

面向数字人民币用户体验的微博情感分析

陆振昇

(湘潭大学公共管理学院, 湖南 湘潭 411105)

摘要: 目前数字人民币仍然处于试点应用阶段, 为了改进数字人民币, 促进数字人民币全国推广应用, 通过分析消费者民众对数字人民币的使用体验感受, 采集微博平台上关于数字人民币的评论文本, 经过数据筛选处理后, 使用改进的文本主题情感分析模型进行主题分类和情感分析。研究结果表明, 根据对主题情感分析的结果进行统计分析, 发现了数字人民币试点现存的不足和痛点, 从应用场景、数币平台、优惠活动 3 个方面提出具体策略建议, 为有关部门推广数字人民币全国应用提供了参考依据, 为研究数字人民币提供了新的思路和技术方法。

关键词: 数字人民币; 用户体验; 情感分析

中图分类号: F49; TP391.1 **文献标志码:** A **文章编号:** 1671-1807(2024)08-0117-08

随着区块链技术的不断发展和应用, 比特币、以太坊等虚拟货币的横空出世对世界各国的货币系统金融系统造成了冲击。世界多国央行因此开始了本国的法定数字货币的开发以应对虚拟数字货币所带来的不利影响, 中国人民银行也加入这一行列^[1]。中国人民银行于 2014 年成立了专门的团队(中国人民银行数字货币研究所)来进行数字人民币的研发工作^[2]。自 2019 年底开始, 数字人民币(e-CNY)相继在一些城市以及北京冬奥会整个场景内开启试点测试^[3]。

2021 年 3 月 11 日全国人大正式表决通过了《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》(以下简称“十四五”规划), 在“十四五”规划的第五篇“加快数字化发展 建设数字中国”提出, 要加快数字经济和数字社会的建设, 促进数字技术与实体经济深度融合, 推动购物消费、交通出行等各类场景数字化, 构建全民数字畅享的新生活, 以数字化转型生产方式、生活方式的变革^[4]。在 2023 年 2 月 27 日, 中共中央、国务院印发了《数字中国建设整体布局规划》, 再次强调要促进数字经济与实体经济深度融合, 以数字化驱动生产生活方式变革, 同时指出, 要做大做强做优数字经济, 培育壮大数字经济核心产业, 全面赋能经济社会发展^[5]。2023 年 4 月 4 日, 中国人民银行召开了货币金银和安全保卫工作电视会议, 要求深入贯彻党的二十大和中央经济工作会议

精神, 强调进一步提高数字人民币研发试点攻坚能力, 稳妥推进数字人民币研发试点。2023 年 9 月 3 日, 中国人民银行数字货币研究所所长在 2023 中国国际金融年度论坛上表示, 要实现用数字人民币作为所有零售场景下的支付工具, 将数字人民币升级成为适应数字经济发展的通用型支付工具。

截至 2023 年下半年, 数字人民币的试点范围扩展为 17 个省份的 26 个地区, 这个数字相对于全国范围来说还是比较小的, 且即便在试点城市中数字人民币的使用热度也并不高。因此, 提高民众对数字人民币的使用意愿和长期使用意愿是数字人民币推广工作的重中之重。在当今互联网社交媒体十分发达的时代, 广大网民往往喜欢在社交媒体平台上发表对于某些事物、事情的看法、日常生活的碎碎念等信息。微博作为当下国内热门的社交媒体平台, 其 2023 年第三季度报告显示微博的月活跃用户数达 6.05 亿人。因此, 可以选取微博平台作为数据来源, 从用户体验角度出发进行数字人民币的应用分析。

目前国内学者关于数字人民币的研究主要侧重于金融学视角、宏观政策和应用场景等。例如, 王劲屹^[6]进行了数字人民币在跨境支付领域的应用情景分析; 寿松涛等^[7]根据现实情况从宏观层面提出了建设粤港澳大湾区数字人民币创新区的路径建议; 杨荣海和李亚波^[8]通过对比世界各国央行数字货币发展情况, 为我国数字人民币的发

收稿日期: 2024-01-24

作者简介: 陆振昇(1996—), 男, 湖南长沙人, 硕士研究生, 研究方向为信息分析与评价、网络舆情分析。

展提出了策略建议;刘凯等^[9]使用经济学计量的方法进行了数字人民币发行和数字人民币支付的宏观政策研究;缪洋和刘莎莎^[10]从商业银行视角进行了数字人民币应用场景的理论探究;宋鹭和李佳林^[11]采用“货币之花”概念模型探究了数字人民币跨境支付应用的未来发展;江映霞^[12]以广州市为例探讨分析了数字人民币试点的现状,为数字人民币推广提供了相应参考建议。这些研究集中在数字人民币的理论分析、制度设计和应用范围推广的政策建议等宏观层面的分析,主要是从经济学角度出发。

同时,也有学者开展了基于公众视角的数字人民币接受意愿及影响因素研究。例如,林宸宇等^[13]在福州市进行了数字人民币使用意愿的影响因素及内在作用机制的研究,发现提高宣传力度可以让民众更多地了解数字人民币、使用数字人民币;施嘉文等^[14]使用问卷调查法对浙江省数字人民币试点城市民众进行调研,分析接受度影响因素,并提出建议;谢英凯^[15]在北京市进行了群众问卷调查,构建数字人民币公众满意度评估模型,进行结果分析;Wu等^[16]采用结构方程模型,解析影响数字人民币接受意愿的主要因素。以上研究均通过问卷调查获取数据,并都采用统计评价模型进行检验,虽解析出了一些影响民众使用数字人民币的因素,但调查人群既有未使用过数字人民币的群众也有使用过数字人民币的群众,覆盖范围也未达到全国,且没有为数字人民币推广提供更多、更深层次的参考建议。这些研究集中于消费者群众对于数字人民币的接受使用意愿,缺乏对数字人民币用户使用体验的研究。

通过微博文本评论进行公众舆情分析、情感分析已经是近些年国内学者在文本情感分析领域应用研究比较多的研究范式。例如,刘桂海等^[17]使用文本特征分析方法和情感分类方法对假房源微博评论信息挖掘,分析民众对于假房源的关注点和情感倾向的影响因素;林伟和陈雁^[18]通过构建融合注意力机制和 Albert-BiGRU 的模型,以及 Bert-BiGRU 和多尺度卷积神经网络(convolutional neural networks, CNN)的模型,实现了对公开微博评论文本数据集的情感分析,取得不错的效果;吴旭旭等^[19]采用 Roberta 预训练模型与 BiLSTM 神经网络相结合,编码及提取微博文本语句特征,构建融合模型在公开微博文本数据集上取得很好的效果;赖宇斌等^[20]使用 Roberta 预训练语言模型进行文

本解析,融合提示词嵌入作为情感分类效果提升的技巧,此方法在突发公共卫生事件微博数据集上得到验证,拥有较好的性能表现;曾莉等^[21]通过构建 LDA-Attention-BiLSTM 的融合模型对某单位热点舆情发生时的微博评论文本进行舆情主题和情感演化研究。可以发现关于微博情感分析的研究方法主要是运用深度学习技术。

1 RobertaWE-LDA-SAPUE 模型构建

1.1 用户体验

用户体验(user experience, UE)的概念是由美国认知心理学家、用户体验设计师 Norman 等^[22]在 20 世纪 90 年代中期提出,是一个涉及心理学、社会学、美学、人体工程学和计算机科学等多个领域的交叉学科概念。目前,国际标准化组织(international organization for standardization, ISO)关于用户体验的定义在学术界和工业界认可度比较高^[23], ISO9241-210 标准将 UE 定义为“用户在使用或预计要使用某产品、系统或服务时产生的主观感受和反应”。此外,ISO 对该定义进行了补充说明,UE 包含使用前、使用时和使用后所产生的情感、信仰、喜好、认知印象、生理和心理上的反应、行为和后果等。

1.2 LDA 主题模型

LDA(latent dirichlet allocation)主题模型由 David Blei 于 2003 年提出^[24],是一种文档主题生成模型,它包含了三层结构,分别是主题、文档、词,是一个贝叶斯概率模型。LDA 模型是一个无监督的机器学习方法,可以用来识别大规模文档集或语料集中的潜在主题信息^[25]。同时,LDA 采用了词袋模型,通过将每一篇文档视为一个词频向量,文档直接用这些向量集合来表示,并且这个词袋方法没有考虑词与词之间的顺序,降低了计算的复杂度。在 LDA 模型中每一篇文档代表一些主题所构成的概率分布,在每一个主题中主题又代表了由很多单词构成的一个概率分布^[26]。LDA 模型的核心是 Dirichlet 分布,在贝叶斯概率理论中被称为共轭先验分布^[27]。

LDA 模型的大体思想为:运用先验分布的理念(即先设定一个猜想值去计算)通过不断迭代调整每个文档中每个词汇对应主题的概率分布和每个主题对应文档的概率分布,使最终结果符合实际的文档集中单词对应文档的分布。用数学公式表示为

$$P(wd) = P(\omega t) \times P(td) \quad (1)$$

式中: w 为表词汇; d 为文档; t 为主题。

1.3 Roberta-wwm-ext 模型

Roberta 模型^[28]是基于 Bert 模型^[29]改进后的更具有鲁棒性的 Bert 模型,其结构相比于 Bert 模型没有改变,由 12 个 Transformer 编码器组成,如图 1 所示。

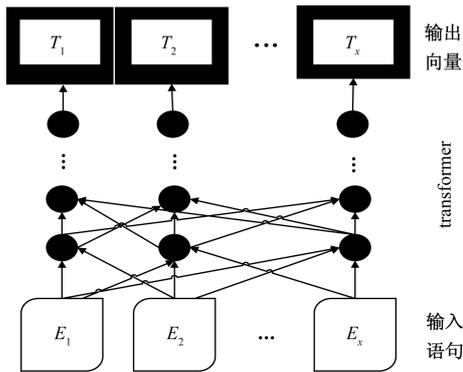


图 1 Bert 模型示意图

图 1 中的 Transformer 结构^[30]在 2017 年由谷歌 AI 研究院提出,是一种带有注意力机制的语言编码模型,被广泛应用于当前各种语言大模型的基础架构。Roberta 模型相较于 Bert 模型^[31]具体做了这些改进:①Roberta 模型采用动态 mask 操作,Bert 模型用的是静态 mask 操作。这样操作可以提升输入文本的随机性,从而使模型学习到更多的语义知识。②Roberta 取消了原 Bert 模型训练时的 NSP(next sentence predict)任务,原因是执行 NSP 任务需要从不同文档中提取信息,增加了判断的难度,取消之后明显模型效果得到提高。③预训练Roberta模型时用到了更大的批处理尺寸 batchsize 和更长的训练时间,同时扩充了预训练的语料规模。在训练时改进了原 Bert 模型的优化器参数,提升了模型训练的稳定性。④Roberta 模型使用了更大的 BPE(byte-pair encoding)词表来处理文本数据^[32]。

Roberta-wwm 模型^[33]在 Roberta 动态掩码操作上运用了全词掩码(whole word masking)的技术,主要是考虑了中文语言的特性^[34],比如说原来的掩码操作是把“跑步”一词中的“跑”或者“步”进行 mask,全词掩码则是“跑步”变成“mask”。Roberta-wwm-ext 模型^[35]则是在 Roberta-wwm 模型预训练过程中增加了一些扩展的数据集。

1.4 RobertaWE-LDA-SAPUE 模型

数字人民币试点工作截至 2023 年底开展了三年以上,尚未全国推广开来,并且数字人民币的研究现状和热点基本上围绕宏观上的路径建议。为了探究消费者群体在使用数字人民币后的体验感

受,通过从用户角度揭示数字人民币的现存优点和不足,从而为国家政府有关部门推广数字人民币应用、促进数字人民币支付体系建设等工作提供一定的参考和建议。

于是,构建了以 Roberta-wwm-ext 预训练语言模型和 LDA 主题分类模型相结合的文本情感分析模型,构造的面向用户体验的情感分析模型 RobertaWE-LDA-SAPUE(Roberta-wwm-ext-LDA sentiment analysis model from perspective of user experience)模型框架流程如图 2 所示。

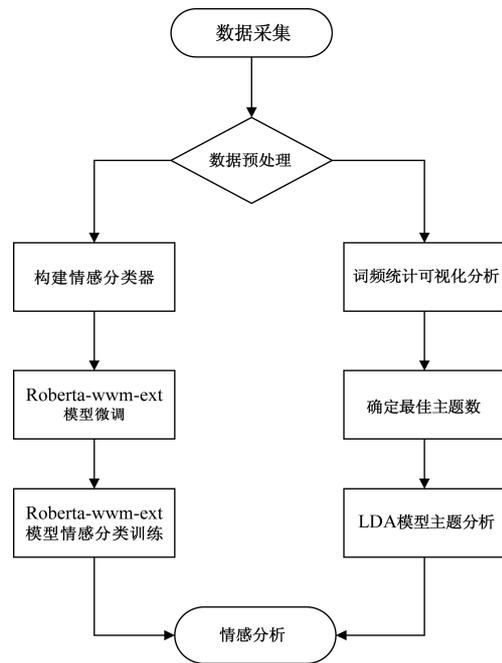


图 2 RobertaWE-LDA-SAPUE 模型流程

此模型分为两个阶段。

第一阶段(主题挖掘)步骤如下:

(1)数据预处理。对待处理文档集进行分词、去停用词等预处理。

(2)词频统计。利用 Python 环境中的数据分析模块及库和词云可视化进行词频分析。

(3)确定 LDA 主题数和主题分析。通过计算主题困惑度和词频分析结果相结合选取最佳主题数,随后得到主题-主题词分布。

第二阶段(情感分类模型训练)步骤如下:

(1)数据预处理。经过筛选和标签分类,得到符合比例的训练集测试集。

(2)构建预训练-微调范式模型的情感分类器。

(3)进行实验。设置最佳实验参数,得到微调后的模型。

结合两个阶段的实验成果来完成最终的主题情感

分析任务,从而实现对各个主题下的文本情感分析。

2 用户体验视角下的主题分析实证

2.1 数据预处理

使用爬虫 selenium 模块进行微博网页数据采集,爬取以“数字人民币”“数币支付”“数币交易”为关键词的所有原创文本和热门博文下的有关评论。本次采集的数据发布时间范围是 2023 年 1 月—11 月,经过筛选与“用户体验”相关且包含情感倾向信息的评论文本,共得到 11 370 条评论文本数据(包含时间、地点信息),设为文档集 L 。为了后续的主题分析研究能更精准地找到数字人民币用户在使用数字人民币后的方面级体验感受,对文档集 L 进行去停用词处理。通过加载 jieba 分词库和网络公开的中文停用词表,同时在停用词表中加入如“数字人民币”“数币”“体验”等对后续分析无意义的词语,最终得到文档集 L 。

2.2 词频统计分析

对文档集 L 的词频分析采用 collections 模块的 Counter 类,开发环境为 Python3.8。将文档集 L 导入 .txt 格式的文件中形成列表,经过 Counter 函数统计得到词频数排名前 10 的结果,如图 3 所示。

使用 wordcloud 库对上述词频统计表进行词云分析,得到如图 4 所示。

由图 3 和图 4 可知,在微博用户群体使用数字人民币的过程中更多关注的是其支付功能是否强大、支付过程是否便捷,应用场景多不多、支持的商家商户多不多,数字人民币 App 和数字人民币软硬件钱包这种数币载体的体验感觉怎么样,有没有优惠活动、满减红包这种可以薅羊毛的活动,活动多不

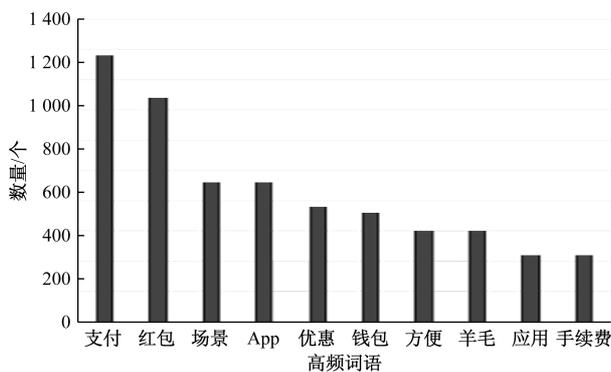


图 3 词频统计



图 4 词频词云图

多、力度大不大等方面。

2.3 确定主题数

由于 LDA 主题模型在训练时需要事先设定好主题分类的个数,困惑度的概念是一种用于评价语言模型的指标^[36]。使用主题困惑度作为确定最佳主题数的辅助指标,其在 LDA 模型中计算公式为

$$\text{perplexity}(D) = \exp \frac{-\sum p(d)}{\sum Nd} \quad (2)$$

$$p(d) = \sum \ln p(w) \quad (3)$$

$$p(w) = \sum z p(zd) p(wz) \quad (4)$$

式中:perplexity 为困惑度; D 为整个文档集; $p(w)$ 为测试集每一个词汇出现的概率; N 为测试集所有词集合; z 为训练过的主题; d 为测试集的每篇文档。

最终计算出来的困惑度代表文档主题的不确定性,因此理论上来说困惑度越小模型性能越好,则在困惑度曲线上显示为最低点或拐点处的主题数是最佳主题数^[37]。主题困惑度曲线如图 5 所示。

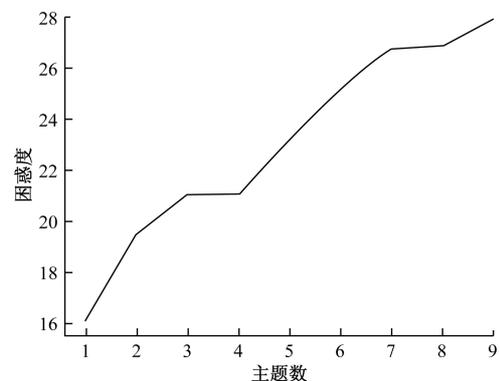


图 5 主题困惑度曲线

由图 5 可知,在主题数 3 或 4 时曲线趋势发生明显改变,由上升趋势变平。同时结合前面词频统计分析的结果,文档集 L 大致可分为四大类或者三大类,据此,最终选择主题数为 3 进行主题分析。

2.4 主题提取挖掘

基于 Python3.8 的 Sklearn 库中 LDA 模型包对文档集 L 进行主题划分,将超参数 α 设置为 0.1, β 设置为 0.01,最大迭代次数设置为 50 次,确定主题数为 3,得到主题-主题词分布如表 1 所示。

表 1 主题-主题词分布

主题	主题词(部分)
Topic 0	消费 商户 商家 地铁
Topic 1	钱包 软件 手续费 银行
Topic 2	优惠 红包 羊毛 活动

结合前文词频分析,将 Topic0、Topic1、Topic2 分别定义为“应用场景”“数币载体”“优惠活动”三大类。其中“应用场景”类有 2 820 条,“数币载体”类有 3 630 条,“优惠活动”类有 4 920 条。

3 面向用户体验的情感分析实证

3.1 数据集与预处理

通过爬虫 selenium 技术获取微博网页中以“数字人民币”“数币支付”“数币交易”为关键词搜索的用户文本数据,数据选取标准:①发布时间范围为 2021—2022 年底;②包含情感倾向信息;③属于用户体验范畴内。根据以上标准,共处理选取得到文本数据 28 425 条,采用三类情感标签进行标注,分别是积极乐观型(positive)、客观中立型(neutral)、负面消极型(negative)。样例如表 2 所示。

表 2 标注情感语句文本样例

情感标签	语句
负向(negative)	不如云闪付,云闪付可以直接扫所以收款码,包括但不限于微信、支付宝、银行卡的收款码
中性(neutral)	我尝试用数字人民币发红包了
正向(positive)	数字人民币钱包设计得很好看,在众多乱七八糟支付界面里是设计得最好的

将此数据集划分为 23 425 条训练集和 5 000 条测试集。

3.2 改进的情感分类器

Roberta-wwm-ext 模型沿用的是 Bert 结构,经过 12 层 Transformer 编码器得到的是 768 维词向量。通常,在自然语言处理领域是用 softmax 层作为分类器来实现情感分类任务^[38]。如图 6 所示。

因此,将 Roberta-wwm-ext 直接连接 softmax 的模型作为基准模型。在深度学习领域,全连接层可以作为一个分类器接在整个主干神经网络的最

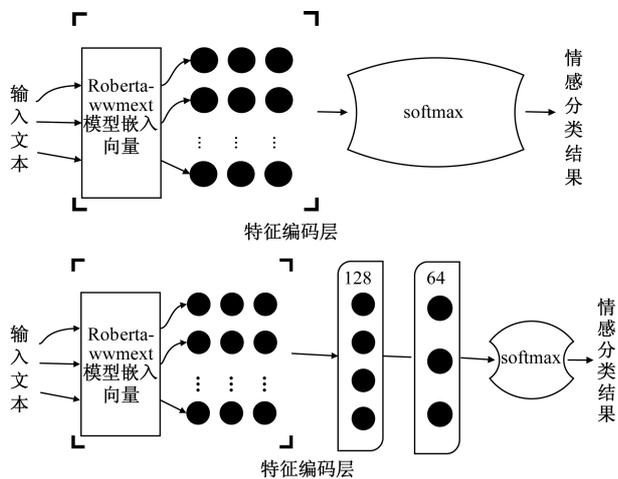


图 6 情感分类模型结构

后,将前层计算得到的特征向量空间映射到标记样本空间,简单说就是将特征融合成一个值,便于分类学习。于是,设计两层神经元个数分别为 128、64 的全连接神经网络(降维表示便于计算和提取有效信息)结合 softmax 函数作为情感分类器。softmax 计算公式为

$$\text{softmax}(x) = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (5)$$

式中:softmax 计算结果在 [0, 1] 的范围内,所有子项相加总和等于 1。以情感三分类问题为例,softmax 的输出会是一个 [0.86, 0.13, 0.01] 的概率分布(分别代表标签 1、标签 2、标签 3 的对应概率),选择最大的概率对应标签作为分类结果。

3.3 实验过程及结果

改进的文本情感分析模型实验环境如表 3 所示。

经过多次实验,选取的实验超参数设置如表 4 所示。

将实验的模型评价指标定为测试集准确率(Accuracy),即模型分类预测正确数量与分类预测任务总量的比例。计算公式为

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \quad (6)$$

式中:TP 为真实值和预测值都为正类的样本数;TN 为真实值和预测值都为负类的样本数;FN 为真实值为正类、预测值为负类的样本数;FP 为真实值为负类、预测值为正类的样本数。

表 3 实验环境

项目	说明
操作系统	64 位 Windows10
CPU	Intel i5-12490F
GPU	NVIDIA GeForce RTX 3060 8 GB
Python	3.8
Pytorch	2.1.0
内存	16 GB
Cuda Runtime Version	11.8
Cuda Driver Version	12.3

表 4 模型超参数设置

参数	说明	值/设置项
Seq_length	单个句子最大长度	128
batch size	一次训练选取的样本数	8
learning rate	学习率	0.000 01
optimizer	优化器	Adam
Epoch	迭代次数	1 000

以 LDA 模型主题分析的文档集 L 作为验证集,基准模型的实验结果为在测试集上准确率达到 80.6%,在验证集上准确率为 80.3%;带有改进情感分类器的模型实验结果为在测试集上准确率达到 82.2%,在验证集上准确率为 82.0%。可以发现,改进的情感分类器在改进的文本情感分析模型上的分类效果较基准模型略有提升。

3.4 结果分析及对策建议

通过对 2023 年 1 月—11 月的文档集 L 的情感分析结果进行统计分析,得到图 7。

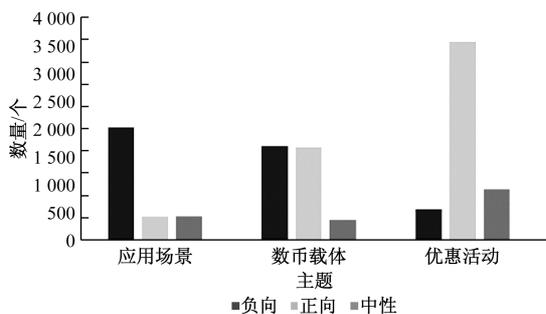


图7 三大主题分布下的情感态势

由图 7 可知,消费者用户对于目前数字人民币的可使用场景是非常不满意的,对于目前数字人民币钱包的使用体验的满意程度比较一般。相对来说,用户对于数字人民币目前可以领到的优惠券、红包还是比较满意的。

结合文档集 L 中在爬取过程所获得的时间信息,得到“应用场景”主题负向情感、“数币载体”主题负向情感、“优惠活动”主题正向情感的时间趋势,如图 8 所示。

由图 8 可知,a 类情感和 b 类情感走势先上升后下降,说明在数字人民币在可使用场景和 App 钱包支付体验方面令消费者用户感受从不满意到稍微满意,但不满意的态势略有回升;c 类情感走势整体趋于下降,但中间有反复,说明在推广数字人民币给到消费者用户的优惠力度方面,呈现逐步下滑的情况。

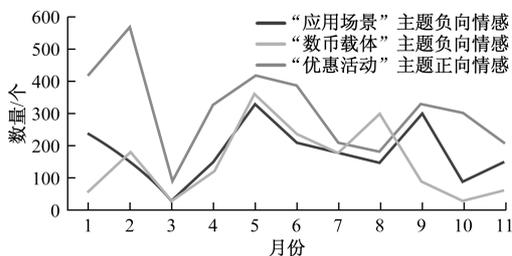


图8 三大主题情感趋势

根据时间标签信息,得到爬取到的数字人民币用户体验的数据量走势如图 9 所示。

由图 9 可知,微博上关于数字人民币的用户体验评论文本数量呈现先上升后下降的态势,因此推断,数字人民币可能存在用户流失或新用户增长缓慢的情况。

结合文档集 L 中在爬取过程中所获得的 IP 地址信息,得到数字人民币用户使用情况地区分布,如图 10 所示。

由图 10 可知,长三角苏浙沪、广东以及北京这类经济发达地区的数字人民币使用热度明显高于其他地区,长三角地区尤为突出。这其中可能与当地政府有关部门、银行等机构开发数字人民币使用场景、适用商户的进度有关,也可能和不同地区之间民众的消费能力水平、对新鲜事物的接受程度等因素有关。

综上,为推广数字人民币试点有关部门提出以下建议。

(1)加快数字人民币应用场景开发拓展。加速扩大数字人民币试点城市范围,利用各省会城市、经济发达城市的“虹吸效应”,将支付场景由省会到地级市到乡镇部署;加强数字人民币生态建设,打通各家银行、第三方支付平台、各类缴费平台等支付通道,推进“一码多扫”,同时,给予商家商户接受数字人民币支付的相应优惠减免优待政策。

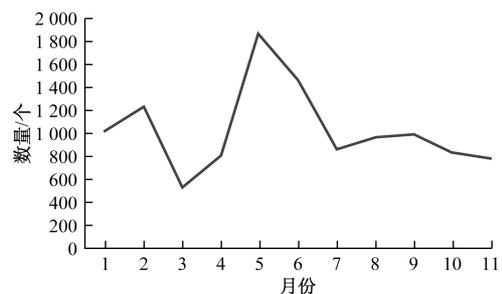


图9 数字人民币用户体验数据量走势

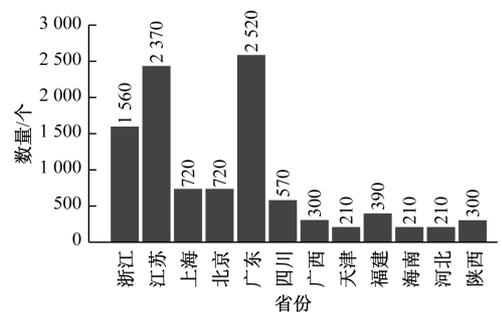


图10 数字人民币用户体验数据量地区分布

(2)继续优化开发数字人民币软硬钱包功能。加快新增支付功能的开发上线,全面升级线上线下终端服务,吸取微信、支付宝等热门支付软件的优点,简化钱包支付、钱包认证等流程,提高不同银行账户之间关联性,使货币流转更智能化。

(3)留住老用户,吸引新用户。长期开展优惠红包活动,不做大数据杀熟行为。持续加码数字红包、消费券的发放,开通更多能够领取到优惠的渠道,并简化领取红包消费券的操作流程。

(4)提高数字人民币宣传力度。内陆地区、欠发达地区的试点城市除了应加快应用场景建设布局,可以北上广深地区当地政府、银行学习数字人民币推广经验,借鉴到本地的数字人民币推广中来。同时,可以在公共场合,例如商场、步行街、旅游景点等进行数字人民币的广告宣传、科普教育等工作,加强银企、银政、银行高校之间的交流合作,制作一些简易的数字人民币开通使用教程,让消费者更多地了解数字人民币的好处。

4 结语

通过提出 RobertaWE-LDA-SAPUE 文本主题情感分析模型,为文本情感分析领域贡献了一种新的技术方法,成功运用到数字人民币的用户体验研究。给文本情感分析领域拓展了新的应用领域的同时,丰富了数字人民币的研究视角和研究方法,为相关机构和部门推进数字人民币试点的过程中提供参考。

随着深圳罗湖的数字人民币产业园成立、数字人民币跨境支付成功案例的出现,数字人民币新增功能、应用场景越来越多,今后会有更多的人加入数字人民币使用行列,可以持续关注研究消费者用户对于数字人民币的体验感受,从而为全面推广数字人民币建言献策。在国家战略的加持下,中国人民银行对数字人民币的试点推广下了很大的决心,相信在中央和地方政府的共同努力下,数字人民币定将在数字经济、数字中国建设中发挥巨大的作用。

参考文献

- [1] 张浩然. 数字人民币全面推广的思考与建议[J]. 新西部, 2022(11): 75-76.
- [2] 刘晓欣. 数字人民币的主要特征及影响分析[J]. 人民论坛, 2020(26): 86-91.
- [3] 王伟玲. 加快数字人民币有序发展的难点和着力点[J]. 经济纵横, 2023(2): 122-128.
- [4] 黄震. 积极稳妥推进数字人民币发展[J]. 中国党政干部论坛, 2021(5): 79-80.
- [5] 李亮亮, 邢云文. 数字经济赋能共同富裕: 逻辑理路、问题指向与实践进路[J]. 经济问题, 2024(1): 10-17.
- [6] 王劲屹. 法定数字货币在跨境支付领域应用前景探析[J]. 征信, 2023, 41(12): 50-59.
- [7] 寿松涛, 杨诗培, 熊巧. 粤港澳大湾区建立数字人民币创新区探究[J]. 对外经贸, 2024(1): 54-57.
- [8] 杨荣海, 李亚波. 全球央行数字货币竞争现状与数字人民币的发展策略[J]. 经济学家, 2023(5): 46-56.
- [9] 刘凯, 郭明旭, 李育. 数字人民币发行与数字支付发展的宏观经济影响研究[J]. 中国工业经济, 2023(3): 39-57.
- [10] 缪洋, 刘莎莎. 商业银行视角下数字人民币应用场景创新研究[J]. 新金融, 2023(7): 41-45.
- [11] 宋鹭, 李佳林. 央行数字货币、跨境支付应用与数字人民币的发展[J]. 新金融, 2023(10): 51-57.
- [12] 江映霞. 数字人民币试点推行的影响与政策建议——以广州市为例[J]. 商场现代化, 2023(17): 56-58.
- [13] 林宸宇, 洪思程, 林本喜. 数字人民币使用意愿的影响因素研究——基于福州试点调查数据的分析[J]. 中国商论, 2024(2): 17-22.
- [14] 施嘉文, 何圣杰, 黄晶莹, 等. 数字支付创新实践——基于 SEM 的数字人民币使用接受度及影响因素研究[J]. 金融科技时代, 2023, 31(9): 43-49.
- [15] 谢英凯. 数字人民币试点公众满意度调查[J]. 合作经济与科技, 2023(7): 70-73.
- [16] WU B, AN X, WANG C, et al. Extending UTAUT with national identity and fairness to understand user adoption of DCEP in China [J]. Scientific Reports, 2022, 12(1): 1-11.
- [17] 刘桂海, 崔福龙, 卢彩菡, 等. 公众对假房源的关注点和态度: 基于微博评论的文本挖掘研究[J]. 管理评论, 2023, 35(11): 153-165.
- [18] 林伟, 陈雁. 融合 BERT-BiGRU 和多尺度 CNN 的中文微博情感分析[J]. 中国电子科学研究院学报, 2023, 18(10): 939-945.
- [19] 吴旭旭, 陈鹏, 江欢. 基于多特征融合的微博细粒度情感分析[J]. 数据分析与知识发现, 2023, 7(12): 102-113.
- [20] 赖宇斌, 陈燕, 胡小春, 等. 基于提示嵌入的突发公共卫生事件微博文本情感分析[J]. 数据分析与知识发现, 2023, 7(11): 46-55.
- [21] 曾莉, 杨添宝, 周慧. 基于 LDA 与注意力机制 BiLSTM 的微博舆情分析模型[J]. 南京理工大学学报, 2022, 46(6): 742-748.
- [22] 邓胜利. 国外用户体验研究进展[J]. 图书情报工作, 2008(3): 43-45.
- [23] 孟猛, 朱庆华. 国内外信息系统用户体验研究综述[J]. 图书馆学研究, 2021(9): 9-19.
- [24] 范宇, 符红光, 文奕. 基于 LDA 模型的专利信息聚类技术[J]. 计算机应用, 2013, 33(S1): 87-89.
- [25] 王鹏, 高铨, 陈晓美. 基于 LDA 模型的文本聚类研究[J]. 情报科学, 2015, 33(1): 63-68.
- [26] 李湘东, 张娇, 袁满. 基于 LDA 模型科技期刊主题演

- 化研究[J]. 情报杂志, 2014, 33(7): 115-121.
- [27] 孙宁宁. 基于主题模型的专利文本分析及应用研究[D]. 北京: 北京工业大学, 2017.
- [28] LIU Y, OTT M, GOYAL N, et al. RoBERTa: a robustly optimized BERT pretraining approach[EB/OL]. (2019-07-26)[2022-01-15]. <https://doi.org/10.48550/arXiv.1907.11692>.
- [29] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg: ACL Press, 2019: 4171-4186.
- [30] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing System. Red Hook: Curran Associates Inc., 2017: 6000-6010.
- [31] 马圆圆, 禹龙, 田生伟, 等. 融合 RoBERTa 和注意力机制的隐喻方面级情感分析[J]. 小型微型计算机系统, 2023, 44(10): 2236-2241.
- [32] 喻金平, 朱伟锋, 廖列法. 基于 RoBERTa-wwm-BiLSTM-CRF 的扶持政策文本实体识别研究[J]. 计算机工程与科学, 2023, 45(8): 1498-1507.
- [33] 王曙燕, 原柯. 基于 RoBERTa-WWM 的大学生论坛情感分析模型[J]. 计算机工程, 2022, 48(8): 292-298.
- [34] XU Z. RoBERTa-WWM-ext fine-tuning for Chinese text classification[EB/OL]. [2021-06-08]. <https://arxiv.org/abs/2103.00492>.
- [35] CUI Y M, CHE W X, LIU T, et al. Pre-training with whole word masking for Chinese BERT[EB/OL]. [2021-06-08]. <https://arxiv.org/abs/1906.08101v3>.
- [36] 谭春辉, 熊梦媛. 基于 LDA 模型的国内外数据挖掘研究热点主题演化对比分析[J]. 情报科学, 2021, 39(4): 174-185.
- [37] 王国睿, 张亚飞, 尚有为, 等. 基于 LDA 主题模型的电子病历热点主题发现[J]. 中华医学图书情报杂志, 2021, 30(2): 33-39.
- [38] 郑淞尹, 王萍, 丁恒, 等. 基于方面级情感分析的博物馆数字化服务用户体验研究[J]. 情报科学, 2022, 40(4): 171-178.

Weibo Sentiment Analysis for e-CNY User Experience

LU Zhensheng

(School of Public Administration, Xiangtan University, Xiangtan 411105, Hunan, China)

Abstract: The e-CNY is still in the pilot application stage, and in order to improve the e-CNY and promote the national popularization and application of the e-CNY. By analyzing the consumer public's feelings about the user experience of e-CNY, the text of comments about e-CNY on the Weibo platform was collected, and after data screening and processing, the RobertaWE-LDA-SAPUE model was used for topic classification and sentiment analysis. The results of the study show that, based on the statistical analysis of the results of the thematic sentiment analysis, the existing deficiencies and pain points of the e-CNY pilot were identified. It puts forward specific strategy suggestions from 3 aspects, namely application scenarios, digital currency platforms, and preferential activities, which provides a reference basis for the relevant departments to promote the national application of e-CNY, and provides new ideas and technical methods for the study of e-CNY.

Keywords: e-CNY; user experience; sentiment analysis