

社交媒体文本的价值测度

——基于深度学习的网络消费舆情指数构建及其作用研究

钟一珂, 廖灵辰

杭州电子科技大学经济学院, 浙江 杭州

收稿日期: 2025年2月8日; 录用日期: 2025年2月28日; 发布日期: 2025年3月31日

摘 要

随着互联网和社交媒体的快速普及, 社交媒体文本因其实时性、互动性和无偏性而受到关注。挖掘量化社媒消费舆情, 并进一步研究其对消费者信心的影响成为有意义的探索。文章首先构建了6个不同主题维度的消费热点关键词库, 爬取了社交平台约178.2万相关热门微博。进一步地, 基于Bert预训练模型挖掘情感倾向, 并引入微博热度得分和小波变换, 创建了一个用来测度消费舆情的网络消费舆情指数。研究发现, 网络消费舆情指数对消费者信心具有显著正向影响且两者存在长期稳定的均衡关系; 进一步将其加入消费者信心预测模型, 所有评价指标均得到显著提升, 且LSTM在不同预测期限内的预测效果和鲁棒性都优于GRU和Random Forest。

关键词

舆情挖掘, 网络消费舆情指数, Bert, 小波变换, LSTM

Value Measurement of Social Media Texts

—Research on the Construction and Function of Online Consumer Public Opinion Index Based on Deep Learning

Yike Zhong, Lingchen Liao

School of Economics, Hangzhou Dianzi University, Hangzhou Zhejiang

Received: Feb. 8th, 2025; accepted: Feb. 28th, 2025; published: Mar. 31st, 2025

Abstract

With the rapid popularization of the Internet and social media, social media texts have attracted attention because of their real-time, interactive and unbiased nature. It has become a meaningful exploration to explore the quantitative public opinion of social media consumption and further study its impact on consumer confidence. The paper first constructed a consumer keyword database with

6 different topic dimensions, and crawled about 1.782 million related popular microblogs on the social platform. Further, based on Bert pre-training model, emotional tendency is mined, and an online Consumer Confidence Index is created to comprehensively quantify consumer confidence by introducing Weibo heat score and wavelet transform. It is found that the online CPOI has a significant positive impact on consumer confidence and there is a long-term stable equilibrium relationship between the two. When it is further added into the consumer confidence forecasting model, all evaluation indicators are significantly improved, and the forecasting effect and robustness of LSTM are better than that of GRU and Random Forest in different forecasting periods.

Keywords

Public Opinion Mining, Online Consumer Public Opinion Index, Bert, Wavelet Transform, LSTM

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

作为推动经济增长的三驾马车之一, 2023 年消费支出对中国经济增长的贡献达到 82.5%, 成为最主要的驱动力。而随着消费者行为理论的发展, 越来越多的研究表明, 消费者对未来经济形势和消费水平的预期会显著影响其消费和投资决策[1] [2], 因此, 消费者信心成为国民经济形势和消费支出的风向标。消费者信心指数(Consumer Confidence Index, CCI)是消费者预期的主要衡量指标, 综合反映消费者对经济形势、生活状况、就业、物价、房市等方面的预期, 但在调查结束后一个月发布, 存在滞后性。

另一方面, 随着互联网的发展, 社交媒体已成为人们讨论公共事务和形成舆论的核心平台, 是舆情最大的数据的重要来源[3] [4]。2024 年, 全球超过 50 亿人使用社交媒体, 日均投入 151 分钟[5]。社交媒体文本具有实时性、互动性和无偏性, 在塑造公众预期方面具备较传统媒体更高的影响力[3] [6]。诺贝尔经济学奖得主 Shiller 提出的叙事经济学认为, 公众通过“叙事”形成预期, 进而影响经济决策[7], 而社交媒体舆情正扮演着这种“叙事”的角色。

因此, 网络舆情对消费者信心的影响不容忽视, 当消费者接收到有关政策利好或经济前景乐观的消息时, 其信心可能会得到提升。为测度网络消费舆情的价值, 本文首先基于新华网构建 6 个主题维度的消费热点关键词库, 爬取约 178.2 万热门微博, 并运用微调后的 Bert 情感模型挖掘情感倾向。进一步地, 结合微博热度和小波变换对消费舆情进行特征提取和量化, 构建网络消费舆情指数(Consumer Public Opinoin Index, CPOI)。研究结果显示, 本文构建的 CPOI 与 CCI 之间存在显著的正相关, 且存在长期稳定的均衡关系。进一步对消费者信心进行预测, 加入网络舆情因素的模型预测能力显著提升, 且 LSTM 模型预测的准确性和鲁棒性优于 GRU 和随机森林。

本文的创新之处在于: 首先, 首次将 Bert 预训练情感模型应用于消费舆情分析, 模型效果显著优于传统深度学习方法; 其次, 融合微博热度和小波变换, 构建了一个综合量化网络消费舆情的指标 CPOI, 更精确刻画消费舆情; 最后, 为评价该指数的有效性, 验证了其与消费者信心的相关性及其长期均衡关系, 证明了网络舆情对消费者预期的预测能力, 填补了相关研究空白。

2. 文献回顾

2.1. 基于深度学习的文本挖掘方法

近年来, 基于深度学习的文本挖掘方法得到了广泛应用, 尤其是在情感分析任务中展现了其优势。

Jelodar 等(2020)利用 LSTM 对 COVID-19 评论进行情感分析, 取得了优于传统机器学习算法的表现[8]; Olabanjo 等(2023)则在总统选举舆论分析中使用神经网络模型, 在多个维度超越 LSVC 模型[9]。2018 年, BERT [10]横空出现, 在 11 项 NLP 任务中做到业界最强, 被广泛应用于各类情感挖掘任务。例如, 王建成等(2020)用神经主题模型识别对话的情感极性[11], 綦方中和田宇阳(2020)基于 Bert 和 LDA 混合模型对酒店评论进行情感挖掘[12], Hartmann 等(2023)提出的 SiEBERT 模型对比词典和传统机器学习模型中表现最佳[13], Olabanjo 等(2023)的 BSERT 模型也超过了 LSTM 模型[9]。

2.2. 社交媒体舆情与消费者信心

随着媒体格局的变化, 社交媒体已成为主要信息来源[14]。越来越多的研究者开始从社交媒体文本中挖掘信息, 探索其价值。例如, Mao 等(2015)基于 Twitter 文本提取的网民情感不仅能够反映投资者情绪的变化, 还能预测股票市场未来表现[15]; 崔炎炎和刘立新(2022)对金融创新股票相关微博进行投资者情绪挖掘和股价预测, 发现负向投资者情绪对收盘价具有负向影响[16]; Angelico 等(2022)则建立通胀预期指标, 显示社交媒体对传统调查得到的通胀预期指标的预测潜力[17]。在消费者信心领域, 研究还主要集中在其对经济变量和周期的驱动作用上[18]-[20], 从社交媒体消费舆情视角探究舆情对消费者信心的研究有限。Lehrer 等(2019)将社交媒体情绪指标 USSI [21]应用于美国消费者信心指数的预测, 显著提升了预测效果[22]; Qiu (2020)进一步优化算法, 提升了准确性[23]; 张一帆等(2023)利用媒体文本构建情绪指数, 挖掘其对消费增速预测的独特价值[24]; Zhang *et al.* (2024)也证明了消费者情绪与社交媒体网红情绪之间的关联[25]。

综上, 本文利用 Bert 预训练情感模型挖掘社交媒体消费舆情, 对其在消费者信心方面的价值进行验证和评价, 期望为该领域研究补充更加全面、准确的视角。

3. 网络消费舆情指数的构建

我们设计构建一个名为网络消费舆情指数(Consumption Public Opinion Index, CPOI)的指标, 量化社交媒体环境中的消费舆情。研究路径可以划分为数据获取与处理、意见挖掘和指数构建三个阶段, 见图 1。本章节对每个阶段的内容进行详细介绍。

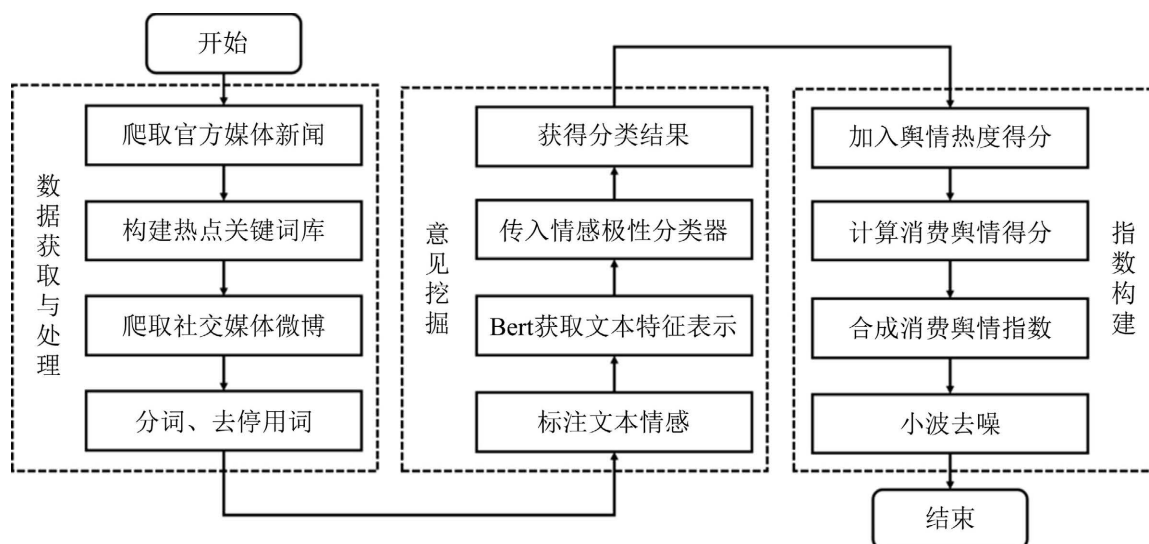


Figure 1. CPOI building framework
图 1. CPOI 构建框架

3.1. 数据获取与处理

3.1.1. 构建消费热点关键词库

为获取当下与消费者信心相关的不同维度的热点主题和关键词，本文选取了中国主流媒体之一的新华网财经频道，爬取了 2023 年 5 月至 2024 年 4 月近一年发布的 1986 条每日新闻。通过词频技术提取文本中的前 1%高频词，并过滤掉与新闻主题无关的停用词和过于泛化的词汇。经过筛选后的 Top 100 高频词云图见图 2：



Figure 2. Cloud map of high-frequency words
图 2. 高频词云图

消费者信心包含对众多宏观经济和家庭层面变量的预期[26]，本文综合参考上述高频词、密歇根指数[27]和中国信心指数[28] [29]的调查内容后，选取了经济形势、生活状况、投资、就业、物价水平和购房 6 个维度的 25 个关键词，构成的消费舆情关键词库见表 1：

Table 1. Consumer public opinion keyword database
表 1. 消费舆情关键词库

主题维度	关键词
经济形势	经济、产业、外贸、利率、GDP、汇率、企业
生活状况	消费、收入、存款、旅游
投资	投资、基金、债券、黄金、股票
就业	就业、失业、创业
物价水平	CPI、物价、价格
购房	房地产、楼市、买房

3.1.2. 微博爬虫

截止 2023 年第四季度, 微博平台月活跃用户达到 5.98 亿, 是当今中国最大的公共信息发布平台之一。网民在平台上发布帖子和交流, 高互动性的热门微博往往承载着更广泛的公众意见。基于已构建的关键词库, 文本以每个关键词作为种子, 爬取了微博平台内 2015 年 1 月至 2024 年 5 月每日前 50 条热门微博, 共计约 178.2 万条, 构成微博文本数据集。各主题维度的文本数量及内容见表 2:

Table 2. Number and content of microblogs

表 2. 微博数量及内容

主题维度	微博数量	例子
经济形势	229,933	(1) 我国经济的亮点是出口的结构更好了, 由低附加值产品转成了 90%机电产品 + 10%劳动密集型产业。 (2) 我觉得 2024 年可能是中国经济最困难的一年。
生活状况	444,802	(1) 去年至今, 我国消费恢复保持良好势头, 各种场景式消费不断创新, 特别是文旅消费热度不减。 (2) 虽然今年收入下滑, 消费也降级, 但好在生活质量没下降太多。
投资	533,206	(1) 截至收盘北向资金净流入 128.09 亿元, 外资积极流入明显, 周三微博里我就明确说了看好 11 月份, 今天依然重复这句话。 (2) 中国股市的冷淡让人担忧, 如果再持续下去, 可能会激起民心的不安, 造成严重影响。
就业	236,853	(1) 今年以来, 我国就业形势继续保持稳中向好态势, 市场用工需求旺盛, 重点群体就业基本稳定, 失业率稳中有降。 (2) 现在就业形势的现实: 从当下开始, 任何人都要做好失业在家的准备。
物价水平	161,572	(1) 国家统计局新闻发言人毛盛勇表示, 尽管 CPI 涨幅扩大, 但主要受结构性因素影响, 随着一系列政策落地, 价格也会回归正常区间。目前既不存在通胀, 也不存在通缩, 物价总体是平稳的。 (2) 近日一名外国游客表示在西湖游玩时一瓶矿泉水要 6 块, 嘉兴西塘酒吧街半打啤酒 600 块, 感觉物价太贵太离谱。
购房	176,060	(1) 重磅利好, 楼市终于出现回暖迹象了, 没想到啊, 你看看, 深圳楼市这人山人海的局面已经好多年都没有看到了。 (2) 为什么这一次楼市市场反应不佳? 因为市场缺的不是需求而是信心, 虽然上半年颁布过一些政策, 但都很“含蓄”, 很难让老百姓有信心。

3.2. 舆情挖掘

3.2.1. Bert 预训练模型

Bert 基于 Transformer 的 Encoder 层堆叠而成, 这种结构使得 Bert 能够学习和抽取不同层次和粒度的语义信息, 获得更丰富的语义表示。同时, 相比 RNN 的顺序计算, Bert 也继承 Transformer 可以并行计算的优点大幅提高了运算速度。Bert 的基本结构见图 3:

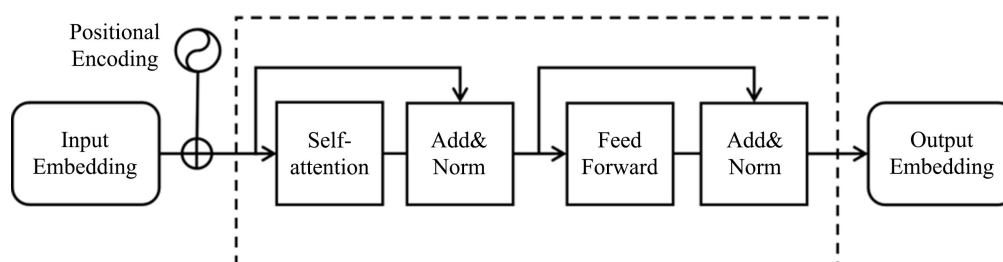


Figure 3. Bert model structure

图 3. Bert 模型结构

在输入层, 位置编码(Position Encoding) [30]额外提供每个 token 的位置信息, 用来识别序列的顺序关系。位置编码使用不同频率的正弦函数和余弦函数计算:

$$PE_{(pos, 2i)} = \sin\left(pos/10000^{2i/d_{model}}\right) \quad (1)$$

$$PE_{(pos, 2i+1)} = \cos\left(pos/10000^{2i/d_{model}}\right) \quad (2)$$

其中, pos 是句子中 token 的位置, d_{model} 是嵌入维度, i 代表向量维度索引, 公式所得到的位置编码维度和嵌入向量维度 d_{model} 相同, 二者相加一起输入至编码层。

在编码层, 自注意力机制(self-attention)是 Bert 的核心部分。用 $S_{embedding}$ 表示句子 S 经过输入层后的最终嵌入特征, 它的维度为 $sequence_{length} \times embedding_{dimension}$ ($sequence_{length}$ 是句子长度, $embedding_{dimension}$ 表示 token 的维度)。分配三个权重矩阵对每一个特征输入作线性映射, 得到 Query 矩阵(Q)、Key 矩阵(K)和 Value 矩阵(V):

$$Q = S_{embedding} W_Q \quad (3)$$

$$K = S_{embedding} W_K \quad (4)$$

$$V = S_{embedding} W_V \quad (5)$$

接着, 基于三个向量做以下计算:

$$Z = \text{attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (6)$$

式(6)中, 用注意力矩阵 QK^T 除以 $\sqrt{d_k}$ 转换为标准正态分布(d_k 通常取 64), 使得 softmax 归一化后的结果更加稳定, 从而帮助模型更易获得平衡的梯度。softmax 变换后, 得到的注意力权重分配与 V 相乘, 得到基于一个 attention 生成的 one-head 表示 Z 。将多个 Z 拼接并进行转换, 最终获得句子 S 的多头表示。

3.2.2. 情感挖掘

(1) 模型参数和评价指标

前期我们在每个主题维度文本中分别抽取 4000 条微博进行人工标注, 共 24,000 条。根据文本情感, 我们将其分为负向、正向和中性, 并用 -1、0、1 表示。负向文本表达负面情绪, 正向文本相反, 中性文本则为无明确情感倾向或与主题无关的内容。

本文采用 bert-base-chinese 模型进行训练。每个主题的标注文本的 80% 作为训练集, 20% 作为测试集。由于数据集中文本类别不平衡, 特别是中性文本占比过大, 可能导致模型偏向于预测大类别。为此, 我们调整了损失函数的权重, 给予正向和负向文本更高的权重, 以提升模型对小类别的识别能力。此外, Bert 模型的最大输入长度为 510 个 token。为保留微博文本中的关键信息, 超过长度的文本选取前 128 个和后 382 个 token。具体的模型参数设置见表 3。

本文基于 Bert 分别对 6 类主题文本进行建模, 并与 TextCNN 进行对比, 以正确率(accuracy)、精确率(precision)、召回率(recall)和 F1 得分(F1-score)作为模型评价指标:

$$\text{accuracy} = \frac{\sum_{i=1}^N TP_i}{\sum_{i=1}^N (TP_i + FN_i + FP_i)} \quad (7)$$

$$\text{precision} = \frac{1}{N} \sum_{i=1}^N \text{precision}_i = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FP_i} \quad (8)$$

$$\text{recall} = \frac{1}{N} \sum_{i=1}^N \text{recall}_i = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FN_i} \quad (9)$$

$$\text{F1-score} = \frac{2}{N} \sum_{i=1}^N \frac{\text{precision}_i \times \text{recall}_i}{\text{precision}_i + \text{recall}_i} \quad (10)$$

Table 3. Bert sentiment classification model parameters**表 3.** Bert 情感分类模型参数

参数名称	参数描述	参数设置
max_length	最大文本长度	510
train_batch_size	每批样本的大小	8
learning_rate	学习率	0.0001
epoch	全部训练集进行几次训练	20

(2) 情感分类结果

模型情感分类结果见表 4。从模型分类效果看, Bert 均优于 TextCNN, 特别体现在 Precision 和 F1-score 的提升。除此之外, Bert 对于不同主题维度文本分类的鲁棒性也显著优于 TextCNN, 经济形势、生活状况、投资、物价水平 4 个主题的所有评价指标都高于 0.85, 就业、购房主题高于 0.8, 而 TextCNN 在这两类主题中的表现与其他主题差距较大。Bert 的表现验证了其在挖掘消费者预期方面的优越性和稳定性。本文用 Bert 对剩余所有未标注文本进行分类。

Table 4. Model classification result**表 4.** 模型分类结果

主题维度	Bert				TextCNN			
	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score
经济形势	0.8887	0.8910	0.8887	0.8890	0.8775	0.8116	0.8775	0.8273
生活状况	0.8787	0.8639	0.8788	0.8694	0.8575	0.7473	0.8575	0.7986
投资	0.8750	0.8700	0.8750	0.8712	0.8425	0.7205	0.8425	0.7767
就业	0.8462	0.8487	0.8462	0.8459	0.7425	0.6350	0.7425	0.6630
物价水平	0.8575	0.8561	0.8575	0.8568	0.8525	0.7601	0.8525	0.7881
购房	0.8025	0.8003	0.8025	0.8013	0.7125	0.6432	0.7125	0.6257

3.3. 网络消费舆情指数构建

3.3.1. 微博舆情得分

首先定义单条微博的舆情得分, 它由情感得分和热度得分组成。情感得分根据 Bert 情感分类结果得到:

$$\text{score}_{\text{sentiment}} = \begin{cases} 1, & \text{正向情感} \\ 0, & \text{中性情感} \\ -1, & \text{负向情感} \end{cases} \quad (11)$$

微博热度的高低能反映网民的情感强度, 具体可以表现在点赞、转发、评论量上。因此, 本文创新性地提出热度得分:

$$\text{score}_{\text{hot}} = \sum_{i=1}^3 \theta_i \text{num}_i, i=1, 2, 3 \quad (12)$$

其中, num_i 是点赞、评论、转发的数量, θ_i 是权重, 本文 $\theta_1 = \theta_2 = \theta_3 = \frac{1}{3}$ 。

舆情得分由热度得分和情感得分相乘得到:

$$\text{score}_{\text{composite}} = \text{score}_{\text{sentiment}} \text{score}_{\text{hot}} \quad (13)$$

舆情得分同时考虑了微博的情感倾向和热度。符号反映情感倾向, 绝对数值反映情感强度。中性微博的舆情得分为 0。

3.3.2. 网络消费舆情指数

按时间窗口划分, 计算不同维度的月度网络消费舆情子指数(Consumer Public Opinion Sub-Index, CPOSI):

$$\text{CPOSI}_i = \sum \frac{\text{total}_{\text{positive}} + \text{total}_{\text{negative}}}{\text{total}_{\text{positive}} - \text{total}_{\text{negative}}}, i=1, \dots, 6 \quad (14)$$

其中, $\text{total}_{\text{positive}}$ 是积极情感微博的消费舆情得分总和, $\text{total}_{\text{negative}}$ 相反。

参考 CCI 的尺度对子指数加总后的结果进行放大, 得到月度网络舆情指数:

$$\text{CPOI} = 100 + \frac{\sum_{i=1}^6 \text{CPOSI}_i}{6} \times 100 \quad (15)$$

与 CCI 一致, CPOI 以 100 为中值, 在 0 至 200 之间变动。CPOI > 100 表明网民对当前经济和消费形势的态度以及未来预期总体积极, CPOI < 100 时则相反。

3.3.3. 小波变换

根据上述网络舆情指数的构建方法和情感分类结果, 本文计算得到了原始的网络消费舆情指数, 但由于随机波动与非系统性噪声呈现出较大的波动。因此, 采用小波变换对其进行去噪和趋势提取。小波变换可通过多分辨率分析捕捉信号特征并去除噪声, 过程包括小波分解、阈值处理和信号重构。多分辨率分析框架下的信号重构表示为:

$$f(n) = \sum_{k=-k_0}^{K_{j_0}} c_{j_0,k} \phi_{j_0,k}(2^{j_0}n-k) + \sum_{j=j_0}^J \sum_{k=-k_0}^{K_j} d_{j,k} \varphi_{j,k}(2^j n-k), n=0, 1, \dots, N_t-1 \quad (16)$$

其中, $\phi_{j_0,k}(2^{j_0}n-k)$, $\varphi_{j,k}(2^j n-k)$ 分别是尺度函数 $\phi(t)$ 和母小波函数 $\varphi(t)$ 的伸缩和平移函数簇; $c_{j_0,k}$ 和 $d_{j,k}$ 是伸缩尺度为 j_0, j , 平移尺度为 k 的展开系数, N_t 是时间序列的点数。

对原始的消费舆情指数进行 ADF 检验, p 值为 0.2297, 是非平稳信号。因此选择具有紧支撑性和正交性、适合非平稳信号处理的 Daubechies 4 (db4) 4 层小波分解, 并进行软阈值处理。小波变换后的 CPOI 见图 4, 降噪后的 CPOI 更清晰地提取了消费舆情的变化趋势。

3.4. 网络消费舆情指数与消费者信心

小波变换后的 CPOI 和 CCI 见图 5。直观地看, 2021 年之前, CCI 呈现先缓慢上升后稳定的趋势。CPOI 虽然有一定波动, 但与 CCI 基本一致。从 2021 年开始, CPOI 将逐渐下降到最低水平, 并从 2023 年底开始略有回升。另一方面, CCI 在 2022 年初急剧下降后, 一直保持在较低水平。我们对两者的相关性进行分析, 结果表明 Pearson 相关系数达 0.8775, p 值远小于 0.01。由此表明, CPOI 与 CCI 可能存在正向作用关系, 随着 CPOI 的下降, CCI 也趋于下降。

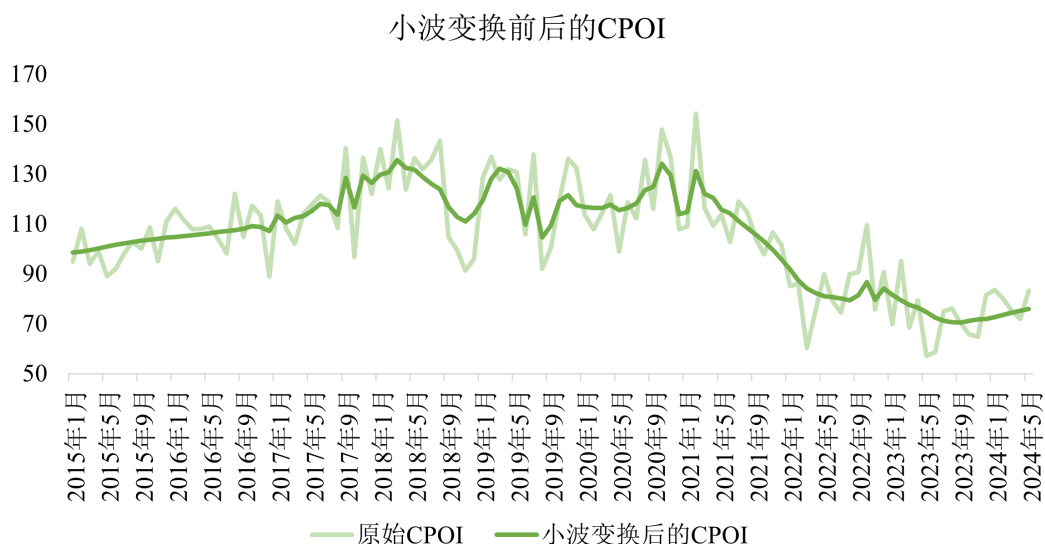


Figure 4. CPOI before and after wavelet transform

图 4. 小波变换前后的 CPOI

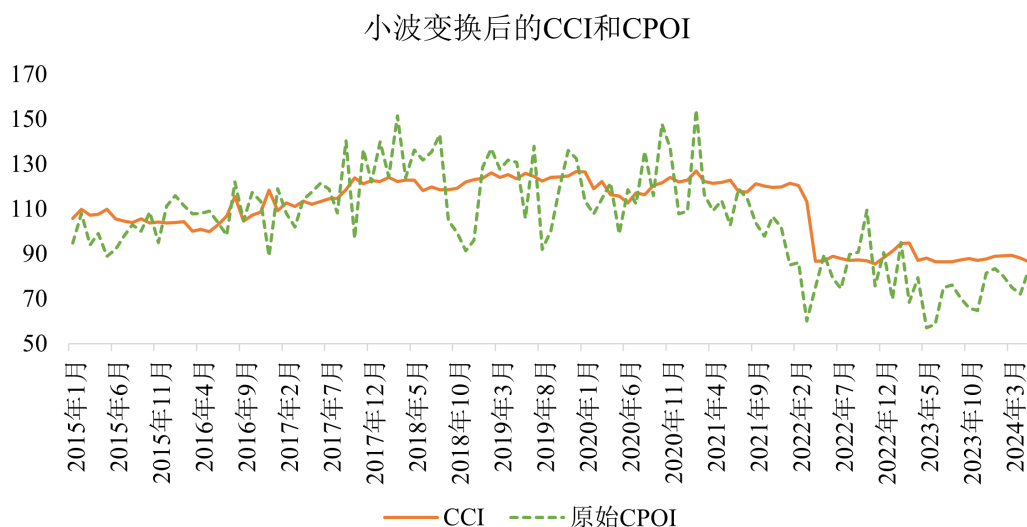


Figure 5. CPOI and CCI after wavelet transform

图 5. 小波变换后的 CPOI 和 CCI

为了证明这种关系是否是长期稳定的, 本文对两者进行协整检验。由于两个序列都是非平稳的, 对两者进行一阶差分, 确定其是同阶单整后采用 Engle-Granger 两步法。对 OLS 回归后的残差进行 ADF 检验, p 值为 0.0033, 因此在 1% 的显著性水平下可以认为调整后的 CPOI 和 CCI 之间存在长期稳定的均衡关系。

4. 基于 LSTM 模型的消费者信心预测

4.1. 模型变量和评价指标

作为一种特殊的 RNN, 长短期记忆神经网络(LSTM) [31]解决了传统 RNN 由于长期依赖导致的梯度

消失和梯度爆炸的问题，非常适合用于时间序列模型预测。与标准的 RNN 单元相比，LSTM 引入了一种门控机制，包括遗忘门、输入门、输出门三个控制门和单元状态。 t 时刻的记忆单元结构见图 6：

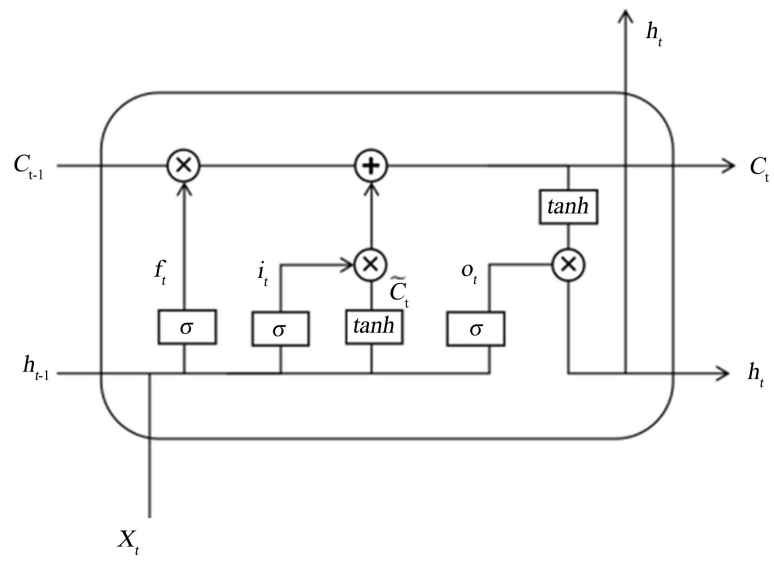


Figure 6. LSTM memory unit structure
图 6. LSTM 记忆单元结构

在时刻 t ，LSTM 接收输入特征 X_t 以及前一时刻的单元状态 C_{t-1} 和输出 h_{t-1} ，并通过三种门控机制过滤和更新信息。遗忘门负责选择需要遗忘的信息，输入门确定需要添加的新信息，并结合遗忘门和单元状态更新当前记忆。输出门则选择哪些信息是当前时刻的输出，最终生成一个新的状态和输出。

本文采用中国国家统计局发布的消费者信心指数作为被解释变量。在解释变量的选择上，考虑到传统的消费者信心指数预测往往以相关经济指标作为解释变量[32]，因此将 GDP 累计同比增速(GDPGt)、16~24 岁劳动力城镇调查失业率(UR16~24t)、居民消费价格指数(CPIt)、居民人均可支配收入累计同比增速(ICGt)、国房景气指数(RECI)、上证综合指数(SCI)分别作为反映经济形势、就业状况、物价水平、生活状况、购房和投资 6 个方面的宏观经济指标，用以建立传统的预测模型，与加入 CPOI 的模型进行对比。其中，GDP 累计同比增速和居民人均可支配收入累计同比增速是季度数据，考虑到数据总体波动不大，故将季度数据视作该季度内每个月的月度数据；16~24 岁劳动力城镇调查失业率存在个别月份的数据缺失，采用三次样条插值进行填补。LSTM 分组实验说明见表 5：

Table 5. LSTM group experiment description
表 5. LSTM 分组实验说明

模型	模型作用描述	模型输入描述
对照组 M1	进行结果对比	CCI _t 、GDPG _t 、UR16-24 _t 、CPI _t 、ICG _t 、RECI _t 、SCI _t
实验组 M2	验证消费舆情因素对模型预测的效果	CCI _t 、GDPG _t 、UR16-24 _t 、CPI _t 、ICG _t 、RECI _t 、SCI _t 、CPOI _t

为消除量纲和异常值的影响，提升模型的预测性能，用式 19 对所有变量进行极值归一化处理，预测完成后再通过逆算得到预测值。

$$N(x_i) = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (17)$$

为了评估不同模型以及同一模型的对照组和实验组的预测效果, 本文采用均方根误差(Root Mean Square Error, RMSE)、平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)、对称平均绝对百分比误差(Symmetric Mean Absolute Percentage Error, SMAPE)三个指标对模型预测效果进行评价:

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{n} \sum_{l=1}^n (y_l - \hat{y}_l)^2} \quad (18)$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{l=1}^n \left| \frac{y_l - \hat{y}_l}{y_l} \right| \quad (19)$$

$$\text{SMAPE} = \frac{100\%}{n} \sum_{l=1}^n \frac{2|y_l - \hat{y}_l|}{|y_l| + |\hat{y}_l|} \quad (20)$$

其中, y_l 是实际值, $\hat{y}_l$ 是预测值, n 是样本量。当 RMSE 越接近于 0、MAPE 和 SMAPE 越接近于 0% 时, 模型的误差越小, 预测精度越好。

4.2. 消费舆情与 CCI 预测

根据上文模型变量设计, 本文设置传统变量构成的对照组 M1 和加入 CPOI 的实验组 M2, 采用 Python 的 Tensorflow 框架分别进行 LSTM 建模。为了优化模型参数, 采用网格搜索(Grid Search)定义参数网格, 遍历所有参数组合并进行交叉验证(Cross-Validation), 从而选择最优参数组合。此外, 引入早停机制防止模型过拟合, 当验证损失在连续 5 个训练周期内没有改善时停止训练。本文最终设置参数见表 6:

Table 6. LSTM model parameter settings
表 6. LSTM 模型参数设置

参数名称	参数描述	参数设置
batch_size	批量大小	4
epochs	训练周期	20
optimizer	优化器	adam
loss	损失函数	均方误差 MSE
units	隐藏层神经元数量	50
learning rate	学习率	0.001
n_past	滑动窗口大小	3

以时间序列数据的前 80% 为训练集、后 20% 为测试集, 分别对对照组 M1 和实验组 M2 建立 LSTM 模型, 最终拟合效果见表 7:

Table 7. LSTM group experiment results
表 7. LSTM 分组实验结果

模型	RMSE	MAPE (%)	SMAPE (%)
对照组 M1	1.3833	1.2682	1.2647
实验组 M2	1.0207	0.9027	0.8989

通过对比对照组 M1 与实验组 M2 的表现可以清晰地看到, 在引入 CPOI 这一变量后, 实验组 M2 在

三个评价指标上的表现均展现出了显著的优越性，相较于对照组 M1 有着明显的提升。具体而言，实验组 M2 的 RMSE 降低了 26.21%，MAPE 下降了 28.82%，而 SMAPE 也实现了 28.92% 的降幅。这一结果不仅彰显了 CPOI 在优化预测模型中的关键作用，而且验证了网络消费舆情对提升消费者信心指数预测精度的积极影响。

4.3. 模型对比与稳健性

分组实验结果已经证实了加入消费舆情指数的实验组 M2 模型对预测样本消费者信心指数的更优表现。但为了进一步考量 LSTM 模型本身的预测效果，本文选择 GRU 和随机森林模型作为对照模型，并在长期(12 期)和短期(6 期)不同期限内检验 LSTM 模型的稳健性。不同期限模型预测结果见表 8：

Table 8. Prediction results of different maturity models
表 8. 不同期限模型预测结果

	短期			长期		
	RMSE	MAPE (%)	SMAPE (%)	RMSE	MAPE (%)	SMAPE (%)
LSTM	0.6786	0.6266	0.6251	0.7217	0.6287	0.6266
GRU	1.3468	1.2795	1.2678	1.0889	1.1325	1.1326
Random Forest	4.5474	5.1539	5.0240	5.0575	5.1895	5.0283

表 8 分别展示了在 3 个不同模型和长、短期限中样本 CCI 的预测结果。从预测效果来看，LSTM 模型的 RMSE、MAPE、SMAPE 无论在长期还是短期均远远优于其他两个模型。LSTM 三项指标均值分别为 0.7002、0.6277%、0.6259%，较 GRU 的 1.2179、1.2060%、1.2002%分别下降 42.51%、47.95%、47.85%，较 Random Forest 的 4.8025、5.1717%、5.0262%分别下降 85.42%、87.86%、87.55%。因此，可以判断，三个模型中 LSTM 在不同预测期限上的预测效果都是最佳的。

从稳健性来看，当增加预测期限时，LSTM 的 RMSE 上升了 6.35%，小于 GRU 和 Random Forest 的 23.68%、11.22%；LSTM 的 MAPE 上升了 0.34%，小于 GRU 和 Random Forest 的 13.0%、0.69%；LSTM 的 SMAPE 上升了 0.24%，小于 GRU 的 11.94%，稍大于 Random Forest 的 0.09%。尽管 LSTM 的 SMPE 指标稳健性表现略逊于 Random Forest，但总体来说 LSTM 的稳健性在三个模型中还是最优的。我们可以进一步得出，基于本文实证数据，LSTM 模型在预测效果和模型稳健性上均为最佳模型。

5. 研究结论

本文立足消费舆情视角，从对社交媒体文本的价值进行测度和评价。本文构建了 6 个主题维度下的 Bert 情感模型，对 2015 年 1 月至 2024 年 5 月的 178.2 万热门微博进行舆情挖掘，并在此基础上融合微博热度和小波变换构建了一个新的网络消费舆情指数量化舆情波动。随后，我们证明了 CPOI 和 CCI 的长期均衡关系并进一步验证了 CPOI 是预测消费者信心的重要指标。得出的主要结论如下：

首先，对 Bert 预训练模型进行微调后的情感建模准确挖掘了社交媒体环境中的消费舆情。其模型结果在准确率、精确率、召回率、F1 等多个评价指标中的表现均显著优于 TextCNN，显示出了预训练模型在社交媒体舆情分析领域的优越性。

其次，网络消费舆情指数量化社交媒体环境中的舆情变化，经检验与消费者信心有显著的正向关系。网络消费舆情指数和消费者信心的变化趋势具有一致性，存在长期稳定的均衡关系。从社交媒体文本挖掘消费舆情可以一定程度上反映消费信心的变动情况。

最后，社交媒体舆情影响着市场主体关于宏观经济的判断和预期，对预测消费者信心具有重要作用。

本文将构建的 CPOI 加入对消费者信心的预测模型中, 预测效果得到显著提升; 同时, 对比 GRU 和 Random Forest, 本文构造的 LSTM 模型在准确性和稳健性都更佳。

基于社交媒体文本挖掘的消费舆情提取可以有效弥补传统消费者信心指数的滞后性, 为进一步准确掌握消费和经济运行趋势, 及时响应市场变化提供有效参考。

基金项目

本研究由浙江省大学生科技创新活动计划项目(批准号: 2024R407C068)资助。

参考文献

- [1] Gelper, S., Lemmens, A. and Croux, C. (2007) Consumer Sentiment and Consumer Spending: Decomposing the Granger Causal Relationship in the Time Domain. *Applied Economics*, **39**, 1-11. <https://doi.org/10.1080/00036840500427791>
- [2] Barsky, R.B. and Sims, E.R. (2012) Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence. *American Economic Review*, **102**, 1343-1377. <https://doi.org/10.1257/aer.102.4.1343>
- [3] Deng, S., Sinha, A.P. and Zhao, H. (2017) Adapting Sentiment Lexicons to Domain-Specific Social Media Texts. *Decision Support Systems*, **94**, 65-76. <https://doi.org/10.1016/j.dss.2016.11.001>
- [4] Yiu, D.W., Wan, W.P., Chen, K.X. and Tian, X. (2021) Public Sentiment Is Everything: Host-Country Public Sentiment toward Home Country and Acquisition Ownership during Institutional Transition. *Journal of International Business Studies*, **53**, 1202-1227. <https://doi.org/10.1057/s41267-020-00380-4>
- [5] Statista (2024) Number of Social Network Users Worldwide from 2017 to 2028 (in Billions). <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>
- [6] Conrad, F.G., Gagnon-Bartsch, J.A., Ferg, R.A., Schober, M.F., Pasek, J. and Hou, E. (2019) Social Media as an Alternative to Surveys of Opinions about the Economy. *Social Science Computer Review*, **39**, 489-508. <https://doi.org/10.1177/0894439319875692>
- [7] Shiller, R.J. (2020) Narrative Economics: How Stories Go Viral and Drive Major Economic Events. Princeton University Press. <https://doi.org/10.1515/9780691212074>
- [8] Jelodar, H., Wang, Y., Orji, R. and Huang, S. (2020) Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach. *IEEE Journal of Biomedical and Health Informatics*, **24**, 2733-2742. <https://doi.org/10.1109/jbhi.2020.3001216>
- [9] Olabanjo, O., Wusu, A., Afisi, O., Asokere, M., Padonu, R., Olabanjo, O., et al. (2023) From Twitter to Aso-Rock: A Sentiment Analysis Framework for Understanding Nigeria 2023 Presidential Election. *Heliyon*, **9**, e16085. <https://doi.org/10.1016/j.heliyon.2023.e16085>
- [10] Devlin, J., et al. (2018) BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding.
- [11] 王建成, 徐扬, 刘启元, 等. 基于神经主题模型的对话情感分析[J]. 中文信息学报, 2020, 34(1): 106-112.
- [12] 蔡方中, 田宇阳. 基于 BERT 和 LDA 模型的酒店评论文本挖掘[J]. 计算机应用与软件, 2023, 40(7): 71-76+90.
- [13] Hartmann, J., Heitmann, M., Siebert, C. and Schamp, C. (2023) More than a Feeling: Accuracy and Application of Sentiment Analysis. *International Journal of Research in Marketing*, **40**, 75-87. <https://doi.org/10.1016/j.ijresmar.2022.05.005>
- [14] Gan, B., Alexeev, V., Bird, R. and Yeung, D. (2020) Sensitivity to Sentiment: News vs Social Media. *International Review of Financial Analysis*, **67**, Article ID: 101390. <https://doi.org/10.1016/j.irfa.2019.101390>
- [15] Mao, H., Counts, S. and Bollen, J. (2015) Quantifying the Effects of Online Bullishness on International Financial Markets. ECB Statistics Paper, 9.
- [16] 崔炎炎, 刘立新. 网络舆情赋能金融科技股票收盘价预测研究[J]. 统计研究, 2022, 39(6): 148-160.
- [17] Angelico, C., Marcucci, J., Miccoli, M. and Quarta, F. (2022) Can We Measure Inflation Expectations Using Twitter? *Journal of Econometrics*, **228**, 259-277. <https://doi.org/10.1016/j.jeconom.2021.12.008>
- [18] Mueller, E. (1957) Effects of Consumer Attitudes on Purchases. *The American Economic Review*, **47**, 946-965.
- [19] Linden, F. (1982) The Consumer as Forecaster. *Public Opinion Quarterly*, **46**, 353-360. <https://doi.org/10.1086/268732>
- [20] Bram, J. and Ludvigson, S.C. (1998) Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race. *Economic Policy Review*, **4**, 59-78.
- [21] Zumbun, J. (2017) A New Index Tracks Our National Mood One Tweet at a Time. *The Wall Street Journal*.

-
- <https://blogs.wsj.com/economics/2017/05/08/a-new-index-tracks-our-national-mood-one-tweet-at-a-time/>
- [22] Lehrer, S., Xie, T. and Zeng, T. (2019) Does High-Frequency Social Media Data Improve Forecasts of Low-Frequency Consumer Confidence Measures? *Journal of Financial Econometrics*, **19**, 910-933. <https://doi.org/10.1093/jfinec/nbz037>
- [23] Qiu, Y. (2020) Forecasting the Consumer Confidence Index with Tree-Based MIDAS Regressions. *Economic Modelling*, **91**, 247-256. <https://doi.org/10.1016/j.econmod.2020.06.003>
- [24] 张一帆, 林建浩, 樊嘉诚. 新闻文本大数据与消费增速实时预测——基于叙事经济学的视角[J]. 金融研究, 2023(5): 152-169.
- [25] Zhang, Z., Keasey, K., Lambrinoudakis, C. and Mascia, D.V. (2024) Consumer Sentiment: The Influence of Social Media. *Economics Letters*, **237**, Article ID: 111638. <https://doi.org/10.1016/j.econlet.2024.111638>
- [26] Armantier, O., et al. (2017) An Overview of the Survey of Consumer Expectations. *Economic Policy Review*, **23**, 51-72.
- [27] Curtin, R.T. (1982) Indicators of Consumer Behavior: The University of Michigan Surveys of Consumers. *Public Opinion Quarterly*, **46**, 340-352. <https://doi.org/10.1086/268731>
- [28] 郭洪伟. 我国消费者信心指数编制及存在的问题[J]. 商业时代, 2010(33): 28-29.
- [29] 郭洪伟. 基于网络大数据的消费者信心指数编制[J]. 统计与信息论坛, 2015, 30(6): 111-112.
- [30] Vaswani, A., et al. (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.
- [31] Hochreiter, S. (1997) Long Short-Term Memory. *Neural Computation*. MIT Press.
- [32] 孙景, 朱建霖, 李挽澜, 等. 基于互联网数据的消费者信心指数滚动预测研究[J]. 西安交通大学学报(社会科学版), 2021, 41(6): 68-77.