

基于Bert-BiLST模型研究投资者情绪对股票收益率的影响

——以工商银行为例

姜 云

贵州大学经济学院, 贵州 贵阳

收稿日期: 2025年3月31日; 录用日期: 2025年4月15日; 发布日期: 2025年5月23日

摘 要

投资者情绪是影响股市走向的因素之一。本文爬取东方财富网工商银行股吧2024年1月1日至2024年12月31日评论, 采用Bert-BiLSTM模型对投资者情绪进行文本分析并判断情绪类型、测算每日投资者情绪指标。再通过基准回归研究投资者情绪指数与工商银行股票的关系。结果发现: (1) 投资者情绪预测中, Bert-BiLSTM模型比Bert模型、BiLSTM模型更有效; (2) 投资者情绪与工商银行股票收益率呈正相关关系。 (3) 不同市场环境下, 投资者情绪对工商银行股票收益率有差异性影响。根据结果对监管机构与投资者分别提出建议。

关键词

投资者情绪, Bert-BiLSTM, 股票收益率, 工商银行

Research on the Impact of Investor Sentiment on Stock Returns Based on the Bert-BiLST Model

—A Case Study of Industrial and Commercial Bank of China

Yun Jiang

School of Economics, Guizhou University, Guiyang Guizhou

Received: Mar. 31st, 2025; accepted: Apr. 15th, 2025; published: May 23rd, 2025

Abstract

Investor sentiment is one of the factors that influence the direction of the stock market. This paper

crawls the comments from the Industrial and Commercial Bank of China's stock bar on Eastmoney.com from January 1, 2024, to December 31, 2024. It uses the Bert-BiLSTM model to analyze investor sentiment from the text and determine the type of sentiment, calculating the daily investor sentiment index. The relationship between the investor sentiment index and the stock performance of the Industrial and Commercial Bank of China is then studied using a benchmark regression. The results show that: (1) The Bert-BiLSTM model is more effective in predicting investor sentiment than the Bert model and the BiLSTM model; (2) Investor sentiment is positively correlated with the stock returns of the Industrial and Commercial Bank of China; (3) The impact of investor sentiment on the stock returns of the Industrial and Commercial Bank of China varies under different market conditions. Based on the results, suggestions are made for regulatory authorities and investors.

Keywords

Investor Sentiment, Bert-BiLSTM, Stock Returns, Industrial and Commercial Bank of China

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. 引言与文献综述

经典的金融学架构中认为投资者都是理性的,投资者的任何投资都不会受到其他因素的影响。然后在行为金融学领域中,投资者往往是非理性的[1]。其中影响投资者决策的因素众多,如个人因素中的个性特征、社会经济地位和收入、教育水平等;心理因素中的情绪与认知偏差;行为因素中的“羊群效应”以及经济与市场环境等[2]-[5]。而股市作为当今企业筹资与个人投资者进行金融交易的重要场所,股价是股市变化最直接的表现与信号。股价的变化也同样受到多重因素的作用,投资者的行为也列为其中。

进行投资决策之前,人们除了通过K线等硬性指标预测股价的走势外,投资者情绪这一软性指标的加持更加事半功倍。对于投资者情绪这一指标的量化方法重要在于机器学习与深度学习中。主要是通过爬虫和自然语言处理技术对各种网络媒体平台与股票网站股民的评论进行挖掘与分析,其中主要通过Twitter、东方财富、雪球网等[6][7]。常用的机器学习方法有朴素贝叶斯与支持向量机(SVM)。随着机器学习的不断发展,深度学习崭露头角,卷积神经网络(CNN)、长短期记忆网络(LSTM)与双向LSTM(BiLSTM)等模型更能准确的进行语言文本的分析。现在,学者更多使用多种模型结合预测情绪,如Wang等提出AE-LSTM模型解决了LSTM模型不能区分不同情感的重要性的弊端,进一步提高了情绪预测的准确性[8]。张科等为探究发帖人中意见主导者的情绪对股票的影响,构建Stack-LSTM模型进行股吧评论分析与社会网络识别发帖人进行探究[9]。在Bert模型与其他模型创新上,杨飘和董文勇用Bert与BGRU-CRF结合[10]、王子牛等提出与BiLSTM-CRF模型结合以及孙弋与梁兵涛采用的Bert-BiGRU-CRF混合模型均表现出Bert与其他文本分析模型结合都要优于基础模型的预测效果[11][12]。

投资者情绪与股票的关系上,不少学者也通过机器学习将文本分析与股票的指标结合,预测股票的走势和探究投资者情绪对股票的影响。大部分学者发现投资者情绪指数对股票的收益率有着显著的影响[13][14]。陈志芳等发现投资者情绪指数与投资者分歧度均对股票收益率有显著的影响,且能预测未来几期股票交易量[15]。赵庆国等通过GED-EGARCH(1,1)-M模型研究中国不同行业股票与投资者情绪的关系,发现大部分行业可以敏感察觉到投资者情绪的变化,进而影响股市的走势。但有一部分学者持相反观点[16]。部慧等发现虽然该指数对收益率有影响,但是对于股票的收益率是没有预测性的[17]。Li等通

过情绪代理变量研究小盘股收益与投资者情绪的关系,发现情绪的变化对小盘股收益没有显著性影响[18]。

根据现有文献发现,较少文章研究投资者情绪在不同市场环境下对股票的影响,较多学者采用混合的语言文本分析模型提高情感分析的准确性。故本文将采用混合的 Bert-BiLSTM 模型对东方财富网股吧评论进行更有效的情感预测,并根据计算形成每日投资者情绪指数,分析该指数对股票收益率是否有影响,并研究不同环境下影响是否有差异。

2. 理论假设

2018 年谷歌提出由 MLM 和 NSP 两个自监督任务组成的 Bert 模型,其架构为输入层、中间层、输出层三层构成。Bert 模型独特之处在于:第一,模型的中间层为一个双向的 transformer 编码器,以提高对文本的识别;第二,为了提高对文本情绪的预测,在 MLM 预训练任务中,模型会先对文本中部分词汇进行遮盖再对被遮盖的词汇预测,从而对情绪与词汇之间的关系更加了解。然后 NSP 模型会通过判断两个句子的连接性,更深入的提高模型对文本的理解。Bert 模型对全局语义的理解更为透彻,往往忽略上下文的信息与细节,同时其对输入的文本长度也有所限制。

BiLSTM 模型是一种特殊的循环神经网络(RNN),它结合了长短期记忆网络(LSTM)的特性和双向数据处理的能力。其创新之处在于从序列的开始到结束的数据处理由正向 LSTM 层完成,而从序列的结束到开始由反向 LSTM 层完成。这种双向 LSTM 的处理方式,使其可以在分析文本信息时能够更加关注到文本的细节信息,这一点就能弥补上 Bert 模型独自进行文本分析的缺点了。同时 BiLSTM 模型可以使用更多的任务中,如:文本分类、情感分析、命名实体识别等任务中;股票价格预测、天气预测等时间序列任务中以及语音识别等。尽管如此,该模型理解的深度和广度确稍逊色,且其在处理长序列时可能会遇到梯度消失或梯度爆炸的问题。

综上 Bert 与 BiLSTM 模型的结合会形成一加一大于二的效果,得出假设 H1。第一,BERT 可以提供更丰富的语义特征,而 BiLSTM 则可以进一步分析这些特征,捕捉更细粒度的上下文关系;第二,BERT 需要大量的标注数据进行微调,而 BiLSTM 可以在较小的数据集上进行训练。在标注数据有限的情况下,可以先使用 BERT 提取特征,然后用 BiLSTM 进行进一步分析,从而在一定程度上降低对标注数据的需求;第三,在某些情况下,BERT 或 BiLSTM 可能会因为各自的特点和局限性而产生错误。结合使用时,一个模型的错误可能被另一个模型的正确判断所补偿,从而提高整体模型的鲁棒性。

H1: Bert-BiLSTM 模型情感分析方法能够提供比基础模型更高的预测准确度。

随着社交媒体和互联网技术的不断发展,投资者和交易者在这些平台上的交流日益频繁。用户在社交媒体上生成的内容蕴含着大量情绪和观点信息,这些信息能够揭示投资者的情绪状态和市场预期[19]。同时社交媒体和网络平台上,投资者的情绪可以通过信息传播迅速扩散,形成一种集体情绪。这种集体情绪能够对市场产生更大的影响,尤其是在市场波动较大的时期[20]。根据行为金融学理论中的羊群效应理论,投资者往往跟随他人行为,形成羊群效应。当大量投资者跟风某一特定资产或投资策略时,市场可能会出现过热现象,资产价格被高估。简而言之,当某些投资者看好某只股票时,他们将在网络上发表自己的言论与观点等行为,由于羊群效应的存在,这种“看好”行为的影响会不断的传播和加深,导致对该股票看好并持有的人增多,使股价高于其实质价值,进而造成市场波动。

在不同的市场环境中,这种效应会加剧市场波动。例如,在上涨行情中,乐观情绪可能引发更多投资者买入,进一步推高股价;而在下跌行情中,悲观情绪可能导致恐慌性抛售,加剧股价下跌。再者,李瑜发现投资者情绪对股票市场收益率存在显著影响[21]。在牛市中,投资者情绪高涨会显著提高市场收益率,且情绪波动幅度的加大会导致股票价格波动幅度的加大。然而,在熊市下,投资者情绪对收益率及波动性的影响并不显著。

- H2: 投资者情绪对股票收益率波动有显著影响。
- H3: 在不同的市场环境下, 投资者情绪对股票收益率产生差异化影响。

3. 实证设计

(一) 样本选择与数据来源

根据 ifind 同花顺 2025 年 1 月 1 日沪深股票总市值的排名, 选择位列第三的工商银行作为本文的研究对象。考虑构建近期的投资者情绪与研究其短期的影响, 仅选取 2024 年 1 月 1 日至 2024 年 12 月 31 日为样本区间, 由于股票只有在交易日进行交易, 则该区间的交易日有 242 天。

通过 Python 对东方财富股吧中工商银行(6001398) 2024 年 1 月 1 日至 2024 年 12 月 31 日的所有评论标题进行爬取, 作为投资者情绪指数的文本。通过 Tushare 网站获取工商银行股票的涨跌幅、交易量、换手率、总市值及沪深 300 指数的涨跌幅五个每日指标。

(二) 变量选择

参考现有研究, 被解释变量 Pct 股票收益率通过股票涨跌幅刻画, 同时选择选择股票交易量 Vol、总市值 Size、换手率 Turn 与市场收益率 T_pct 作为控制变量, 其中为统一量化标准, 对 Vol、总市值取对数(具体见表 1)。对于核心解释变量 SENT, 通过 Bert-BiLSTM 模型判断出每一条评论的情绪类别, 参考金德环和李岩做法[22], 以下公式(1), 计算每日投资者情绪价值:

$$SENT_t = \frac{pos_t - neg_t}{pos_t + neg_t}$$

(1)

其中, pos_t 为第 t 期正面情绪的股票评论数, neg_t 为第 t 期负面情绪的股票评论数, $SENT_t$ 为第 t 期投资者情绪指数, 且取值在-1 至 1 区间。

Table 1. Variable description
表 1. 变量说明

变量		来源与计算
Pct	股票收益率	Tushare 数据库获得的涨跌幅作为指标
SENT	投资者情绪指数	东方财富网爬取评论后, 采用 Bert-Bilstm 模型与公式(1)计算
Vol	股票交易量	Tushare 数据库获得, 取对数 $\ln(1 + vol)$
Size	总市值	Tushare 数据库获得, 取对数 $\ln(1 + size)$
Turn	换手率	Tushare 数据库获得
T_pct	市场收益率	以沪深 300 指数作为市场指数收益率
pos	正面情绪评论数	Bert-Bilstm 模型预测结果
neg	负面情绪评论数	

(三) 模型设计

1. Bert-Bilstm 模型设计

(1) 文本处理

对爬取到工商银行 2024 年 42,709 评论清洗: 重复出现的标题只留下第一次出现的、标题中含有特殊标记字符如【】与\$、标题只包含图片、表情与符号数字的、删除官方的公告与政策解读等。同时, 由于股票仅在交易日进行交易, 则剔除非交易日评论, 以便后续完成基准回归。一系列清洗后, 数据剩余

34,261 条。

在清洗后的数据中，为保证时间序列特征，按照时间分层抽样选择。其中 1) 按月份计算每一天的评论数量，然后按照比例从每个月的评论中抽取样本；2) 根据每一天的评论数占当月总评论数的比例，抽取当天的样本；3) 保证每一天至少抽取一条样本，即每一天都包含在抽样结果中。选择出 1920 条并进行人工标注情绪类型，规定情绪类型为“正面”、“负面”、“中性”三类。

(2) Bert-BiLSTM 模型情绪训练

将 1920 条按照 8:2 的比例划分训练集与测试集。进行 Bert 模型的预训练与微调两步。其中，预训练中调用“bert-base-Chinese”模块，参数设置为：隐藏层的 hidden_dim = 128，最大文本长度为 max_length = 128，训练样本的数量 batch_size = 32，transformer = 12。在微调阶段使用 Adam 优化器，hidden_dim = 768，bidirectional = True，学习率为 lr = 2e-5，训练次数为 epoch = 5。

(3) 构建模型评估指数

其一，构建的 Accuracy (准确率)、Precision (精确率)、Recall (召回率)和 F1 值(F1-Score)四个指标来评价实验模型。

其二，通过 Bert 模型、BiLSTM 模型与 Bert-BiLSTM 模型预测评价的情绪，对模型的性能进行评价比较。

(4) 保存模型参数，批量文本处理

完成上述操作，保存已训练好的各模型参数。将未预测的样本导入模型中，进行文本解读与情绪类型预测。

2. 基准回归

构建投资者情绪对股票收益率的基准回归，见公式(2)：

$$SENT_{t+1} = \alpha_1 Pct_t + \alpha_2 Vol_t + \alpha_3 Turn_t + \alpha_4 T_pct_t + \alpha_5 Size_t + \varepsilon_t \quad (2)$$

其中， t 表示时间。被解释变量为股票收益率 Pct_t 表示第 t 期工商银行的股票收益率，核心解释变量为投资者情绪指数 $SENT_{t+1}$ 表示第 $t+1$ 期投资者的情绪， ε_t 为随机扰动项。

4. 实证估计

(一) 投资者情绪指数模型预测

根据上述模型设计，完成 Bert、BiLSTM 与 Bert-BiLSTM 三个模型的训练与预测。获得各模型的准确率(Accuracy)、精确率(Precision)、召回率(Recall)和 F1 值，见表 2 与图 1。

Table 2. Prediction effects of bert, BiLSTM, and Bert-BiLSTM models

表 2. Bert、BiLSTM 与 Bert-BiLSTM 模型的预测效果

模型	Accuracy	Precision	Recall	F1
Bert	0.7135	0.7117	0.7135	0.7136
BiLSTM	0.7507	0.6392	0.7710	0.6990
Bert-BiLSTM	0.7699	0.8118	0.7823	0.7310

可知，Bert-BiLSTM 模型在准确率、精确率、召回率和 F1 值为 0.7699、0.8118、0.7823、0.7310。与 Bert、BiLSTM 两个模型的指标相比，更为优越，且在精确率 Precision 中有明显的提高。由此进一步证明本文提出的假设 H1，Bert-BiLSTM 模型情感分析方法能够提供比基础模型更高的预测准确度。

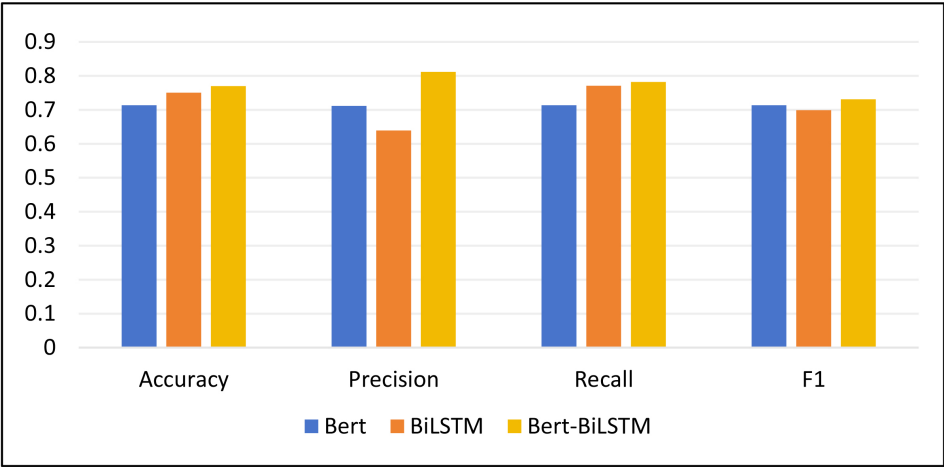


Figure 1. Comparison of prediction effects of Bert, BiLSTM, and Bert-BiLSTM models
图 1. Bert、BiLSTM 与 Bert-BiLSTM 模型预测效果对比图

(二) 投资者情绪对股票收益率的影响

1. 描述性统计

回归前进行描述性统计，见表 3。通过分析发现股票收益率(Pct)的标准差为 1.2831，波动范围为 -4.1602 至 4.8951，显示收益率波动较大，投资者面临一定的风险。投资者情绪指数(SENT)均值为 -0.0238，标准差为 0.2527，最小值为 -0.8378，最大值为 0.6154。这说明整体上该变量的值偏负但幅度不大，即整体投资者情绪偏负向。换手率(Turn)的标准差为 0.0665，最小值与最大值为 0.0529 与 0.6262 均表示存在一定的波动。市场收益率(T_pct)标准差为 1.3472，最小值与最大值为 -7.0514 到 8.4826，说明市场存在较大的波动范围。对正面情绪评论数(pos)与负面情绪评论数(neg)两者之间的数据进行对于，从均值比较，为 40.9959 与 46.1942，表示投资者负面情绪更明显；从标准差比较，为 76.3249 与 92.8217，即投资者正面负面情绪表达有显著的波动性。

Table 3. Descriptive statistics
表 3. 描述性统计

变量	观测值	均值	标准差	最小值	最大值
Pct	242	0.1919	1.2831	-4.1602	4.8951
SENT	242	-0.0238	0.2527	-0.8378	0.6154
Vol	242	14.9916	0.3661	14.1711	16.6418
Size	242	19.1301	0.0846	18.9534	19.3278
Turn	242	0.1301	0.0665	0.0529	0.6262
T_pct	242	0.0656	1.3472	-7.0514	8.4826
pos	242	40.9959	76.3249	2	797
neg	242	46.1942	92.8217	2	783

2. 基准回归

使用 stata18 进行股票收益率 Pct 滞后一期的基准回归，得到结果表 4。由于本文选取的是每日股票收益率，则会出现涨跌幅为 0 即某日股票收益率为 0 的情况，所以基准回归时，样本量从 242 将至 190。

通过基准回归结果表示，SENT 的回归系数为 1.1048 即投资者情绪每增加一个单位，股票收益率平均增加 1.1048 个百分点。且其在 1%的水平下显著，即核心解释变量投资者情绪 SENT 对股票收益率 Pct 有显著的正向影响。除印证了本文的假设 H2：投资者情绪对股票收益率波动有显著正向影响外，这与行为金融学的观点一致，投资者的乐观情绪往往推动股价上涨。再者，股票交易量 Vol 也在回归中呈现在 5%的水平下正向显著，意味着较高的交易量反映了市场的活跃程度，进而对股票收益率产生积极影响。换手率 Turn 系数为-12.1022，t 值为-1.85，表示换手率与股票收益率之间存在负相关关系，且在统计上接近显著水平。这可能意味着高换手率的股票可能面临较低的收益率，这可能由于频繁交易导致的价格波动或市场的不确定性增加。同时高换手率也反映投资者短期投机行为，增加市场波动性，对股票收益率产生负面影响。

Table 4. Baseline regression results
表 4. 基准回归结果

变量	Pct
SENT	1.1048*** (3)
Vol	1.9448** (1.95)
Size	1.7586 (1.1609)
Turn	-12.1022** (-1.85)
T_pct	0.0319 (0.42)
常数项	-60.9767*** (-2.52)
样本量	190
R ²	0.083

注：***、**、*分别代表在 1%、5%和 10%水平上显著；括号中的值为 t 值。

(三) 异质性检验

在股票市场中，市场环境可分牛市(即股价上升期)、熊市(即股价下跌期)、震荡三种。当市场环境不同时，投资者也会产生不同的情绪反映。为讨论在这三种环境下，投资者情绪对股票收益率是否会产生影响，本文对所选样本区间内，根据股票收盘价与道氏理论，即从不再创新低的点一直持续到不创新高的点作为一个完整的牛市；类似的，从不在创新高的点一直持续到不在创新低的点则是一个完整的熊市。通俗的来讲：只要每一个后续价位弹升比前一个弹升水平更高，而每一个次等回撤的低点(即股价自上至下的反转趋势)均高于上一个回撤，这一基本趋势就被认为是上升趋势，即为牛市。反之，每一个中等下跌都将价格水平逐渐压低，这种趋势则被认为是下降趋势，即为熊市。从中选出了样本区间内代表性上升期、下跌期、震荡期。由于时间设置为一年范围，难以分清股市的三个时期，所以将时间范围规定在 2023 年 12 月 27 日至 2025 年 1 月 3 日，以更好区分三个时期(见表 5 与图 2)。

Table 5. Rising, falling, and fluctuating time tables of the selected industrial and commercial bank of China in 2024

表 5. 所选工商银行 2024 年上升、下跌、震荡时间表

时期	时间
上升期	11 月 14 日~12 月 31 日
下跌期	8 月 28 日~9 月 20 日
震荡期	4 月 29 日~6 月 19 日



Figure 2. Division of rising, falling, and fluctuating areas of the industrial and commercial bank of China
图 2. 工商银行上升、下跌、震荡三个区域划分图

由回归结果(见表 6)可知,在股票上升期时,SENT 的回归系数为 0.3565 且在 1%水平下显著,即在牛市时,投资者情绪与股票收益率成正相关。根据过度自信理论分析,这可能是因为投资者可能高估自己的判断能力,低估风险。特别是在牛市中,这种心理倾向就会加剧市场的上涨趋势,以导致股票的收益率也不段上升。再根据羊群效应,牛市时投资者可能盲目跟随他人行为,忽视自身信息和判断。同时,在股票走势一直向好的时候,乐观情绪驱动投资者积极买入,推高股价,会形成一个正向循环,促使投资者情绪与股票收益率一起上升。再者当股市处于牛市态势下,持有积极情绪的投资者会提高未来的预期回报、低估风险,发生更多的交易行为[23]。

在股票下跌时,SENT 的回归系数为-0.1267 且仍然在 1%水平下显著,这表明在熊市时,投资者消极情绪可能会导致股票价格进一步下跌。在熊市中,投资者情绪的消极变化会加剧市场的恐慌情绪,导致投资者过度反应负面信息,引发更多的卖出行为,从而进一步压低股票价格。根据羊群效应理论,与牛市与之相反的是,羊群效应此时可能会导致集体抛售,从而加剧市场下跌。

在股票处于震荡期时,SENT 不显著,这可能是由于震荡期的变动幅度较小或维持在一个范围内,投资者情绪对股票收益率的影响微小。此时,市场可能更多地受到其他因素如宏观经济、公司基本面等的影响,削弱了投资者情绪的作用。综上,三种不同市场环境下,投资者情绪对股票收益率的影响也有所不同,由此佐证了本文提出的假设 H3:在不同的市场环境下,投资者情绪对股票收益率产生差异化影响。

Table 6. Regression results of rising, falling, and fluctuating periods
表 6. 上升、下跌与震荡期的回归结果

变量	上升	下跌	震荡
SENT	0.3565*** (2.09)	-0.1267*** (3.8)	0.5667 (1.52)
控制变量	Yes	Yes	Yes
样本量	34	16	34
R ²	0.3391	0.7846	0.5031

注：***、**、*分别代表在 1%、5%和 10%水平上显著；括号中的值为 t 值。

5. 结论与建议

(一) 结论

为了研究投资者情绪对股票收益率的影响，以工商银行为例，构建针对工商银行的转述投资者情绪指数，将探究具体化细节化。其中，通过 python 爬取东方财富网工商银行 2024 年所有评论，在 Bert-BiLSTM 模型训练微调结束后，调用该模型对爬取的所有文本进行分析与预测。然后采用基准回归的方式，研究投资者情绪与股票收益率的关系，并探讨了不同的市场环境下两者的关系。

结论：(1) 投资者情绪预测中，Bert-BiLSTM 模型比 Bert 模型、BiLSTM 模型更有效。即 Bert-BiLSTM 模型结合了 Bert 在自然语言处理中的强大语义理解能力和 BiLSTM 在处理序列数据中的优势，能够更准确地捕捉文本中的情绪信息；(2) 投资者情绪与工商银行股票收益率呈正相关关系，即投资者情绪越高，股票收益率越高。投资者情绪能够影响其投资决策，进而影响股票价格；(3) 不同市场环境下，投资者情绪对工商银行股票收益率有差异性影响。具体看来，当市场处于上升阶段，投资者的乐观情绪提升，预期未来回报增加，可能低估风险，进而增加交易活动，推动股价进一步上涨。市场下跌时，投资者情绪转为悲观，可能导致恐慌性抛售，加剧股价下跌，形成负反馈循环。震荡期市场波动较小，投资者情绪可能未形成一致预期，情绪对收益率的影响减弱。

(二) 建议

1. 对于监管机构

(1) 鼓励金融科技部门结合多种文本分析方法，如结合自然语言处理(NLP)中的情感分析、主题建模、语义分析等技术，创新投资者情绪的测算模型。例如，可以利用深度学习技术，如 BERT、GPT 等，来更准确地捕捉文本中的情绪信息。此外，还可以结合机器学习算法，如随机森林、支持向量机等，对情绪指标进行优化和校准，提高模型的预测准确性和鲁棒性。扩大数据范围，不仅包括社交媒体、新闻报道等传统数据源，还可以整合论坛、博客、评论等多渠道数据，丰富情绪分析的数据基础，提高模型的泛化能力，从而更有效地判断投资者情绪，及时发现和预警可能对市场造成巨大变动的风险。

(2) 根据不同股票市场的发展情况，采取差异化关注投资者情绪。实时关注市场变动，稳定投资者情绪。在市场整体表现较差，如处于熊市时，监管机构应加强对投资者情绪的监测和引导，通过发布积极的市场信息、举办投资者教育活动等方式，给予投资者信心，稳定其情绪，防止恐慌情绪蔓延导致市场进一步下跌。在市场处于牛市时，监管机构则应适当提醒投资者注意风险，通过发布风险提示、加强投资者教育等措施，引导投资者理性投资，避免过度乐观和羊群效应的产生，降低市场泡沫的风险。

2. 对于投资者

(1) 提升金融知识储备。投资者应不断学习和积累金融知识，了解市场的基本规律、投资工具的特点

以及风险与收益的关系。可以通过阅读金融书籍、参加金融培训课程、关注权威财经资讯等方式,提高自身的金融素养。在面对网络上的股票走势评论和信息时,保持批判的态度,学会甄别信息的真伪和可靠性。对于未经证实、来源不明或过于夸张的信息,要保持警惕,避免盲目跟风。坚持自己的投资决策,不轻易受到市场“热浪”的影响,克服羊群效应和过度自信的心理,做到理性投资。

(2) 抓住市场环境的契机。投资者应根据不同的市场环境,灵活调整自己的投资策略。例如,在年末市场收益率不断上涨或股市处于牛市时,市场整体向好,大部分股票的收益率也较好,此时可以适当减少更换股票的频率,即降低换手率。因为频繁交易会増加交易成本,而且在市场整体上涨的情况下,频繁换股可能会错过一些股票的上涨机会。投资者应更加关注长期的投资价值,选择优质股票进行持有,分享市场的整体上涨收益。同时,也要注意分散投资,降低风险,避免过度集中投资带来的风险。

(三) 研究不足与未来展望

本研究在构建基于股吧评论的高频情绪指标时,虽通过包含了 242 个观测点(按 2024 年全部交易日计算)的日度数据细粒度捕捉了投资者情绪的短期波动特征,已满足时间序列分析的基本要求。但受限于观测周期集中于 2024 年,可能尚未充分揭示情绪因子在不同市场周期中的长期传导规律。同时,由于观测周期较短可能导致异质性分析中选择的上涨期、下跌期与震荡期的代表性不充足。下一步计划将当前方法论扩展至更广泛样本区间与多只股票的运用中,进一步验证情绪指标对资本市场的时变影响效果,并考虑加入中介变量以研究影响路径。

参考文献

- [1] Sedor, L.M. (2002) An Explanation for Unintentional Optimism in Analysts' Earnings Forecasts. *The Accounting Review*, 77, 731-753. <https://doi.org/10.2308/accr.2002.77.4.731>
- [2] Jiang, Z., Peng, C. and Yan, H. (2024) Personality Differences and Investment Decision-making. *Journal of Financial Economics*, 153, Article ID: 103776. <https://doi.org/10.1016/j.jfineco.2023.103776>
- [3] 刘孟娟, 李刚. 基于微观数据的个人投资决策影响因素的实证分析[J]. 南京财经大学学报, 2015(5): 13-18, 108.
- [4] Bakar, S. and Yi, A.N.C. (2016) The Impact of Psychological Factors on Investors' Decision Making in Malaysian Stock Market: A Case of Klang Valley and Pahang. *Procedia Economics and Finance*, 35, 319-328. [https://doi.org/10.1016/s2212-5671\(16\)00040-x](https://doi.org/10.1016/s2212-5671(16)00040-x)
- [5] Fateye, T.B., Peiser, R.B. and Ajayi, C.A. (2024) Behavioral Factors Influencing Investment Decision-Making in the Nigerian Real Estate Stock Market. *Scientific African*, 24, e02238. <https://doi.org/10.1016/j.sciaf.2024.e02238>
- [6] Wysocki, P.D. (1999) Cheap Talk on the Web: The Determinants of Postings on Stock Message Boards. University of Michigan Business School Working Paper.
- [7] 蒋翠清, 梁坤, 丁勇, 等. 基于社会媒体的股票行为预测[J]. 中国管理科学, 2015, 23(1): 17-24.
- [8] Wang, Y., Huang, M., Zhu, X. and Zhao, L. (2016) Attention-Based LSTM for Aspect-Level Sentiment Classification. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, 1-5 November 2016, 606-615. <https://doi.org/10.18653/v1/d16-1058>
- [9] 张科, 李心丹, 方晓, 等. 股票网络论坛中的意见领袖: 慧眼识珠还是吸引眼球[J]. 管理科学学报, 2022, 25(9): 90-107.
- [10] 杨飘, 董文永. 基于 BERT 嵌入的中文命名实体识别方法[J]. 计算机工程, 2020, 46(4): 40-45, 52.
- [11] 王子牛, 姜猛, 高建瓴, 等. 基于 BERT 的中文命名实体识别方法[J]. 计算机科学, 2019, 46(S2): 138-142.
- [12] 孙弋, 梁兵涛. 基于 BERT 和多头注意力的中文命名实体识别方法[J]. 重庆邮电大学学报(自然科学版), 2023, 35(1): 110-118.
- [13] Antweiler, W. and Frank, M.Z. (2004) Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *The Journal of Finance*, 59, 1259-1294. <https://doi.org/10.1111/j.1540-6261.2004.00662.x>
- [14] 李瑞前, 许竞文, 徐光华. ESG 评级分歧对投资者情绪的影响研究[J]. 金融理论与实践, 2024(10): 91-103.
- [15] 陈志芳, 李佳员. 基于 BERT-BILSTM 模型的投资者情绪对股价短期预测[J]. 财经理论研究, 2024(6): 28-38.
- [16] 赵庆国, 曲晓宇. 投资者情绪对中国不同行业股票收益率的影响[J]. 会计之友, 2022(14): 54-60.

-
- [17] 部慧, 解峥, 李佳鸿, 等. 基于股评的投资者情绪对股票市场的影响[J]. 管理科学学报, 2018, 21(4): 86-101.
- [18] Di, L., Shaiban, M.S. and Hasanov, A.S. (2021) The Power of Investor Sentiment in Explaining Bank Stock Performance: Listed Conventional vs. Islamic Banks. *Pacific-Basin Finance Journal*, **66**, Article ID: 101509. <https://doi.org/10.1016/j.pacfin.2021.101509>
- [19] 柯春花. 基于 BERT-BiLSTM 文本情绪挖掘的 Black-Litterman 投资组合研究[D]: [硕士学位论文]. 太原: 山西财经大学, 2023.
- [20] 武广林. 投资者情绪对股票收益率的影响研究——基于对东方财富网顺丰控股评论的分析[J]. 运筹与模糊学, 2024, 14(2): 352-359.
- [21] 李瑜. 投资者情绪对我国股市收益率及波动性的影响研究[D]: [硕士学位论文]. 昆明: 云南财经大学, 2023.
- [22] 金德环, 李岩. 群体智慧: 同伴观点与价值发现——来自社交媒体的经验证据[J]. 经济管理, 2017, 39(12): 157-173.
- [23] 陈文慧. 投资者情绪对股票指数回报率影响的差异分析——基于沪深 300 指数的实证研究[J]. 金融, 2023, 13(2): 334.