

基于大数据分析的金融类APP用户需求研究

马英乔

大连东软信息学院健康医疗科技学院, 辽宁 大连

收稿日期: 2024年10月7日; 录用日期: 2024年10月22日; 发布日期: 2025年1月3日

摘要

在大数据时代,各大金融机构均在数字化转型的道路上持续发力,而面对移动互联网的挑战,数字金融已介入了金融机构各项业务的方方面面。本文运用Python算法对于金融机构APP的用户评论进行数据挖掘与分析,对收集到的30,142条文本数据展开LDA主题模型分析以找出用户的需求,并将用户需求划分为功能类、服务类和推广类。以自然语言处理中的SnowNLP算法了解用户的情感倾向,发现APP用户满意度不高,急需改善用户体验。对此,本文在量化分析的基础上提出了金融机构APP的优化建议,力图提升APP的整体运营。

关键词

金融机构, APP优化, LDA主题模型, 情感倾向

Research on User Demands of Financial APP Based on Big Data Analysis

Yingqiao Ma

School of Healthcare Technology, Dalian Neusoft University of Information, Dalian Liaoning

Received: Oct. 7th, 2024; accepted: Oct. 22nd, 2024; published: Jan. 3rd, 2025

Abstract

In the era of big data, all major financial institutions continue to make efforts on the road of digital transformation. In the face of the challenges of the mobile Internet, digital finance has intervened in all aspects of the business of financial institutions. In this paper, we use Python algorithm to mine and analyze the user comments of financial institutions' APPs. We analyze the collected 30,142 text data with LDA topic model to find out the users' needs and classify them into function, service and promotion categories. Then we use SnowNLP algorithm to understand the users' emotional tendency, finding that the APP user satisfaction is not high. As there is an urgent need to improve the

user experience, based on the quantitative analysis, optimization suggestions for financial institutions' APPs have been put forward to improve the overall operation of these APPs.

Keywords

Financial Institutions, APP Optimization, LDA Topic Model, Emotional Tendency

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

为了推进数字化转型，各大金融机构早已纷纷推出官方 APP 并不断升级换代，力图实现智能技术与金融业务场景的深度融合，从而为用户创造更加便捷的服务体验[1]。在各大银行不断推进数字赋能、服务创新的同时，APP 用户在使用中所产生的看法与需求值得被开发者和运营者重视。除了经典的调研与访谈方式，互联网所带来的海量用户生成内容(UGC)反映了用户真实的使用体验、兴趣与需求，分析应用市场中 APP 页面的用户评论可以帮助银行机构更好地了解用户反馈和需求，从而进行产品研发和改进，改善市场推广和提升用户满意度[2]。本文以此为切入点，运用大数据算法进行金融机构 APP 用户的需求与情感倾向分析，力图提升 APP 的使用体验。

2. 研究设计与数据来源

2.1. 研究方法与研究步骤

为研究用户对于金融机构 APP 的真实需求以及情感倾向，本文将采用文本数据分析方法，通过数据挖掘获取金融机构类 APP 的用户评论，并运用 LDA 和 SnowNLP 模型及算法进行数据分析。

文本分析法旨在从文本数据的挖掘、处理和分析中获取有价值的信息，从而帮助研究者深入理解文本数据中所隐藏的信息与趋势，提升决策的效率[3]。基于文本分析法的研究思路，本文将对收集到的大量文本数据展开 LDA 主题模型分析以找出用户的需求，以 SnowNLP 算法了解用户的情感倾向。

LDA (Latent Dirichlet Allocation)是一种无监督学习算法。LDA 主题挖掘在文本分析中已得到了广泛的应用，并在研究用户需求方面起到了重要的作用。例如，在旅游研究方面，已有多篇文献通过 LDA 模型分析 OTA 平台上的用户评论，了解各旅游目的地的感知形象[4]；在电子商务研究方面，LDA 模型可用于分析电商平台和相关 APP 中产生的用户评论[5]。而在金融研究方面，已有学者对于金融类 APP 的隐私条款进行文本分析[6][7]。本文将通过 LDA 模型确定金融机构类 APP 评论文本的主题分类，从而了解用户的真实需求。

随着人工智能和深度学习领域的飞速发展，自然语言处理(NLP)越发受到各行各业的重视，衍生出多种细分算法。其中，SnowNLP 针对于中文文本处理，能够获取中文文本数据中的情感特征并量化情感倾向，反映用户对于 APP 的满意度，应用较为广泛[8]。但模型内置的训练语料库为电商评论文本，因此需要根据研究内容进行预训练，以获得更为准确的情感倾向结果。因此，本文在使用 SnowNLP 算法前，先拆分获取的金融机构类 APP 文本数据进行了预训练。

2.2. 数据来源

为达到上述研究目的，需要对金融机构 APP 所产生的用户评论进行收集和整理。本文通过八爪鱼数

据采集器和七麦数据收集该类用户的公开评论。七麦数据是国内专业的移动应用数据分析平台，汇集了 IOS 和安卓主流应用商店的各项实时数据及排名，在该平台进行数据挖掘能够较为全面地获取大众用户对于金融机构 APP 的真实需求与看法。本文在 IOS、华为、小米等应用商店根据各金融机构 APP 的热度和下载量进行了评论采集，评论的时间跨度为 2022 年 2 月至 2024 年 2 月，在去除无效、广告和重复文本后，最终获取有效文本数据共 30,142 条，涵盖了多个互联网银行 APP 及其旗下信用卡 APP、非银行机构金融 APP 等。

3. 金融机构 APP 用户需求分析

3.1. 词频分析

在文本分析中，词频统计与分析必不可少，其中的高频词分析能够反映用户的普遍需求。本研究使用 Python 库中的 Jieba 分词工具，在去除副词、停用词、无效词和合并同义词后，对文本数据进行分词处理，并统计高频词，整理为表 1，然后绘制词云图如图 1 所示。由高频词可见，用户评论涵盖了 APP 使用的方方面面，对于优点和缺点都直言不讳。在高频词中，“服务”、“功能”、“界面”、“活动”、“客服”和“业务”体现了用户对于银行 APP 覆盖的场景、提供的功能和推广活动的重视与了解，而在 APP 的使用过程中，“密码”的忘记与找回则成为用户最为关注的问题，“体验”反应了用户对于整体使用流程的感受，而其中“客服”的水平则成为影响用户体验的重要环节。客服的优异表现结合功能与界面的升级会让用户感到 APP 整体服务的人性化，而无法解决问题的客服则更容易引起用户的不满，影响 APP 的使用体验。

Table 1. High-frequency words statistics
表 1. 高频词统计

高频词	频数
密码	2705
垃圾	2142
服务	1654
功能	1619
界面	956
活动	881
体验	828
客服	807
版本	700
验证码	676
办理	644
人脸识别	626
业务	608
信息	594
用户	566
人性化	551
限额	525
页面	524
短信	489
柜台	488



Figure 1. Overall word cloud distribution

图 1. 总体词云分布

3.2. LDA 主题挖掘与可视化分析

在 Python 中调用 LDA 模型进行主题挖掘，提取 1000 个特征词并迭代 1000 次，获取不同主题分类的结果。由于 LDA 模型为非监督学习，需要通过可视化辅助确定最优主题个数，即 k 的取值。运用 pyLDAvis 对生成的不同主题个数进行多次可视化模拟，结果如图 2 所示，气泡代表聚类生成的主题，气泡的大小代表主题的重要性，气泡的距离代表各主题间的差异性，结合困惑度，当 $k=9$ ，即主题数量为 9 时，各主题重叠较小，差异较大，能够有效地概括总体。由此，将金融机构 APP 的需求主题概括为 9 种类型。

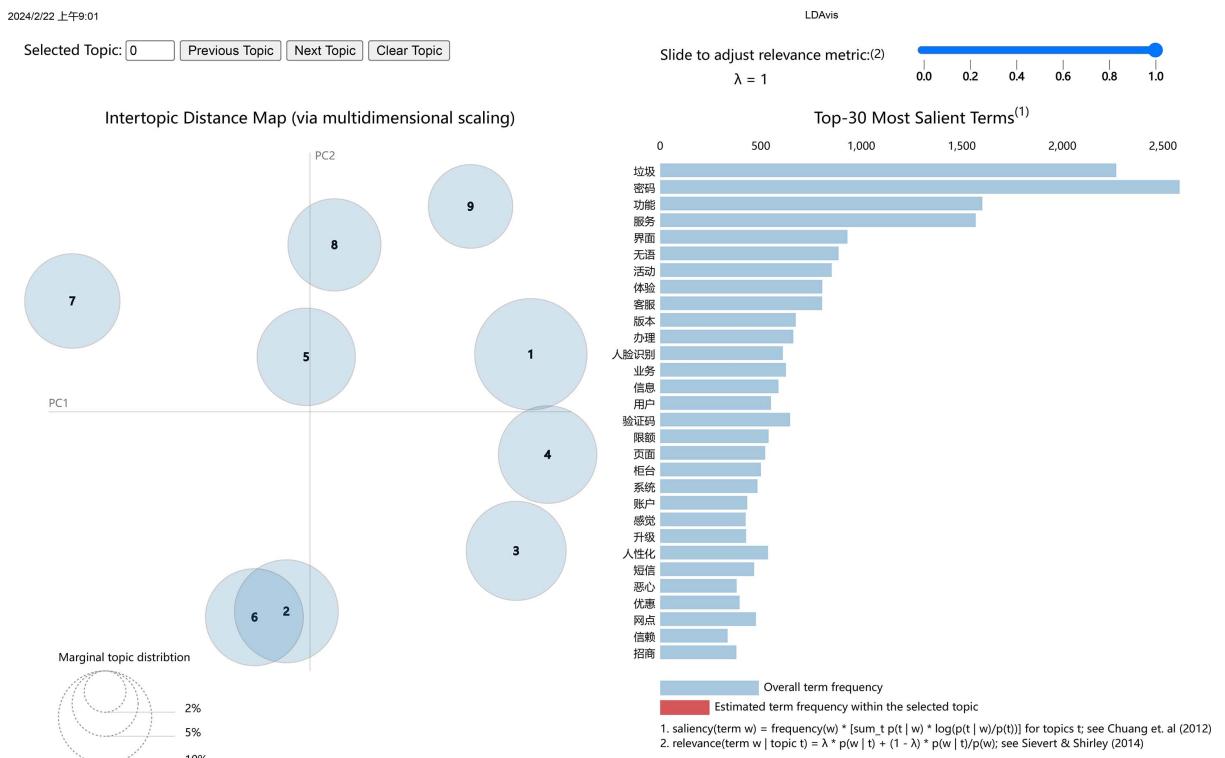


Figure 2. Visualisation results based on LDA model

Figure 2: Visualisation results based on LDA topic model.

主题一涉及的词汇主要为“密码”、“验证码”等，与身份验证相关，因此概括为“身份验证”。主题二的词汇主要为“客服”、“人工”和“页面”等，以提供基础服务的人工客服为主，因此将主题二概括为“基础客服”。主题三的词汇主要为“办理”、“业务”和“系统”等，涉及各项基础业务的办理与线下网点的查询，因此概括为“业务流程”。主题四的词汇主要为“活动”、“优惠”和“生活”等，与银行的推广活动相关，表现了用户对于活动和优惠信息的关注，因此概括为“活动参与”。主题五的词汇主要为“界面”、“体验”和“人脸识别”等，反映了用户对于界面设计与相关功能的体验，因此概括为“UI 界面”。主题六中的负面形容词体现了用户对于该类别涉及服务的严重不满，其他主要词汇为“账户”、“资金”和“余额”等，主要与基础的账户服务相关，因此概括为“基础服务”。主题七的词汇主要为“信赖”、“公司”和“信誉”等，体现了用户对于银行机构的信任程度，因此概括为“机构信誉”。主题八的词汇主要为“功能”、“信息”、“建议”、“空间”和“客户端”等，与 APP 在应用商店的运营较为相关，因此概括为“运营管理”。主题九的词汇主要为“服务”、“客户”、“理财产品”、“收益”和“基金”等，因此概括为“理财服务”。

综上，从 LDA 主题模型的聚类结果来看，主题一和主题五主要与 APP 的使用功能相关，主题四、主题七和主题八主要与 APP 的宣传和推广相关，主题二、主题三、主题六和主题九则主要与金融服务相关，因此将用户对于金融机构 APP 的主要需求划分为功能类、推广类和服务类三大主题，如表 2 所示。

Table 2. Topic categories and feature words based on LDA model
表 2. LDA 模型主题类别及其关键词

需求类别	主题	关键词
功能类	身份验证	密码、验证码、短信、网络、设置、重置、程序、流量、图表、指纹、登录密码、浪费
	UI 界面	界面、体验、人脸识别、用户、人性化、办法、方式、习惯、美观、照片、工具、直观
推广类	活动参与	活动、版本、优惠、生活、设计、优化、时间、人性化、交易、学校、实惠、只能、优惠券、力度
	机构信誉	垃圾、信赖、地方、公司、服务器、影响、纪念币、服务质量、外包、团队、效率领导、信誉
运营管理	运营管理	功能、信息、建议、账单、消费、五星、身份、鸡肋、客户端、空间、充值、收支、视频、用心
	基础客服	客服、页面、省级、招商、电话、人工、设备、指纹、速度、平台、后台、权限、重新安装、定位、解决问题
服务类	业务流程	办理、业务、限额、柜台、系统、网点、线下、营业厅、工作人员、交易、流程、原因、工作、时间、排队
	基础服务	账户、服务态度、无语、自动、资金、余额、管理、关键、办卡、电子、借记卡、评分、收费
理财服务	理财服务	服务、客户、理财产品、收益、基金、经理、内容、热情、态度、手续费、专业、速度快、投资、利率

3.3. 情感分析

首先，为获取更有针对性的研究结果，对 SnowNLP 模型进行训练。在获取的文本数据中抽取 1200 条进行人工情感分类，然后将所得的积极文本和消极文本迭代入 SnowNLP 模型完成训练，随机抽取测试

数据进行模型检测，发现该模型的准确率达到了 85% 以上，适合用作本研究的情感倾向分析。

接下来正式运用 SnowNLP 模型进行情感分析。SnowNLP 模型计算得出的情感倾向数值范围为 0~1，数值越大情感倾向越正向，反之，数值越低情感倾向则越负向。以 0.5 为区分，数值大于 0.5 为正向情感，低于 0.5 则为负向。金融机构 APP 所产生的评论文本总体情感倾向得分为 0.3972，表明用户对于该类 APP 具有较大的负面情感，满意度不高，抱怨较大，急需改善用户体验。

由于情感倾向能够反映用户对于 APP 的满意度，对于相关机构而言，积极评论能够体现 APP 的优势，而了解消极评论中隐含的用户抱怨，有助于进一步优化 APP，提升数字化推进的质量。为此，本研究对积极和消极评论分别进行了高频词统计，以了解目前用户的真实体验。由表 3 和图 3、图 4 所示，积极评论与消极评论的高频词存在相似之处，例如“服务”、“功能”和“办理”均出现在前十五名之内，这说明服务与功能是金融机构 APP 的核心，优质的服务与人性化的功能设计将带来良好的用户体验，极大地提升用户的评价和满意度，使用户对于 APP 更加信任；反之，消极的体验将对用户产生更大的负面影响，更易产生用户抱怨，降低 APP 的满意度与口碑。

Table 3. High frequency words statistics for positive and negative comments
表 3. 积极评论与消极评论高频词统计

积极评论高频词	频数	消极评论高频词	频数
服务	930	密码	2609
活动	628	垃圾	2134
功能	584	功能	1035
界面	538	客服	750
优惠	362	服务	724
感觉	275	人脸识别	609
人性化	269	版本	606
生活	256	体验	582
招商	251	验证码	527
信赖	250	信息	524
体验	246	办理	477
业务	185	限额	477
额度	181	用户	453
办理	167	柜台	449
设计	152	网络	433

4. 金融机构 APP 优化策略与建议

4.1. 优化 APP 的用户使用体验

从用户对于 APP 的功能类需求可见，简化各项流程、方便用户操作是提升用户整体使用体验的核心 [9]。其中，身份验证是用户提及最多的功能类需求，与之相关的用户注册、登录和密码找回是几乎所有用户都要进行的操作流程，应引起产品经理的重视。除此之外，应设计简洁直观的 APP 界面，方便用户快速找到需要的功能和信息。具体而言，可设计便捷的导航栏和快速入口，让用户能够快速访问常用功

能；通过提供快捷的搜索功能，让用户能够快速找到需要的信息和服务。必要时可设置清晰简明的新手指引，帮助新用户迅速了解 APP 的各项功能入口，减少用户的学习成本。



Figure 3. Word cloud distribution of words with positive tendency

图 3. 积极评论词云图



Figure 4. Word cloud distribution of words with negative tendency

图4. 消极评论词云图

UI 的美观度也是影响用户满意的重要因素[10]。界面设计应简洁清晰，突出重点功能，避免过多的装饰和复杂的布局，并使用统一的风格和色彩搭配，保持整体的视觉统一性，提升用户体验。

4.2. 提升数智化服务质量

服务质量关乎于银行的满意度与口碑，而提升 APP 的数智化服务质量，是金融机构数字化转型和智能化发展的必经之路，既能够满足用户日益增长的金融服务需求，也有助于提升银行的服务水平和品牌价值[11]。

为提升数智化服务质量，首先，银行 APP 需合理进行用户数据挖掘，建立完善的数据采集系统，运用机器学习等大数据分析技术，对用户数据进行深度分析，了解用户需求和行为规律，为产品与功能的设计提供支持，从而实现更精准的个性化推荐服务，满足用户的多元化需求。其次，可引入智能客服系

统，通过自然语言处理和人工智能技术，提供智能化的客户服务支持作为人工客服的补充，并及时更新技术，丰富线上服务的场景，提升智能客服的交互能力，在便捷用户的同时降低银行的服务成本；还可开发自助服务功能，如智能问答系统、在线办理业务等，提升用户体验和服务效率。

与此同时，也应注意到目前的人工智能还不足以取代人工客服，因此对于人工客服的培训和管理仍应引起重视，确保用户能够及时获取人性化的服务，提升对于 APP 的服务体验。

4.3. 建立数字化的运营推广体系

建立数字化的运营推广体系对于银行 APP 的推广和用户获取至关重要[12]。由前文分析可见，活跃的用户参与和全方位的运营推广能够在获取新用户的同时，提升用户对于 APP 的关注，并从活动参与中获得良好的体验，产生对于 APP 和银行机构的信任感，从而转化为忠诚客户。为建立数字化的运营推广体系，首先需要对 APP 市场进行深入分析，了解目标用户的需求和偏好，针对不同用户群体推送个性化的推广活动和内容[13]。在获取用户方面，可增加流量入口，在各大互联网平台投放链接广告，通过有趣、新奇的活动吸引新用户的关注和下载；或通过优化 APP 在各大应用商店的排名，制定有效的 SEO 策略，提高 APP 的曝光度和搜索可见性，吸引更多用户了解和使用银行 APP。在维系用户方面，全面利用社交媒体平台，发布有吸引力的内容与活动，如省钱攻略等，保持与用户的互动，提升用户的黏性。建立用户口碑和社区管理机制，关注用户评价和分享，及时收集用户的反馈，进行数据分析和评估，并适时调整和优化数智化服务策略，从而提升银行 APP 的口碑和信誉。

4.4. 提升数据安全和隐私保护力度

随着数字化技术的发展，社会对于数据安全和隐私保护更为关注，而银行 APP 有责任维护用户的相关权益[14]。为此，银行 APP 应提升数据加密和安全传输技术，保护用户数据在传输和存储过程中的安全性，防止数据被恶意攻击者窃取或篡改；采用多因素身份认证技术，确保用户身份的真实性和安全性，限制未授权用户的访问权限，防止数据被盗用；严格遵守相关隐私保护法规，明确用户数据收集和使用目的，保护用户的隐私信息不被滥用或泄露。与此同时，也应设立相关员工的安全意识培训，强化对数据安全和隐私保护的重视，提升员工的安全意识和防范能力。

5. 结语

数字化转型的推进让各类官方 APP 成为金融机构与客户沟通的窗口，APP 的使用体验与服务的整体质量密不可分，值得引起各大金融机构的重视[15]。通过大数据挖掘，本文对于获取的 APP 评论文本数据进行了 LDA 主题构建，将金融机构 APP 的用户需求分为功能类、服务类和推广类，并细化了每一个类别内的需求主题。然而，尽管目前金融机构 APP 的应用广泛，但用户评价却不高。SnowNLP 情感分析发现，用户在 APP 的使用中面临着诸多问题，体验欠佳。因此，各机构应当进一步优化 APP 使用体验，不断提升数智化服务质量，并在大力推广 APP 的同时提升数据安全和隐私保护的力度，从而真正贴近用户，为用户创造价值。

参考文献

- [1] 黄海聪. 我国银行数字化转型的研究热点与演进趋势——基于可视化知识图谱分析[J]. 时代经贸, 2023, 20(11): 64-67.
- [2] 楚东晓, 王雯露, 穆勤远. 基于 LDA 和语义网络的产品感知价值维度研究[J]. 包装工程, 2023, 44(S1): 47-55.
- [3] 李亚超, 罗昆, 王玉法. 上市公司财务报告问询函与年报前瞻性信息披露——基于文本分析的研究方法[J]. 南方金融, 2024(1): 68-82.
- [4] 邹永广, 雷振仙, 廖金斤, 等. 新媒体时代旅游资源持续关注损失测度及影响因素——基于网络舆情大数据的

- 证据[J]. 自然资源学报, 2024, 39(7): 1512-1530.
- [5] 魏丽. 基于 LDA 主题模型的电商评论数据分析[J]. 信阳农林学院学报, 2023, 33(3): 112-116.
- [6] 马永保. 隐私政策视角下第三方支付行业用户个人信息保护[J]. 金融科技时代, 2022, 30(4): 67-73.
- [7] 徐昌龙, 夏火松. 金融科技 APP 隐私政策对企业绩效作用机制[J]. 武汉纺织大学学报, 2024, 37(2): 91-99.
- [8] 蔡增玉, 韩洋, 张建伟, 等. 基于 SnowNLP 的微博网络舆情分析系统[J]. 科学技术与工程, 2024, 24(13): 5457-5464.
- [9] 郑亚萍. 商业银行移动金融发展策略[J]. 合作经济与科技, 2024(11): 60-62.
- [10] 周雷, 邱勋, 刘欣, 等. 数字经济时代金融科技 App 用户满意度与忠诚度研究[J]. 价格理论与实践, 2023(4): 175-179.
- [11] 周雷, 王慧聪, 毛晓飞, 等. 金融科技背景下开放银行构建模式与发展路径研究[J]. 新金融, 2021(12): 21-25.
- [12] 交通银行深圳分行课题组, 李灿宙. 金融机构线上线下一体化经营模式探索[J]. 新金融, 2021(1): 56-58.
- [13] 陆岷峰. 商业银行金融市场业务数字化转型路径研究[J]. 金融与经济, 2023(8): 89-96.
- [14] 黄莉群, 官心果, 钟宇. 数字经济时代的数据安全研究——以金融行业为例[J]. 商业经济, 2024(2): 174-179.
- [15] 张凤娜. 信用管理视角下商业银行数字化转型演化与挑战研究[J]. 工信财经科技, 2024(2): 18-32.