挖掘，电子商务

From UGC to Refined Operations: Key Attribute Identification and Optimization Strategies for Paid Study Rooms Based on E-Commerce Review Mining

Zhiping Yan

School of Management, Jiangsu University, Zhenjiang Jiangsu

Received: March 5, 2026; accepted: March 17, 2026; published: April 16, 2026

Abstract

This study addresses the homogeneous competition and efficiency challenges faced by the paid study room industry under the O2O model, constructing an analytical framework from UGC

文章引用: 颜志萍. 从 UGC 到精细化运营: 基于电商评论挖掘的付费自习室关键属性识别与优化策略[J]. 电子商务评论, 2026, 15(4): 633-641. DOI: 10.12677/ecl.2026.154439

mining to operational decision-making. By collecting platform review data and comprehensively utilizing LDA, sentiment analysis, and IPA methods, five core attributes—environment, facilities, services, location, and experience—were identified and quantified. The study found that "environment" is the current core advantage, while "facilities" is a key weakness and a priority for improvement. Based on this, and using the 4P theory, a data-driven strategy encompassing product optimization, tiered pricing, channel integration, and community operation is proposed, providing an empirical path and methodological reference for paid study rooms to shift from "traffic operation" to "value operation".

Keywords

Paid Study Rooms, User-Generated Content (UGC), Text Mining, E-Commerce

Copyright © 2026 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

在终身学习与职场竞争加剧的背景下，付费自习室作为满足年轻人学习“刚需”的 O2O 服务电商迅速兴起。消费者通过美团等平台完成全流程消费，使得在线评论(UGC)成为影响其决策的核心口碑资产[1]。行业在经历用户规模快速增长后，正面临盈利模式单一、竞争同质化与客流周期性波动的严峻挑战，粗放运营难以为继，亟待向数据驱动的精细化运营转型[2]。当前，电商营销逻辑正从“人找服务”转向“服务找人”，这要求商家主动将 UGC 从反馈信息转化为驱动服务优化与品牌建设的战略资产[3]。

既有研究为本文提供了坚实基础，但也揭示了当前研究的不足与本文的切入点。在付费自习室研究方面，现有文献多聚焦于其社会属性、消费动机与环境设计[4]，但缺乏利用消费者在线评论(UGC)进行数据驱动的实证分析，难以动态捕捉用户真实、细粒度的体验反馈。在 UGC 与电商评论挖掘研究方面，利用 LDA 主题模型从 UGC 中提取服务属性[5]、通过情感分析量化满意度[6]、并借助重要性–绩效分析(IPA)进行诊断[7]，已成为电子商务与服务业领域洞察消费者需求的成熟范式。然而针对“学习型消费空间”这一新兴场景的评论语料，尚缺乏系统性的挖掘与分析框架。在精细化运营与数据驱动策略方面，尽管学界普遍认同基于用户洞察优化资源配置的价值，且已有研究尝试将 UGC 用于服务改善[8]，但在付费自习室等细分场景中，如何构建从用户评论挖掘到可落地运营策略的完整闭环模型，仍有待深入探索。

鉴于此，本研究旨在弥合上述研究缺口，系统回答两个核心问题：第一，如何利用电商平台 UGC，科学识别付费自习室服务的关键属性及其对满意度的相对影响权重？第二，如何将这些数据洞察转化为系统、可落地的精细化运营策略？为此，本文构建了“UGC 数据挖掘→关键属性与权重识别→IPA 绩效诊断→4P 整合优化策略”的闭环研究框架。本文的边际贡献在于：理论层面，首次系统整合 LDA、情感分析与 IPA 模型，构建了从 UGC 中自动提取属性、量化绩效并确定战略优先级的完整分析链条；实证层面，精准识别了付费自习室“设施”属性的关键杠杆作用，揭示了其“高重要性–低满意度”的显著特征，为理解该业态消费者满意度的动态演化提供了新证据；实践层面，提出了基于 4P 营销组合的、可操作且优先级明确的整合优化策略，实现了从“数据洞察”到“行动路线图”的直接转化。本研究以期为该行业从“流量运营”转向“价值运营”提供兼具理论深度与实践价值的参考。

2. 研究设计

2.1. 研究方法

本研究依次采用 LDA 主题模型、情感分析与重要性–绩效分析(IPA)三种方法，构建从属性识别到战略诊断的完整分析链条。首先，运用 LDA 主题模型对预处理后的评论文本进行无监督学习，提取出代表消费者关注核心的潜在主题，并通过语义归纳将其界定为具体的服务属性。其次，通过情感分析对归属于各属性的评论进行情感倾向量化，计算其情感得分均值，以此作为各属性的“绩效”表现指标。具体计算式为：对于每条评论其情感得分 S

$$S = \frac{\sum_{i=1}^{n} W_i * P_i}{n}$$

其中：n 为评论中匹配到的情感词总数；W_i 为第 i 个情感词的权重(积极词为+1，消极词为−1)；P_i 为第 i 个情感词的强度(基础强度为 1，根据程度副词加权，如“非常”为 1.5，“有点”为 0.8)。最后，引入重要性–绩效分析(IPA)模型，以各属性对应的评论数量占比衡量其“重要性”，与前述“绩效”值共同构成二维决策矩阵，从而将各属性划分为优势区、改进区、机会区与劣势区，明确资源优化与运营改进的优先顺序。

2.2. 数据来源与预处理

本研究的数据主要来源于中国主流生活服务电商平台——美团与大众点评。为保证样本的代表性与多样性，研究选取北京、上海、深圳、成都等八个不同经济与地域层级的城市作为样本区域，并以评论数不少于 100 条的付费自习室商户为分析对象。通过 python 进行网络爬取，共收集原始评论数据 39,820 条。

在数据清理过程中，首先删除了空值、乱码及明显无效的评论(如极短的“好评”“一般”等无语义内容)；其次，去除了重复记录及机器人评论；最后，对文本进行了规范化处理，包括全角半角统一、字符大小写标准化、错别字修正以及表情符号与特殊符号的去除。经过上述预处理后，获得有效评论 26,682 条，构建成后续文本挖掘分析的高质量语料库。整体数据的时间跨度覆盖近两年，兼具广度与时效性，能够较为全面地反映消费者对付费自习室服务体验的真实反馈。

2.3. 文本分析与特征挖掘

采用 Python 的 Jieba 分词工具(精确模式)对评论文本进行分词，并辅以自定义词典(纳入“自习室”等专有名词)与哈尔滨工业大学停用词表，以提升分词准确性并过滤无意义词汇[9]。进而，通过词频统计生成高频词表与词云图，见图 1。结果显示，“学习”、“自习室”、“环境”、“安静”、“氛围”等词频最高，初步表明消费者核心关注点集中于学习环境与氛围，且整体情感倾向偏积极。

Figure 1. Online comment word cloud
图 1. 在线评论词云图

为深入探究词语间的语义关联，利用 ROST CM 软件构建了社会语义网络图，见图 2。分析发现，“自习室”处于网络中心，与“安静”、“环境”、“干净”等环境属性词，“服务”、“热情”等服务属性词，以及“方便”、“地铁”等位置属性词紧密关联。这进一步验证了环境、设施、服务、位置及体验是构成消费者感知的核心维度，为后续的模型分析奠定了基础。

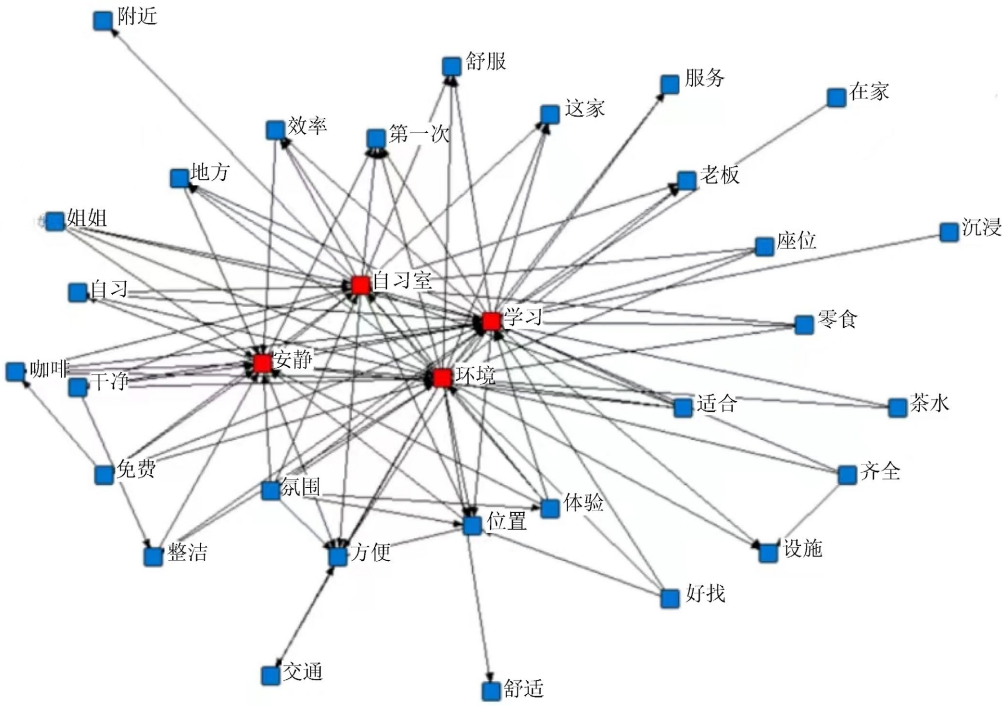

Figure 2. Online commentary social semantic network graph
图 2. 在线评论社会语义网络图

2.4. 研究方法

本研究依次采用 LDA 主题模型、情感分析与重要性－绩效分析(IPA)三种方法，构建从属性识别到战略诊断的完整分析链条。首先，运用 LDA 主题模型对预处理后的评论文本进行无监督学习，提取出代表消费者关注核心的潜在主题，并通过语义归纳将其界定为具体的服务属性。其次，通过情感分析对归属于各属性的评论进行情感倾向量化，计算其情感得分均值，以此作为各属性的“绩效”表现指标。最后，引入重要性－绩效分析(IPA)模型，以各属性对应的评论数量占比衡量其“重要性”，与前述“绩效”值共同构成二维决策矩阵，从而将各属性划分为优势区、改进区、机会区与劣势区，明确资源优化与运营改进的优先顺序。

3. 研究实施与结果分析

3.1. LDA 主题模型分析：关键属性识别

为从非结构化评论文本中识别消费者关注的核心维度，本文采用 LDA 主题模型对 26,682 条有效评论进行主题挖掘。通过计算不同主题数(k)下的困惑度与主题一致性得分，发现当 $k=5$ 时，模型在保持较低困惑度的同时，主题一致性得分达到局部最优，表明此时主题内聚性与区分度最为平衡，见图 3。因此，确定最优主题数为 5。

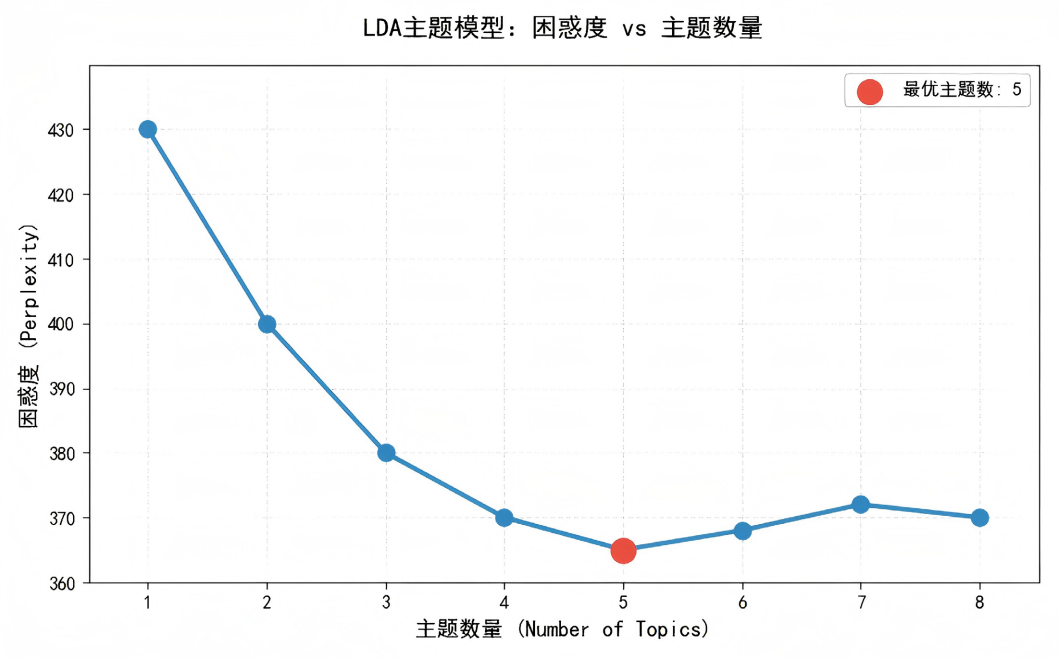

Figure 3. Theme consistency calculation results chart
图 3. 主题一致性计算结果图

Table 1. LDA topics and corresponding keywords
表 1. LDA 主题及对应的特征词

主题 1	概率	主题 2	概率	主题 3	概率	主题 4	概率	主题 5	概率
干净	0.066	位置	0.044	免费	0.038	小姐姐	0.089	学习	0.033
环境	0.053	方便	0.027	咖啡	0.037	服务	0.053	环境	0.033
推荐	0.018	好找	0.029	设施	0.029	前台	0.036	安静	0.016
舒适	0.016	交通	0.021	提供	0.026	热情	0.036	自习室	0.015
空调	0.013	吃饭	0.018	零食	0.024	管家	0.031	氛围	0.012
安静	0.012	地铁站	0.016	齐全	0.020	温柔	0.020	适合	0.010
座位	0.011	区域	0.011	茶水	0.018	态度	0.019	推荐	0.010
氛围	0.010	便利	0.011	贴心	0.013	贴心	0.018	体验	0.009
满意	0.008	附近	0.010	充电	0.012	及时	0.011	特别	0.009
透气	0.008	地方	0.008	自助	0.012	好评	0.008	下次	0.008

模型提取的主题及其代表性高权重特征词，结果见表 1。主题命名遵循以下逻辑：首先，提取每个主题下概率最高的前 15 个特征词如表 1 所示；其次，结合领域知识，归纳这些词汇共同指向的核心语义；最后，为每个主题赋予一个简洁、客观的概括性名称。具体而言：

主题 1 的特征词(如干净、环境、舒适、安静、氛围)主要描述学习空间的物理条件与整体感受，故命名为“环境”。

主题 2 的特征词(位置、方便、交通、地铁站、附近)聚焦于地理可达性与周边便利性，故命名为“位置”。

主题 3 的特征词(免费、咖啡、零食、提供、齐全)涉及商家提供的附加物品等基础设施，故命名为“设施”。

主题 4 的特征词(小姐姐、服务、前台、热情、态度)明确指向工作人员及其服务行为，故命名为“服务”。

主题 5 的特征词(学习、环境、安静、氛围、适合、体验)虽然部分词汇与主题 1 重叠，但其核心词“学习”、“体验”、“下次”等更侧重于消费者在此空间中的核心活动目的、主观综合感受及未来行为意向，因此命名为“体验”。主题 1 与主题 5 的部分重叠反映了物理环境是构成整体体验的基础，但主题 5 更侧重于从用户视角出发的整合性评价。

3.2. 情感分析：属性满意度评估

为进一步评估消费者对各属性的满意度水平，本研究在主题划分结果基础上开展评论情感分析。通过构建情感词典与方向标注模型，对各主题评论进行情绪极性判定与得分量化，结果见表 2。

Table 2. Statistical analysis results of sentiment analysis for each attribute
表 2. 各属性情感分析结果统计

主题	总评论数	积极评论数	消极评论数	中性评论数	情感得分均值
环境	122,66	10,473	1317	476	20.3789271
位置	1673	1433	165	75	18.782539
设施	1951	1608	254	89	10.4366311
服务	3416	2864	419	133	16.4578329
体验	2131	1809	235	87	8.2351971

结果显示，整体满意度处于较高水平，积极评论占比 84.8%，但属性间评分差异显著。其中，“环境”属性的评论数量与情感得分均居首，显示其为最显著的竞争优势；“位置”和“服务”的平均得分高于整体均值，说明这两个方面的消费者认同度较好。相对而言，“设施”和“体验”两个主题的情感得分显著低于平均值，成为主要的满意度短板，需要在后续经营中优先改进。这一结果表明，硬件完善度与个体体验感是制约付费自习室口碑提升的关键环节。

3.3. IPA 分析：战略优先级诊断

基于上述结果，本文进一步将五个属性的重要性(由评论比例表征)与绩效(由情感得分反映)纳入 IPA (Importance-Performance Analysis)模型，以识别各属性的战略优先级。分析结果见图 4。

由图可以看出，优势区(高重要性–高绩效)：环境。环境是付费自习室的核心竞争力与口碑基础，需保持现有高标准并持续精细化优化。改进区(高重要性–低绩效)：设施。该属性对满意度具有高影响力，但当前表现不佳，应优先投入资源改善硬件与辅助功能，以提升整体体验价值。机会区(低重要性–高绩效)：位置与服务。消费者当前对这两方面满意度较高，可作为差异化卖点延伸与营销强化方向。劣势区(低重要性–低绩效)：体验。讨论频率较低但得分偏低，可能与性价比感知及学习沉浸感不足相关，应通过优化设施与服务间的协同提升整体体验感知。综上，IPA 分析揭示了从“数据认知”到“策略行动”的逻辑转换路径。其中，“环境”是需保持的核心优势，“设施”是最迫切的优化突破口，而“位置”“服务”“体验”则分别承担品牌差异化与用户价值体验的潜在增长点。该结论为后续基于 4P 视角的精细化运营策略设计提供了明确的实证依据。

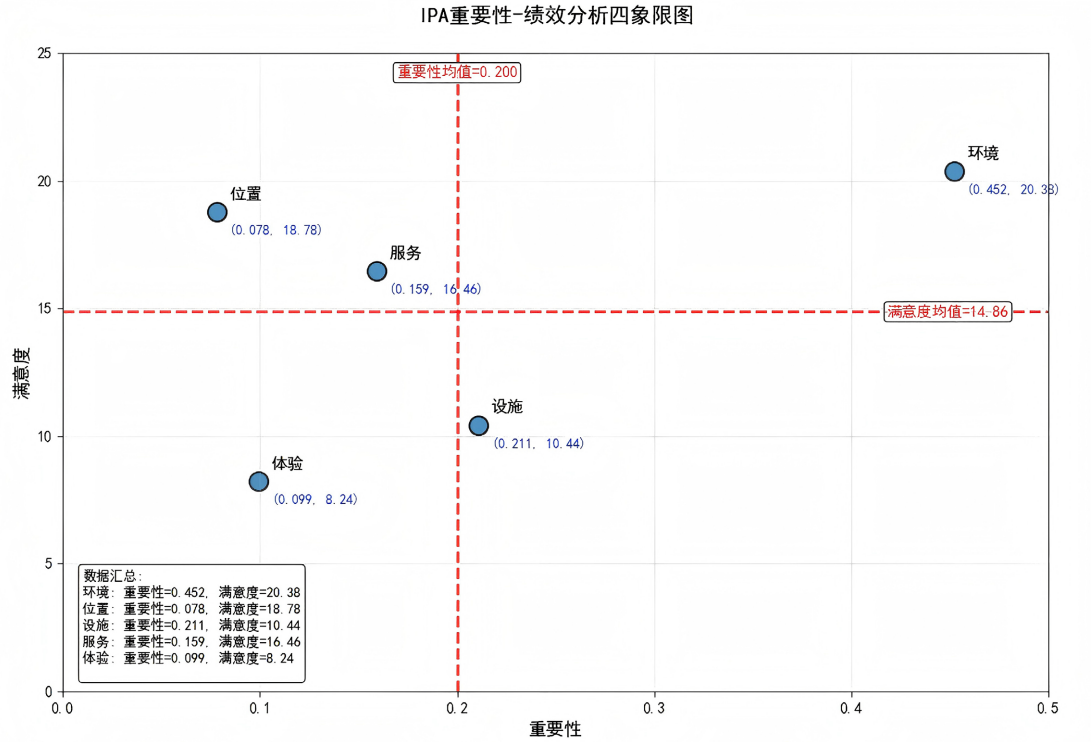

Figure 4. Consumer satisfaction IPA four quadrant chart
图 4. 消费者满意度 IPA 四象限图

4. 基于数据洞察的精细化运营策略

基于 LDA、情感分析与 IPA 模型的诊断结果，本研究识别出环境、设施、服务、位置与体验五大核心属性，并明确了其战略优先级。其中，环境为当前优势，设施为关键短板。为将数据洞察转化为可落地的行动方案，本文依据 4P 营销理论，构建如下精细化运营优化路径[10]。

4.1. 产品策略：巩固优势与补齐短板

强化环境核心壁垒，“环境”是当前高满意度与高关注度的优势区。因此，产品设计应持续强化环境体验的稳定性与差异化。确保空间的安静、整洁与舒适，形成稳定的“高质量学习氛围”；通过主题细分实现体验延伸，如划分“深度专注区”“轻讨论区”“放松休憩区”等功能分区，以满足不同学习状态与社交需求，提升场景多样性与空间使用效率。与此同时，“设施”是消费者高期待但低满意度的关键短板。重点投入设施改进，基础层面可优先升级照明、座椅、人机工学桌面、充电端口与网络稳定性等核心硬件，以减少物理性不适与中断感；增值层面可提升茶饮、储物、打印等附属设施的质量与便利度，塑造“效率友好”的学习环境。通过“环境稳固 + 设施升级”的产品策略组合，既能巩固优势认知，又可以可见改进带动整体满意度提升。

4.2. 管理启示价格策略：分层定价与价值重塑

“体验”维度满意度偏低主要受“性价比认知偏差”影响。为此，应在价格机制中引入价值导向与分层逻辑，实现价格与体验的再匹配。定价结构采用分区定价与时段定价，针对静音区、独立包间等差异化场景，按空间稀缺与体验深度设价位，结合时间实行动态调价，如工作日优惠、考试季溢价，提高时段利用率和坪效收益。其次，通过套餐与价值沟通策略重塑用户支付意愿，企业设计灵活卡种与会员

体系，将“优质环境”和“升级设施”整合为核心卖点，突出“学习效率提升价值”。以价值认知为中心的价格体系比单纯折扣刺激更能强化用户长期留存与品牌忠诚度。

4.3. 渠道策略：线上线下融合

“位置”被评价为高满意度但低关注度属性，表明“便利性”被用户视为基本前提而非差异化优势。因此，渠道策略应兼顾线下触点优化与线上流量运营，实现“物理接触 + 数字可达”的融合布局。线下方面，应聚焦高潜客群区域进行精准选址，如高校聚集区、办公区及交通枢纽，以降低通勤成本并贴近目标群体。通过智能预约、小程序导航与实时座位监控等功能，简化入店流程并提高资源利用效率。线上方面，应在美团、大众点评等主流生活服务平台强化内容运营，将“环境安静”“设施完善”“服务贴心”等关键词转化为可感知的品牌标签。通过强化 UGC 展示和高评分评论的精细运营，提升搜索曝光度与转化效率，从而实现线上引流与线下转化的闭环。

4.4. 促销策略：社群化与人格化运营

情感分析显示，用户在“服务”与“体验”主题中高度关注人际互动与归属感，表明促销活动的价值应从价格激励转向关系共建。首先，可通过人格化服务塑造提升场所温度。运营方应强化前线员工培训，树立“学习助理”或“空间管家”的角色形象，使服务过程更具陪伴感与目标导向性，从而提升信任与情感黏性。其次，应建立以学习为中心的社群化用户运营机制。通过会员社群、学习打卡、主题分享与互助活动，构建基于共同目标与轻社交互动的用户网络，激发正向口碑与自发传播动力。此类“社群式促销”不仅能延长用户生命周期，也能在无形中形成“同伴学习压力”，强化用户与空间的情感连接，实现口碑沉淀与品牌生态的自循环。综上，促销策略的本质不再是让利行为，而是社区文化与学习氛围的持续营造。通过社群共创与人格化服务的结合，付费自习室可在同质化竞争中塑造差异化形象与持续价值。

5. 结论与展望

本研究证实，电商平台 UGC 是解码付费自习室消费者满意度、驱动精细化运营的宝贵资源。结合 LDA 与 IPA 模型，识别出环境、设施、服务、位置、体验五大关键属性，精准定位“设施”为运营改进杠杆点。据此提出 4P 整合优化策略，为自习室创业者提供实操路线图。

不过，本研究有局限性。方法论上，IPA 分析用“评论比例表征”作属性“重要性”代理指标，可能与消费者真实重视程度不符，未来可结合问卷调研或实验法验证。数据源于单一电商平台 UGC，用户群体和评论风格有平台特定偏差，结论普适性需多平台数据检验。情感分析基于词典方法，对复杂语境理解有局限，未来可引入深度学习细粒度情感分析模型提升精度。

未来研究可从三方面深化：一是纵向追踪，分析关键属性重要性及满意度随市场周期的动态演变；二是将研究框架应用于不同城市层级比较，探索地域性差异化运营策略；三是探索前沿 AI 技术在 UGC 分析中的应用，实现前瞻性产品创新。总之，在服务电商成熟背景下，基于 UGC 的精细化运营能力是付费自习室及线下体验服务业构建核心竞争力的关键。

基金项目

国家社会科学基金项目“数字经济背景下平台垄断的多元共治研究”(项目编号：21BGL037)。

参考文献

[1] Lin, J., Li, L., Wang, C., Jiang, X. and Xiao, S. (2026) Uncovering the Role of Reviewer Heterogeneity in Customer Needs Identification and Customer Satisfaction Improvement: A Perspective from Target Customers. *Information Processing & Management*, **63**, Article ID: 104496. https://doi.org/10.1016/j.ipm.2025.104496

[2] 王皓, 王茂福. 付费学习空间: 消费时代城市青年的空间重构[J]. 城市发展研究, 2025, 32(3): 35-39.

[3] 汪仁琪等. 融合情感偏离度特征的评论有用性预测模型——基于深度学习框架[J]. 电子商务评论, 2025, 14(4): 1147-1154.

[4] 张智林. “格子间”里的梦想: 当代青年迈入共享自习室的文化逻辑[J]. 湖北社会科学, 2024(12): 87-94.

[5] 黄沈权, 朱晓辉, 陈子瑞, 等. 基于双向序列特征和主题语义模型的制造服务推荐方法[J]. 计算机集成制造系统, 2026, 32(2): 772-785.

[6] 聂慧如, 王建军, 阮佳乐, 等. 基于在线评论的节能家电用户情感分析与偏好研究[J/OL]. 系统科学与数学, 1-25. https://link.cnki.net/urlid/11.2019.o1.20260228.1445.008, 2026-02-11.

[7] 郑淳佳, 钱万惠, 杨清, 等. 基于语义情感的森林康养满意度分析——以西樵山森林公园为例[J]. 中南林业科技大学学报, 2026, 46(2): 204-214.

[8] Srianan, S., Nanthaamornphong, A. and Phucharoen, C. (2025) Analyzing Emotional Experiences through UGC: Insights from Bali and Phuket. *Journal of Quality Assurance in Hospitality & Tourism*. https://doi.org/10.1080/1528008x.2025.2484576

[9] 王克勤, 高智姣, 乔亚楠, 等. 在线评论中的用户需求识别及其演化趋势挖掘[J]. 机械科学与技术, 2023, 42(7): 1070-1080.

[10] Liu, H. (2025) Research on Digital Marketing Strategy Optimization Based on 4P Theory and Its Empirical Analysis. *Modern Economics & Management Forum*, **6**, 110-114. https://doi.org/10.32629/memf.v6i1.3564