

AI 大模型赋能金融市场量化投资？基于另类数据与传统金融数据的研究

何 勇^{1,2}, 焦 丽³, 杨 艺¹, 祝怡菲¹

(1. 山东大学金融研究院, 济南 250100; 2. 山东国家应用数学中心, 济南 250100;
3. 山东大学数学学院, 济南 250100)

摘 要 当前, 以 ChatGPT (chat generative pre-trained transformer) 为代表的大语言模型迅速发展, 被广泛用于股市投资算法交易、风险管理等多个领域. 这为金融投资者提供了新的决策工具和投资途径. 本文基于 BERT (bidirectional encoder representation from transformers) 模型和 ChatGPT 构建了适用于中国股票市场的投资交易模型, 实现从财经新闻文本数据以及传统金融数据中获取交易信号. 对于文本数据, 首先抓取每日的财经新闻将其与对应的股票代码相匹配. 其次将新闻文本数据输入至训练好的 FTBERT (fine-tuning BERT) 模型中, 得到每条新闻的情感倾向, 选择积极情感的财经新闻作为正的投资交易信号. 对于传统金融数据, 借助 ChatGPT 的高级解析能力, 对中国股票市场的历史数据进行深入分析. 通过调整 prompt 读取数据, 从而构造出用于股票投资的关键因子, 输出每日各股票的得分. 最终根据不同数据类型得到每日各股票的投资交易信号, 并将其作为构建投资组合的依据, 构建有效的投资策略. 实证结果表明, ChatGPT 能有效判断文本情感倾向, 且经过微调后的大语言模型能有效助力量化投资, 为投资者带来超额收益. 本研究尝试将大语言模型运用于金融投资领域, 展现了其在生成股票投资信号方面的潜在价值. 随着技术的不断发展和市场环境的变化, 这种基于人工智能的投资策略将不断演进, 为投资者创造更多价值.

关键词 中国股票市场; 财经新闻分析; 大语言模型; 量化投资

收稿日期: 2023-12-29

基金项目: 国家自然科学基金 (12171282)

Supported by National Natural Science Foundation of China (12171282)

作者简介: 通信作者: 何勇, 教授, 山东大学齐鲁青年学者, 研究方向: 概率统计、计量经济学, E-mail: heyong@sdu.edu.cn; 焦丽, 硕士研究生, 研究方向: 金融统计, E-mail: 202211956@mail.sdu.edu.cn; 杨艺, 硕士研究生, 研究方向: 应用统计, E-mail: 202332009@mail.sdu.edu.cn; 祝怡菲, 硕士研究生, 研究方向: 应用统计, E-mail: 202332022@mail.sdu.edu.cn.

AI Large Language Model Empowers Quantitative Investment in Financial Markets? Research Based on Alternative Data and Traditional Financial Data

HE Yong^{1,2}, JIAO Li³, YANG Yi¹, ZHU Yifei¹

(1. Zhongtai Securities Institute for Financial Studies, Shandong University, Jinan 250100, China; 2. National Center for Applied Mathematics, Shandong University, Jinan 250100, China; 3. School of Mathematics, Shandong University, Jinan 250100, China)

Abstract At present, chat generative pre-trained transformer (ChatGPT) as a representative of the rapid development of large language models, is widely used in stock market investment, algorithmic trading, risk management and other fields. This provides financial investors with new decision-making tools and investment paths. In this paper, we construct an investment trading model based on the bidirectional encoder representation from transformers (BERT) model and chat generative pre-trained transformer (ChatGPT) for the Chinese stock market, and realize the trading signals from financial news text data and traditional financial data. For the text data, the daily financial news is captured and matched with the corresponding stock codes. Secondly, we input the news text data into the trained fine-tuning BERT (FTBERT) model to get the sentiment tendency of each news item, and select the positive financial news as the positive investment trading signals. For the traditional financial data, we use the advanced parsing capability of chat generative pre-trained transformer (ChatGPT) to analyze the historical data of Chinese stock market. By adjusting the prompt to read the data, the key factors for stock investment are constructed, and the daily scores of each stock are output. Finally, the daily investment signals of each stock are obtained based on different data types, which are used as the basis for constructing investment portfolios and building effective investment strategies. The empirical results show that chat generative pre-trained transformer (ChatGPT) effectively determine the sentiment tendency of text. The fine-tuned model can effectively assist quantitative investment and bring investors excessive returns. This study attempts to apply big language modeling to financial investment and shows its potential value in generating stock investment signals. With the continuous development of technology and changes in the market environment, this artificial intelligence-based investment strategy will continue to evolve and create more value for investors.

Keywords China stock market; financial news analysis; large language models; quantitative investment

1 引言

近年来, 预训练基础模型技术作为新一代人工智能技术的主流探索方向, 其发展受到了学术界与工业界的广泛关注. 而在这场发展浪潮中, 大语言模型 (large language models, LLMs) 在自然语言处理领域所取得的成绩尤其瞩目 (赵朝阳等, 2023). 当前, 大语言模型在金融中的应用处于探索阶段, 学者们逐渐尝试将其应用于文本数据的情感分析. 新闻舆情作

为反映社会民意和市场预期的重要指标之一, 股票市场自然也会受到其影响。因此, 对于金融机构、投资者和研究人员来说, 通过使用大语言模型提取关键信息、处理大量新闻舆情, 能够节省时间和精力, 提高工作效率, 并为他们的决策提供更全面和准确的信息依据, 具有重要意义。

然而传统的文本情感分析方法主要基于情感词典和传统的机器学习方法, 容易受到语言表达方式的限制, 且无法捕捉上下文信息。相比之下, 以 BERT (bidirectional encoder representation from transformers) 和 GPT 为代表的大语言模型成为自然语言处理任务的重要基准模型。通过在大规模的无标签数据集上进行模型的预训练任务, 可以学习到通用的语言表示。这种表示能够捕捉到丰富的上下文信息, 并在多种下游任务中表现出色。现有研究中, 鲜有学者考虑使用 ChatGPT 来进行文本数据分析, 大部分学者仍采用 BERT 作为主要的分析模型, 但这往往需要手动标注数据。而 ChatGPT 的优势在于能够减少或消除手动标注的需求, 从而提高分析效率和减轻工作量。因此, 本文创新性地将 ChatGPT 与 BERT 模型组合, 提出 FTBERT (fine-tuning BERT) 模型。此模型是在熵简科技发布的中文版开源 FinBERT 模型基础上进一步微调得到, 将其应用到投资策略的构建中, 旨在探究 AI 大模型能否基于另类数据与传统金融数据构建出具有超额收益的投资策略, 赋能金融市场量化投资。

本文的研究主要分为两个步骤: 第一步是基于 FTBERT 模型的文本情绪分析。首先利用 GPT-4 判断新闻文本数据的情感倾向, 输出带情感标签的数据集。其次将带有情感标签的数据集输入 BERT 预训练模型, 然后进行微调以适应特定的下游任务。在微调过程中, 调整最终的输出层为二分类任务, 用于计算每篇新闻文本数据的情感倾向。最终将新闻文本数据中的积极情绪作为文本数据的投资交易信号。第二步是 ChatGPT 生成因子过程。本文考虑风险情绪类因子、估值与市值类因子等多种数据类型, 将其输入至代码解释器 (code interpreter) 让 GPT-4 读取数据。通过接收用户指令, 让 ChatGPT 生成一个独特的选股因子。根据选股因子获取股票交易信号, 随后做投资策略的回测, 基于回测结果判断构建的股票投资因子是否满足要求。如果回测结果优于大盘指数股 (中证 500), 则输出选股因子, 形成传统金融数据的股票交易信号。反之, 回到指令输入的部分重新生成因子。至此, 本文实现了从新闻文本数据和传统金融数据中提取股票交易信号。

本文研究的数据包含另类数据中的文本数据和传统的金融数据, 时间区间均为 2019 年 1 月至 2023 年 7 月。其中, 本文选择国内权威财经网站的新闻作为文本数据源, 经过数据清洗后的新闻总量为 119,738 篇。将清洗后的文本输入至 GPT-4 判断其文本的情感倾向。为了做对照, 本文将财经新闻情感分类公开数据集也输入至 GPT-4, 验证其判断效果。结果显示, 与常见的分类器相比, 无论是在训练集还是测试集上, GPT-4 在情感分类任务方面表现出色。为了验证 GPT-4 选择因子的可靠性, 本文采用因子有效性检验, 使用信息系数 IC 值与因子相关性判断因子的有效性, 按照高 IC 值和高相关性原则下筛选出的因子与 GPT-4 重合度较高, 说明 GPT-4 选择的因子具有一定的可靠性。在投资策略构建方面, 本文分别构建了文本数据交易策略、传统金融数据交易策略和“文本 + 传统”的复合交易策略。为了验证本文构建的投资策略是否有助于个体投资者的投资, 本文进行了一年期回测。结果显示, 测试期内基于文本数据和混合数据类型的投资策略表现均优于传统金融数据和仅投资大盘指数股 (中证 500)。特别地, 复合交易策略的投资效果 (年化收益率、累计对数收益率等) 显著优于单一策

略的投资,虽然最大回撤略高于文本数据构建的信号,但也小于 5%,属于低风险投资(在合理区间内),说明此投资策略有良好的稳健性和明确的投资价值。

本文的贡献主要包括两个方面。在数据方面,模型数据输入考虑了传统金融数据与另类数据,提高了模型的普适性。数据来源包括金融类文本、技术指标等金融数据,联合预测股票收益,为 A 股市场投资交易策略研究提供了新思路。本文创新应用股票匹配算法,节省计算机算力的同时,使数据处理更具有针对性,并且在训练中减少了主观因素的影响。在模型方面,本文提出的 FTBERT 模型,创新性地基于 GPT-4 框架的 ChatGPT 与 BERT 模型相结合,该组合模型可以通过处理新闻文本数据来构建投资交易模型,以更准确地捕捉市场情绪变化;同时本文初步尝试将 GPT-4 用于选股因子的构建,并将其与 FTBERT 模型输出的市场情绪结合构建投资策略,为投资决策提供参考。

后文安排如下:第二部分是文献综述,整理了新闻舆情与股市收益率、投资组合理论与机器学习以及大语言模型在金融中应用的相关研究成果;第三部分介绍基于 FTBERT 模型的文本情绪分析和基于 ChatGPT 的选股因子生成方法;第四部分为实验设计与实证分析,阐述本文的数据处理过程和交易信号提取的训练过程,根据交易信号构建股票投资策略,报告实验结果,并进行模型鲁棒性测试;最后是本文的研究结论。

2 文献梳理

2.1 新闻舆情与股市收益率影响的研究

近年来,随着人工智能技术的发展,大量学者开始研究新闻数据中的情绪信息,已有研究表明,投资者的情绪影响公司股价走势、股票收益等多方面(何诚颖等,2021)。在金融经济领域,已有学者论证了股票市场中的投资者情绪是系统性因子,它通过作用股票的交易策略,从而影响预期收益,因此是资产定价中的重要因素(陈荣达等,2019)。在技术层面上,文本情感分析技术快速发展,自 Tetlock (2007) 首先对市场中的新闻报道影响做出研究以来,人们已经意识到新闻媒体中的情绪信息会影响整体经济走势,越来越多的学者开始了对新闻舆情的研究,致力于挖掘媒体情绪与股市走向的关系。例如,在国内的相关研究中,游家兴和吴静(2012)发现,在公司信息不透明时,媒体情绪影响了股票价格偏离基本价值;汪昌云和武佳薇(2015)认为媒体语气会影响投资者的决策行为,并且媒体语气越强烈,对 IPO 抑价率和投资者情绪的影响越明显;许海川和周炜星(2018)基于股票市场的横截面分析,提出了新的情绪代理指标——中国波动率指数,研究情绪指标对未来股票市场的预测收益;杨晓兰等(2016)还研究了股吧评论数据中的媒体情绪对股市收益率的影响;部慧等(2018)利用文本挖掘技术,对股评情绪进行分类计算,提出新的投资者情绪指标。

大量报道证实了新闻舆情会作用于投资者情绪,进一步对股市收益情况造成影响。媒体在报道时往往会放弃中立立场,进而产生特有的“媒体情绪”(游家兴和吴静,2012)。姚尧之,王坚强和刘志峰(2018)发现混频情绪显著影响股市收益率与波动。Veldkamp (2006)所构建的理论模型表明,在经济繁荣时期,媒体的报道更加乐观,而乐观的新闻舆情造成了资产价格的抬升与收益波动的加剧;相反地,Tetlock (2007)构建了媒体悲观情绪指数,表明在股票大盘指数下跌时,悲观媒体情绪是其重要的影响因素。因此,当新闻舆情表现为乐观时,狂热的投资者主导了交易市场,股票价格在一定程度上反映了狂热投资者的意见,造成了股价高估

的现象, 当新闻舆情表现为悲观时, 结论与之相反, 在此不再赘述。总之, 研究新闻舆情与股市收益率之间的关系, 在金融市场中有十分重要的应用价值。

2.2 投资组合理论与机器学习应用研究

2.2.1 投资组合理论发展研究

如何合理配置现有金融资产, 在尽可能降低金融市场风险的同时实现最大化收益, 这一议题在金融领域始终备受关注。对于机构和散户投资者而言, 制定合理的投资组合战略尤为重要。Markowitz (1952) 提出均值-方差模型, 首次将数理统计方法用于投资组合, 正式拉开现代投资组合理论 (modern portfolio investment theory) 序幕。然而, 该模型存在许多局限性, 所考虑的市场环境因素过于理想化, 因此学者们不断提出各种改进方案。考虑影响投资预期收益率的因素, 部分学者不断探索, 如 Sharpe (1963) 提出资本资产定价模型 (capital asset pricing model, CAPM), Ross (1976) 提出套利定价理论 (arbitrage pricing theory, APT) 等。为解决资产收益率的分布通常具有非对称性这一问题, 学者们从偏度、峰度等角度入手, 如 Lai (1991) 研究了带偏度风险的投资组合问题。针对不同风险度量情况, 相关学者不断尝试, 先后提出了半方差、绝对偏差、半绝对偏差等方法 (Konno et al. (1991), Mansini et al. (1999))。20 世纪 80 年代, 风险管理领域出现了一种新的风险度量方法, 即风险价值 (value at risk, VAR) 模型。然而, 由于 VAR 存在缺陷, Rockafellar et al. (2000) 提出了改进的 CVAR 投资组合优化模型。随着统计思想与投资组合的不断结合, Black et al. (1992) 从贝叶斯理论出发, 构建了 black-litterman (BL) 模型。Sahamkhadam et al. (2022) 为进一步解决非正态收益数据问题, 在 Copula 模型基础上扩展了 BL 模型。基于多阶段投资组合问题, Li et al. (2000) 利用嵌入法, 提出离散时间的多期动态 MV 模型。张鹏 (2010) 为解决“维数灾难”问题, 提出了限制交易成本和交易量的多阶段投资组合策略。Sun et al. (2011) 借助元启发式算法, 将一种新的漂移粒子群优化 (DPSO) 方法用于多阶段投资组合, 优势性明显。经过不断发展, 投资组合的理论研究已经取得了丰富成果, 这些理论在实践中也被广泛应用。

2.2.2 机器学习在投资组合领域的应用研究

近年来, 随着数据处理技术的提升和算法的不断优化, 机器学习展现出卓越效果并持续渗透到各个领域。

在投资组合领域, 机器学习在股价预测、投资组合优化等方面展现出巨大潜力, 已成为许多金融机构和投资者的重要工具。基于此, 关于投资决策的制定和股票未来收益的预测, 现有的文献研究可大致分为以下三个方面: 一是以自回归移动平均模型为代表的传统计量模型。传统投资者在很大程度上依赖金融市场的历史数据, 并通过建立严格而复杂的计量经济学模型来积累投资决策的“经验”。二是以支持向量机 (SVM) 等为代表的经典机器学习模型。机器学习方法打破了传统方法成本过高等局限性, 实现了股票预测能力的优化, 能够进一步提升预测精度和稳定性 (林耀虎等, 2022)。许多学者对此进行了探索, 如蔡毅等 (2023) 提出一种反向混频数据抽样模型与机器学习算法结合的新模型, 用于模型预测与风险调整优化; 周亮 (2021) 发现随机森林模型相较传统模型实现了对个股相对收益率更好的拟合和预测。更进一步地, Leippold et al. (2022) 将机器学习应用于资产定价领域, 建立并分析了一套全面的因子组合以预测收益。三是以神经网络为核心的深度学习模型。由于深度神经网络在金融

数据预测建模方面取得成功,深度学习已成为一种前景广阔的金融预测方法(卜湛等, 2023; 苏治等, 2017)。深度学习在各种时间序列, 图像分类和自然语言处理问题上取得突破, 能够实现对金融复杂数据的捕捉与分析, 在金融领域前景广阔。在传统金融时序领域, 欧阳红兵等(2020) 将小波分析与 LSTM 神经网络相结合, 构建金融时间序列数据预测模型, 提升了对长短期动态趋势的预测精度。马甜等(2022) 基于时序数据, 采用生成式对抗网络模型(GAN), 显著提高了模型预测精度和因子投资绩效。近年来, 随着深度学习在图像识别领域的不断突破, 其在金融图像识别中的应用也越来越广泛。Jiang et al. (2023) 将股票历史价格转换为图像, 再借助卷积神经网络(CNN) 来识别和解析这些复杂模式, 实现市场动态预测能力的提升。然而, 金融信息复杂多样, 局限于传统金融数据已无法满足需要。基于深度学习的自然语言处理技术使得金融文本和多模态数据信息得以发掘和利用。金融投资者可以借助自然语言处理模型来分析文本情绪在投资决策中的重要作用(姜富伟等, 2021)。Gutiérrez-Fandiño et al. (2022) 结合自然语言处理模型提出了一种基于监督微调句子嵌入的金融情绪分析模型, 对投资决策的制定影响深远。石善冲等(2018) 基于自然语言处理模型利用微信公众号上的推文和评论文本, 结合上证指数的交易数据进行对比分析, 发现投资者情绪对股票市场有显著影响, 并且这种情绪能够在一定程度上预示市场风向。胡楠等(2021) 采用文本分析和自然语言处理技术研究管理者短视主义对企业长期投资的影响, 具有一定的现实意义。

机器学习算法应用广泛, 刘景江等(2023) 认为考虑如何将机器学习算法赋能金融及管理等多个领域意义深远。随着机器学习算法的不断发展, 机器学习与经济问题的结合改变了传统的经济学研究范式(蒋锋等, 2022)。因此, 将机器学习应用于金融投资领域, 进而实现投资优化具有重要意义。

2.3 大语言模型在金融中的应用现状及研究

大语言模型是一类基于深度学习的自然语言处理(natural language processing, NLP) 模型。通过在大量的文本数据上训练模型, 在语言理解、文本生成和情感分析等任务上表现出色(Carlini et al., 2021)。得益于 LLMs 复杂的网络结构和大规模的预训练数据, 研究者仅需少量标注数据即可微调模型, 实现优越的任务性能(王颖洁, 2022)。

传统经济学研究在数据获取、模型预测等方面面临诸多挑战。大语言模型的出现为经济学研究提供了全新的视角和方法, 为深入理解经济现象提供有力支持。洪永淼和汪寿阳(2024) 指出, 以 ChatGPT 为代表的大语言模型正在推动经济学研究范式的变革。大语言模型是大数据与人工智能催生的一种新的系统分析方法, 适合于研究复杂人类经济社会系统。在金融领域, 大语言模型的出现为非结构化数据、金融情绪分析等研究提供了新思路, 在股市投资、风险管理、金融知识传输与服务、客户咨询等方面应用前景广阔(汪寿阳等, 2023)。Peters et al. (2018) 利用大规模无监督语料, 使用双向 LSTM 训练语言模型, 得到 ELMo 模型, 并将其用于自然语言处理任务, 发现其在多项 NLP 任务中取得了显著的性能提升。此外, Transformer 凭借其在长期依赖性方面的优势迅速发展。在此背景下, GPT 预训练语言模型和 BERT 模型产生。Devlin et al. (2019) 提出了一种语言训练表示模型, 通过堆叠 Transformer 构建基础 BERT 模型, 解决了一词多义的问题。Liu et al. (2020) 在金融语料上训练 BERT 模型, 提出了 FinBERT 模型并发现模型在金融事件分类任务上的准确率提升

了 14 个点. 李旭晖等 (2021) 深入探讨金融事件抽取问题, 融合预训练模型与多层神经网络模型, 成功实现事件识别与要素抽取两大核心任务. 与此同时, 业界也开始注意到预训练语言模型的重要性. OpenAI 提出了基于 GPT 系列框架的大型生成式语言模型 ChatGPT (chat generative pre-trained transformer), 凭借其大算力、大算法和大数据, 赋予机器理解或学习人类可以完成的任何智力任务的能力. 已有学者将 ChatGPT 用于金融市场股票投资 (陆岷峰, 2023). Ko and Lee (2023) 通过实验发现, 与随机选择的资产组合相比, ChatGPT 构建的资产组合配置更具多样化, 在统计学上效果更为显著. ChatGPT 可以分析客户交互模式识别欺诈行为, 从而减轻金融投资者的损失和风险. Lopez-Lira et al. (2023) 使用 ChatGPT 来判断金融新闻的情感倾向并将其用于股市回报的预测. 结果表明, 将先进的大语言模型纳入投资决策过程中可以产生更准确的预测, 并增强量化交易策略的表现. 汪寿阳等 (2023) 总结了 ChatGPT 在金融领域的应用情况和发展前景, 认为以 ChatGPT 为代表的大语言模型可以辅助提高研究人员对金融语言的理解, 帮助金融投资者分析非结构化数据. 总之, 大语言模型在金融中的应用前景广阔, 有望为投资者、金融机构和决策者提供更准确、实时的信息和工具来辅助决策.

3 理论模型

3.1 基于 FTBERT 模型的文本情绪分析

目前文本特征提取的主流方法有 Word2Vec 与 GloVe 等基于神经网络的方法, 但其最明显的缺点在于无法解决一词多义的问题 (谌志群等, 2020). 而以 BERT 等为代表的预训练语言模型, 通过引入动态代表机制, 使同一词语的词向量随其在文本中的位置不同而动态调整, 可以更好地捕捉文本上下文关系, 从而解决一词多义等问题. 因此本文使用 BERT 模型作为基准模型处理财经新闻文本数据.

BERT 模型的结构框架如图 1 所示, 由输入层、编码层、输出层构成. 输入层用于将文本数据表示成计算机可读的嵌入式向量. 为了适应不同语义, BERT 模型的输入包括三个部分, 分别为词向量 (token embedding)、段向量 (segment embedding) 和位置向量 (position embedding), 如图 2 所示. 令 L_i 代表第 i 条文本的词数, 不包括 BERT 模型的特殊标记. 在文本通过 BERT 模型的输入层后, 会首先添加一个 [CLS] 标记在开头, 以及一个或多个 [SEP] 标记来分隔句子或标示文本的结束. 如果第 i 条文本由 S_i 个句子组成, 则在处理完成后得到的文本向量 $\mathbf{X}_i$ 的维度将为 $(L_i + S_i + 1) \times H$, 其中 H 是 BERT 模型隐藏层的维数. 编码层是由多个 Transformer 的 encoder 部分叠加组成, 核心结构是“自注意力机制 + 残差连接和层归一化 + 前馈神经网络”, 计算步骤如下: 第一步, 对 $\mathbf{X}_i$ 做线性映射, 得到 $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ 矩阵, 计算公式如 (1) 式所示, 其中 $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V$ 是权重矩阵.

$$\begin{cases} \mathbf{Q} = \mathbf{X}_i \mathbf{W}_Q, \\ \mathbf{K} = \mathbf{X}_i \mathbf{W}_K, \\ \mathbf{V} = \mathbf{X}_i \mathbf{W}_V. \end{cases} \quad (1)$$

第二步, 构建加权注意力矩阵, 获得第 i 条文本数据的全部向量的信息. 公式如 (2) 式所

示. 其中 d_k 是隐含层的维度.

$$\text{Attention}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \text{soft max} \left(\frac{\boldsymbol{Q}\boldsymbol{K}^T}{\sqrt{d_k}} \right).$$

(2)

第三步, 残差连接和层归一化. 经过注意力矩阵加权后与 $\boldsymbol{X}_i$ 相加得到残差连接, 然后做归一化处理, 可以加速模型的训练过程, 使其更快的收敛.

第四步, 经过前馈神经网络得到 $\boldsymbol{X}_i$ 的隐含层编码矩阵 $\boldsymbol{X}_{\text{hidden}}$.

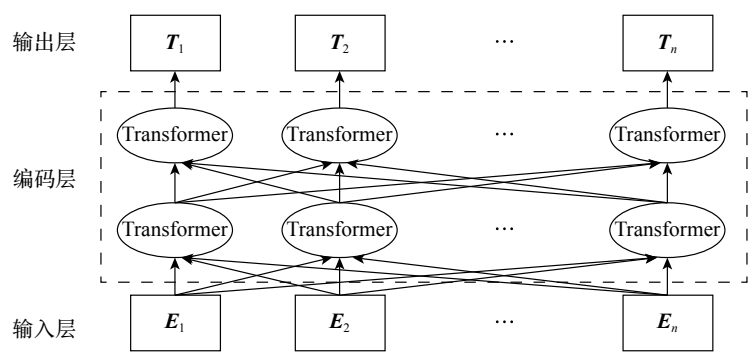


图 1 BERT 模型结构

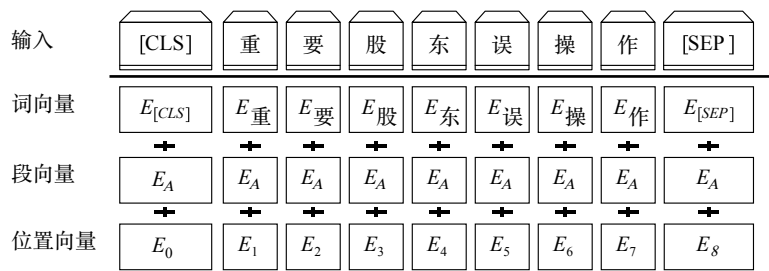


图 2 BERT 模型的输入

BERT 的输出层一般包括两类: 其一, Pooler Output. 源自 [CLS] 标记的隐藏状态, 该标记旨在捕捉整个输入序列的总体上下文. 此状态经过全连接层和激活函数 (如 $\tanh$) 处理后得到的输出, 主要用于需要整个序列概览的任务, 如文本分类. 其二, Sequence Output. 包括对序列中每个标记的隐藏状态输出, 为序列中每个元素提供了其上下文中的细节表示, 适用于标记任务, 如实体识别. 本文使用第一类输出作为基准 BERT 模型.

将 BERT 用于具体的文本分类任务一般遵循“预训练 + 微调”的结构范式. 预训练过程是在大规模文本数据集上进行模型的无监督学习, 通过 masked language model (MLM) 任务和 next sentence prediction (NSP) 任务学习文本语言的通用逻辑与结构. 在 MLM 任务中, 输入文本中的一些单词会被随机“遮蔽”, 然后通过上下文预测该单词, 从而学习输入文本的内在逻辑. 具体处理方法为: 一部分 (80%) 被遮盖, 替换为 [MASK]; 一部分 (10%) 随机替换为其他任意词; 一部分 (10%) 保留原词. 在 NSP 任务中, 模型获取一对句子, 并需要判

断第二句是否是第一句的下文, 即每个样本都是由 A 和 B 两句话构成, 如果句子 B 确实是句子 A 的下一句话, 样本标签为 Is Next; 反之, 如果句子 B 并非紧跟句子 A 而是随机选取的无关句子, 样本标签为 Not Next. 在样本集合中, 两种类型的样本占比均为 50%. 通过此任务, 模型能从上下文中推断句子间的逻辑关系. 微调过程中, 首先用预训练得到的参数来初始化模型, 然后利用针对特定任务的标注数据进行有监督学习. 期间模型的所有参数都会被调整和优化, 以提升其在特定任务上的表现.

本文以熵简科技发布的 FinBERT 模型为基础, 将 GPT-4 与 BERT 模型相结合, 构建了 FTBERT 模型做财经新闻情感分类任务. 首先让 GPT-4 判断每篇新闻文本的情感倾向, 将其作为 BERT 微调过程中的标签集. 然后将带情感标签的数据集输入至 BERT 预训练模型, 再进行模型微调, 调整输出层为二分类任务, 输出最优参数. 于是给定一篇新的新闻文本, 就可以使用该模型得到其情感倾向, 最终输出积极与消极情绪概率.

3.2 基于 ChatGPT 的选股因子生成过程

3.2.1 ChatGPT 的工作原理

截止至目前, OpenAI 陆续推出四代 GPT 模型. 最新推出的 ChatGPT 是基于 GPT-4 架构的大型预训练语言模型. GPT 系列采用 Transformer 深度学习网络架构, 可以利用自注意力 (self-attention) 机制来捕捉输入序列中的长距离依赖关系, 使得模型能够更好地理解文本语境并生成相关的回答.

其主题框架遵循“基础语料 + 预训练 + 微调”的理论范式, 并按照以下步骤实施: 第一步, 收集海量高质量的基础语料数据集. 这些数据集根据应用需求被细分为预训练数据集和微调数据集两大类. 其中, 预训练数据集汇聚了互联网上各类文本数据, 如书籍、杂志、新闻报道、维基百科文章和社交媒体帖子等, 旨在使模型掌握语言的逻辑结构和表达方式. 微调数据集包括了编程语言代码、专业标注以及用户对话等经过精细加工处理的高质量标注文本数据, 其目的是进一步提升模型人机交互对话能力. 第二步, 预训练模型. 采用无监督学习的方式在大规模的数据集上训练模型, 使模型深入理解并掌握从基础语法结构到复杂文本语境的各层次语言特征, 同时积累了丰富的通用知识和特定领域专业知识, 初步获得理解人类语言和学习上下文的能力. 第三步, 模型微调. 在带有标注的特定下游任务数据集上进一步训练模型, 根据任务的需求调整部分参数, 以使模型性能更加贴近所期望达到的目标. 值得注意的是, 根据不同任务类型的特点, 微调过程可以进行相应地修改和优化, 确保模型在各类任务中都能展现出优异的性能.

ChatGPT 是在 GPT 系列框架基础上, 将下游任务以提示 (prompt) 的形式融入到预训练模型中, 使模型具备零样本学习能力. 零样本学习是指模型在未接触特定任务样本数据的前提下, 依靠在大规模数据集上进行预训练所积累的丰富语言知识和理解能力, 实现对新型任务的有效应对和处理. 这一能力突显了模型在处理各类语言任务时展现出的泛化和推理能力, 反映了其对语言结构和语义信息深刻的理解与把握. 因此, ChatGPT 具备了生成问答式对话文本的能力, 其工作机制可以被归纳为“预训练 + 提示 + 预测”的新范式.

3.2.2 选股因子生成

借助 GPT-4 的强大模型架构, ChatGPT 在自然语言处理和人工智能交互对话方面取得了显著进展. 该模型在 GPT-4 模型基础上融入了提示学习、思维链以及人类反馈强化学习的核心理念, 展现出在处理复杂语言结构和理解深层语义方面的出色能力. 对于尚未接受模型训练的新任务, 用户只需要提供该任务的语言描述以及指令微调, 即可让模型在指定情境下输出恰当的回答结果, 这大大降低了 ChatGPT 的使用门槛. 在这一背景下, 提示工程(prompt engineering) 作为一门专注于优化语言模型提示机制的学科领域, 其重要性愈发凸显. 通过运用并精通提示工程的技术, 可以进一步提升系统的准确性和性能, 在降低成本的同时优化用户体验. 对研究人员来说, 这也有助于更深入地理解和改进模型的性能.

下面将描述如何利用提示词来生成选股因子, 具体方法如下:

第一步, 读入数据. 使用 GPT-4 的代码解释器功能, 将传统金融数据文件传输至 ChatGPT 的代码运行环境中.

第二步, 检查数据. ChatGPT 会自行编写 Python 代码, 浏览数据的结构以及内容, 检查数据是否存在缺失值, 是否存在数据频率不一致的问题.

第三步, 数据预处理. 根据输入数据的特点, 采用适用于此数据类型的填充方式处理缺失值, 并将数据频率调整一致, 这一过程由 ChatGPT 自行判断得到.

第四步, 生成选股因子. 向 ChatGPT 的对话框输入生成选股因子指令. 随后使用该选股因子进行投资策略的回测, 如果回测结果优于大盘指数股, 则输出选股因子, 形成股票交易信号. 反之, 回到指令输入的部分重新生成选股因子.

4 实验设计与实证分析

4.1 数据准备

本文借助财经数据接口包 Tushare 提取了“东方财富网”的新闻快讯, 将其作为本文的文本数据集; 从 Tushare 的“沪深股票”接口中提取了传统金融数据, 将其作为 GPT-4 读取的原始数据集¹.

4.1.1 文本数据及其处理

Tushare 是一个基于 Python 的免费、开源的财经数据接口包, 主要实现股票等金融数据的采集、清洗加工以及数据存储等任务, 能够为金融分析人员提供快速、整洁和多样的便于分析的数据. 本文主要使用 Tushare 的“新闻快讯”接口, 其数据来源为主流新闻网站的新闻快讯. 通过设置接口参数返回相应的国内财经新闻, 返回值包含新闻类别、新闻标题、发布时间、新闻链接、新闻内容等字段. 为了选取合适的财经新闻网站, 本文综合多个信息评价指标来评价网站, 具体如下:

1) Alexa 排名, 是一个全球排名系统, 通过计算过去三个月内网站的每日唯一身份访问者和网页浏览量的几何平均数得到, 每三个月更新一次全球网站排名.

¹ 本文所使用的数据和代码请参见科学数据银行 (ScienceDB) 期刊社区, DOI: 10.57760/sciencedb.j00214.00068 和 CSTR: 31253.11.sciencedb.j00214.00068. 若使用文中数据信息, 请注明引文和数据出处.

- 2) 百度权值, 主要通过网站的关键词排名来估算能给网站带来的流量, 用于评估网站的受欢迎程度, 其划分等级为 0~10.
- 3) PR 值, 全称为 Page Rank. 是 Google 用于评测一个网页“重要性”的一种方法, 其划分等级为 0~10. PR 值越高说明该网页越受欢迎.
- 4) 反向链接, 是其他网站指向该网站的链接数量, 代表该网站被其他网站认可, 从而提高网站的权重和排名.
- 5) 有效数据含量, 指清洗后的数据占原始数据的比例. 在 Tushare 接口下有四个主流财经新闻网站, 本文选择 2021 年 3 月 1 日至 2021 年 3 月 31 日的数据进行清洗, 计算有效数据含量.

四个网站的相应指标如表 1 所示. 表中显示“东方财富网”在 Alexa 排名和反向链接的数量上表现很好, Alexa 排名虽然稍逊于“新浪财经网”, 但优于“云财经”排名. 关于有用数据比例, “云财经”为 32%, 是四个平台中最好的一项, 但是它的 Alexa 排名和反向链接数量较差. 综合考虑, 本文最终选取“东方财富网”作为财经新闻文本的数据来源. 表 2 展示了接口提取的原始数据, 输出三个字段新闻发布时间、新闻标题以及新闻内容. 尽管接口开发人员已对网站数据做了初步整理, 但接口返回值仍存在新闻内容为空、新闻字段不完整等问题, 因此对数据清洗是很有必要的. 清洗过程分为两步: 第一步, 根据财经新闻标题及内容匹配相应的股票代码, 本文研究的股票代码范围为非金融行业的 A 股市场, 并且排除了股票名称中含有 ST (special treatment)、*ST、SST、S*ST 及 PT (particular transfer) 等关键字段的股票; 第二步, 删除匹配结果不唯一的新闻, 确保数据一致性和分析的准确性. 这意味着虽然一篇新闻仅对应一支股票, 但一支股票可以关联多篇新闻. 清洗后的数据示例如表 3 所示. 最终, 从“东方财富网”按日度频率提取 2019 年 1 月 1 日 00:00 至 2023 年 7 月 31 日 24:00 的财经新闻, 作为本文的文本数据源. 提取的数据总量为 615,830 篇, 清洗后为 119,738 篇.

表 1 权威财经网站综合排名指标					
	Alexa 排名	百度权值	PR 值	反向链接	有效数据含量
东方财富	172	8	7	58,041	27%
新浪财经	22	8	8	51,880	16%
云财经	73,487	5	1	1,630	32%
同花顺	992	8	6	32,359	26%

4.1.2 传统金融数据及其处理

对于传统金融数据, 本文利用 Tushare 的“沪深股票”接口得到 2019 年 1 月至 2023 年 7 月的日度数据, 包括风险情绪类因子、估值与市值类因子、盈利因子与现金流因子、基本面类因子和技术类指标因子五类因子数据, 各类数据分项因子如表 4 所示. 这类数据通常是数值型数据, 本文将这类数据输入至 ChatGPT 的代码解释器中, 由 GPT-4 进行数据预处理的工作. 预处理过程分为三步: 第一步, 加载五个数据集并查看其基本信息, 目的是了解数据的结构与分布; 第二步, 分析数据集中的缺失值, 并填充缺失值; 第三步, 消除量纲的影响, 将数据标准化, 使每个因子数据的均值为 0, 标准差为 1.

表 2 未经预处理的财经新闻文本数据示例

发布日期	新闻内容	新闻标题
2019/1/1 23:12:55	随着城市商业银行资金清算中心改制完成,我国清算机构服务供给的分工界限更加明确. 2018 年 12 月 29 日,央行发布消息称,城银清算服务有限责任公司成立大会在上海举行. 该公司是经央行批准设立的特许清算机构之一,将承接城商行清算中心业务资质,进一步畅通城商行的资金汇路. (北京商报)	城商行清算中心改制完成清算服务分工明确
2019/1/1 23:30:34	年终奖,这个本应喜气洋洋的词,近年来却成了一个伤心的话题. 往年传说中动辄十几万至几十万的金融业年终奖,仿佛就只是一个“传说”. 2018 年,金融业寒风瑟瑟,金融从业人员的年终奖是否会再降一个台阶? (中国证券网)	扫楼探秘: 银行保险证券年终奖能拿多少?
2019/1/1 23:46:44	2019 年 1 月 1 日,三聚环保 (300072) 公告,公司与粤财信托签署合同,将持有的部分应收账款转让给粤财信托,应收账款账面原值为 30 亿元,转让价款为 25.86 亿元. (中证网)	三聚环保转让面值 30 亿元应收账款 粤财信托以近 26 亿元承接

表 3 预处理后的金融新闻文本格式示例

发布日期	新闻内容及标题	股票代码
20190101	6 个要点帮你读懂个税变化: 超 1 亿人有望变 0 税一族 2019 年 1 月 1 日,新修改的《中华人民共和国个人所得税法》将正式实施. 新个税法中的个税起征点提高、六项专项附加扣除等规则的落地,减轻了纳税人的负担,也鼓了老百姓的钱袋子. 随着新个税法的实施,你的个税能少缴多少? 能不能享受专项附加扣除? 填报申请需要注意什么? 快收好本期个税“大礼包”. (新京报)	603883.SH
20190101	环保税开征一周年为绿色发展提供新动力 2018 年 1 月,我国首部环境保护税法正式施行,在全国范围对大气污染物、水污染物、固体废物和噪声等 4 大类污染物、共计 117 种主要污染因子进行征税. 作为我国第一个体现“绿色税制”的综合税种,环保税平稳实现了排污费“费改税”并顺利完成首年征收,为推进我国生态文明建设提供了全新动力. (新华网)	603888.SH
20190201	这家公司债台高筑! 两天谈妥的“白衣骑士”肠子都悔青了大股东几近爆仓! 索菱股份走到今天与其主业不振有关,在招股书中,公司曾表示“做具有国际领先技术的车载电子系统供应商”,但在手机导航高度进化后,车载导航日渐衰落. 本来,中山乐兴进来后,给了公司一次重生的机会,但目前的债务规模超出了其当初预料的 2、3 亿元,并吓住了这位“白衣骑士”. (证券时报)	002766.SZ

表 4 候选因子表

因子分类	因子名称
风险情绪类因子 (6 个)	成交量, 成交额, 换手率, 涨跌幅, 量比, 振幅
估值与市值类因子 (13 个)	市盈率, 市销率, 市净率, 股息率, 总股本, 流通股本, 自由流通股本, 总市值, 流通市值, 总资产, 流动资产, 固定资产, 公积金
盈利因子与现金流因子 (7 个)	每股收益, 未分配利润, 每股未分配利润, 收入同比增长率, 利润同比增长率, 毛利率, 净利润率
基本面类因子 (6 个)	收盘价, 昨日收盘价, 涨跌额, 开盘价, 最高价, 最低价
技术类指标因子 (5 个)	平均线指标, 相对强弱指标, 乖离率, 顺势指标, 布林线指标

4.2 实验设计

4.2.1 文本数据交易信号提取

本文基于 GPT-4 和 BERT 模型实现财经新闻文本情感分析, 分为两个部分: 情感标签集的制作和 BERT 模型微调. 在情感标签集的制作部分, 本文将财经新闻文本数据输入至 GPT-4, 由 GPT-4 判断其情感倾向, 得到带情感标签的数据集. 输入的文本数据集经由 Tushare 接口提取, 其中训练数据集时间跨度为 2019 年 1 月 1 日 00:00 至 2019 年 12 月 31 日 24:00, 共 14,912 篇新闻. 同时为精确评估 GPT-4 模型在判断财经新闻文本数据情感倾向 (积极或消极) 的效能与准确性, 本文将雪球网上的财经新闻情感分类公开数据集输入至 GPT-4. 此数据集共计 17,149 篇新闻数据, 包括日期、公司、代码、正/负面、标题、正文 6 个字段, 其中正面新闻 12,514 条, 负面新闻 4,635 条. 本节将此数据集的 10% 划分为验证集, 其余部分用于模型训练, 将 GPT-4 的预测标签与真实标签作比较. 同时将 GPT-4 与当今广泛使用的 SVM、CNN 和 XGBoost 这三种分类器对比, 结果如表 5 所示. TF-IDF 算法在文本挖掘中被广泛应用, 特别是在特征向量化方面, 它能有效地揭示文档中词语在语料库中的重要程度. 因此, 本节在使用 SVM、CNN 和 XGBoost 算法之前, 均通过 TF-IDF 算法来提取文本特征. 这一过程涉及计算词汇在文档中的频率 (TF) 及其在语料库中的逆文档频率 (IDF), 从而确定词汇的重要性. 高 TF-IDF 值的词语通常与强烈情感相关. 将这些词与情感词汇表对比, 可判断文本的正面或负面倾向. 上述方法综合考虑词频和语料库的独特性, 为情感分析提供全面视角. 实验对比发现, GPT-4 打分不论在训练集还是在测试集上的表现均优于 SVM 和 CNN. XGBoost 算法作为一种集成学习方法, 相比单一有监督学习器, 拥有计算速度快、泛化性能更好等高性能的特征, 而 GPT-4 的表现与其相当, 所以可以认为 GPT 的表现良好.

在 BERT 模型微调部分, 本文以熵简科技发布的面向金融领域的中文预训练语言模型 FinBERT 为基础, 进行模型参数的微调. FinBERT 模型是基于 BERT 模型架构, 采用了 12

表 5 GPT-4 与其他情感分类任务效果对比

指标	训练集				测试集			
	GPT-4	SVM	CNN	XGBoost	GPT-4	SVM	CNN	XGBoost
准确率	99.82%	86.15%	95.14%	98.20%	92.02%	82.88%	54.46%	96.81%
精确率	99.86%	82.71%	98.11%	97.19%	92.97%	80.44%	99.39%	97.53%
召回率	99.79%	95.67%	92.53%	99.45%	96.62%	92.89%	42.00%	98.21%
F ₁	99.83%	88.72%	95.24%	98.31%	94.76%	86.22%	59.05%	97.87%

层的 Transformer 结构, 隐藏层维度为 768, Multi-headAttention 的参数为 12, 预训练语料主要包含三大类金融领域的语料, 具体包括金融财经类新闻、研报与上市公司公告、金融类百科词条等. 本文的 FTBERT 模型在此基础上做了以下微调:

第一步, 确定数据集. 本文使用从 Tushare 接口提取的 2019 年度“东方财富网”的新闻快讯. 这些新闻文本数据经过 GPT-4 的情感分析处理, 生成一个带有情感标签的数据集. 本文保留 10% 的数据集作为验证集, 剩余文本数据作为训练集.

第二步, 参数训练. 为了提高模型训练速度, 选择使用 GPU P100 进行加速运算. 从 PyTorch 导入经过预训练的 BERT 模型, 将训练集输入至此模型做参数微调的工作. 在第 5 个训练周期结束后模型收敛. 最终模型训练得到最优参数, 学习率为 1×10^{-5} 、批量大小为 16、选用 AdamW 为优化器. 训练集的准确率稳定在 99.962%, 验证集的准确率稳定在 99.158%. 同时, 本文计算了训练集和测试集上的召回率, 分别稳定在 99.961% 和 99.145%. 至此本文完成了 BERT 模型的微调, 微调后的模型记为 FTBERT 模型. 将清洗后的文本数据输入至训练好的模型, 输出结果为积极情绪概率与消极情绪概率, 从而完成文本的情绪分类, 得到每日各股票的得分, 分别为 0 和 1. 本文将文本数据传递的积极情绪作为正的投资交易信号.

4.2.2 金融数据交易信号提取

本节将描述如何利用提示词生成选股因子, 建立可以从海量金融数据中获取交易信号的模型. 受 Lopez-Lira et al. (2023) 的启发, 本文将 GPT-4 的人机交互与信息读取的优势用于交易信号提取, 并将其用于中国股票市场投资. 首先将从 Tushare 接口提取到的金融数据输入至代码解释器 (code interpreter). 代码解释器, 是一个可以将 Python 代码写入 Jupyter notebook, 并在 Linux 沙盒中运行的 ChatGPT 模型. 当请求 ChatGPT 执行一项任务时, Code Interpreter 会首先解析这一任务并匹配数据分析方法, 然后生成相应的 Python 代码并在虚拟环境中执行. 同时, 它还能上传文件至临时磁盘空间, 从而让 GPT-4 有能力读取并处理多种类型的数据. 总之, 代码解释器是“GPT-4 + Python + Jupyter notebook”的结构框架. 然后向 ChatGPT 输入提问指令, 提问指令按照“角色 + 目标 + 背景 + 任务”的框架, 构造一个选股因子. 本文假定 ChatGPT 是一名资深的量化研究员, 拥有丰富的金融专业知识与中国 A 股市场的交易经验. ChatGPT 可以根据候选因子池 (如表 4 所示), 构造选股因子, 选股因子满足如下条件: 第一, 符合中国股票市场的市场运行规律, 不要随意杜撰捏造因子; 第二, 不要使用现有的金融指标直接作为选股因子, 要根据对中国股票市场的理解构造一个创新的因子; 第三, 不要使用单一的候选因子, 尽可能考虑多个维度的金融数据; 第四, 假定投资者是理性投资者, 没有特定的股票偏好与行业偏好. 本文经过多次调试指令, 最终指令如图 3 所示. GPT-4 根据指令构建综合性选股因子, 该因子考虑了风险情绪类因子、估值与市值类因子、盈利因子与现金流因子、基本面类因子以及技术指标因子多个维度的因子. 为进一步确保 GPT-4 回复的稳定性, 本文对股票因子池进行限制. 在多次试验后, 最终选取了 GPT-4 回复中高频出现的因子来构建综合性因子, 此因子由四部分构成, 分别如下:

1) 风险情绪与市场活跃度 (RSE): 市场情绪和股票的活跃度对股价的短期波动有显著影响. 活跃的交易和波动性往往表明投资者对股票的兴趣增加. 一般来说, 成交量、换手率、涨

LU YOU

从现在开始,假定你是一名资深的量化研究员,拥有多年中国A股市场的交易经验.我希望你可以根据我提供的候选因子表以及对中国股票市场的理解,手动构造出一个行之有效的选股因子,并给出详细的构建方式以及构建原因.要求做到以下几点:

- 1、符合金融市场逻辑,不要杜撰;
- 2、不要使用现有的指标直接作为选股因子,要有一定的创新;
- 3、尽可能考虑多个维度的数据;
- 4、假定投资者是理性投资者,没有特定的股票偏好.

候选因子表如下所示:

风险情绪类因子(6个)成交量,成交额,换手率,涨跌幅,量比,振幅;

估值与市值类因子(13个)市盈率,市销率,市净率,股息率,总股本,流通股本,自由流通股本,总市值,流通市值,总资产,流动资产,固定资产,公积金;

盈利因子与现金流因子(7个)每股收益,未分配利润,每股未分配利润,收入同比增长率,利润同比增长率,毛利率,净利润率;

基本面类因子(6个)收盘价,昨日收盘价,涨跌额,开盘价,最高价,最低价;

技术类指标因子(5个)平均线指标,相对强弱指标,乖离率,顺势指标,布林线指标.

图 3 选股因子指令输入界面

跌幅、量比、振幅这些指标能够从多个角度反映市场对股票的关注度和情绪变化.本文通过分析成交量、换手率等,获取市场活跃度和投资者情绪的信号.

2) 估值与市场价值 (VMV): 估值是投资决策的核心,合理的估值反映了市场对未来收益预期的共识,而低估或高估则可能指示未来股价调整的方向.本文使用市盈率、市净率、股息率这三个指标评估股票的价格合理性,可以帮助投资者判断股票是否被高估或低估.

3) 盈利能力与现金流 (PCF): 公司的盈利能力和现金流状况是其长期成功的关键.盈利能力强、现金流稳定的公司更有可能在市场中保持竞争力,为投资者带来持续的回报.本文使用每股收益、净利润率、收入同比增长率这些指标反映公司的盈利能力和成长性,以评估公司的基本面信息.

4) 技术动量 (TM): 技术分析是根据历史价格和交易量信息来预测股票未来走势的一种方法.动量指标如相对强弱指标和乖离率可以帮助投资者识别趋势的强度和潜在的反转信号,是短期或中期交易策略中常用的工具.

四个维度的得分加权结合成一个综合得分,作为传统金融数据的选股因子,其中权重根据历史数据的回测结果进行优化.综合得分的计算公式如 (3) 所示:

$$\text{Factor} = \omega_1 \times \text{RSE} + \omega_2 \times \text{VMV} + \omega_3 \times \text{PCF} + \omega_4 \times \text{TM}. \quad (3)$$

为验证 GPT-4 选择的因子是否可靠,本文做了因子有效性检验.当选择的因子与其对应的股票收益率有着高相关性,就认为该因子是有效的.一般来说,主要使用 IC 值与因子相关性判断因子的有效性,具体指标介绍如下:

IC 值即信息系数,指所选股票的因子值与股票下期收益率的截面相关系数,通过 IC 值可以衡量因子的股票预测能力和实际结果一致性.其值的计算主要有两种方式: normal IC 和 rank IC (秩相关系数).因为 normal IC 的计算有数据要服从正态分布的前提条件,然而实际市场并不满足这一条件.因此本文使用 rank IC 来判断因子的预测能力,计算公式如 (4)

所示. 其中 order_{t-1}^f 表示 $t-1$ 时期各股票的因子排名值, order_t^r 表示 t 时期各股票的收益率排名值.

$$\text{rank IC} = \text{corr}(\text{order}_{t-1}^f, \text{order}_t^r). \quad (4)$$

rank IC 的取值范围在 $[-1, 1]$ 之间, 其绝对值越大, 表明该因子在该期对股票收益的预测能力越强. 一般来说, IC 值大于 0.1 时, 就认为该因子具有很强的选股能力; IC 值大于 0.05 时, 就认为因子是有效的; IC 值接近 0 时, 因子被认为是无效因子.

因子间相关系数, 计算公式如 (5) 所示. 其中 $r(X, Y)$ 是因子 X 和 Y 间的相关系数, $\text{Cov}(X, Y)$ 是因子 X 与 Y 间的协方差, $\text{Var}(X)$ 是因子 X 的方差. $r(X, Y)$ 的取值在 $[-1, 1]$ 之间, 其绝对值越大, 说明因子 X 与 Y 的相关性越高.

$$r(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}. \quad (5)$$

按照 IC 值大于 0.02 且因子相关性不低于 0.9 的原则筛选因子, 最终筛选出 10 个因子, 分别是市盈率 (PE)、市销率 (PS)、市净率 (PB)、对数市值 (logMV)、流通市值 (circMV)、平均线指标 (MACD)、相对强弱指标 (RSI)、乖离率 (Bias)、顺势指标 (CCI)、随机线指标 (KDJ) 上述因子与 GPT-4 选择的因子有 4 个相重合, 分别是市盈率 (PE)、市净率 (PB)、相对强弱指标 (RSI)、乖离率 (Bias) 这说明 GPT-4 选择的因子有一定的可靠性.

4.3 实证结果分析

在得到财经新闻文本数据和传统金融数据的得分后, 根据其 与股票匹配结果, 分别得到对应股票的投资信号, 通过预测的投资信号建立股票投资组合并验证模型的有效性, 具体步骤如下:

步骤一: 设 t 为交易日, 收集 t 日开盘前一天中国股票市场相关的文本数据与传统金融数据, 并根据数据类型计算各自股票得分, 形成对应的股票交易信号.

步骤二: 分别对两种数据指定交易策略. 策略一: 对于文本数据, t 日投资情感倾向为正的股票; 策略二: 对于传统金融数据, 将股票按其得分从高到低依次排序, t 日等额投资交易信号为正且排名前 100 的股票. 策略三: t 日等额投资金融数据交易信号为正的前 100 支且财经新闻情绪为正的股票.

步骤三: 股票持有一天, 在 $t+1$ 日开盘后卖出, 重复步骤一和步骤二.

本文使用年化收益率、夏普比率、最大回撤、日均基点作为评价投资策略好坏的指标, 原因如下:

年化收益率: 指投资期限为 1 年所获得的收益率, 计算公式如 (6) 式所示. 其中 p_r 是年化收益率, p_{end} 是投资策略最终总资产, p_{start} 是投资策略初始总资产.

$$p_r = \left(\frac{p_{\text{end}} - p_{\text{start}}}{p_{\text{start}}} \right) \times \frac{365}{\text{投资天数}} \times 100\%. \quad (6)$$

夏普比率: 指承担一单位风险所获得的超额收益率, 其公式如 (7) 式所示. 夏普比率为正值, 代表标的报酬率高过波动风险; 夏普比率为负值, 代表标的波动风险高过报酬率. 对于投

资者而言, 这个比例越高, 投资组合越佳. (7) 式中 $E(R_p)$ 表示投资组合预期报酬率, R_f 表示无风险利率, σ_p 表示投资组合的标准差. 本文考虑无风险收益率为 3.35%.

$$\text{SharpRatio} = \frac{E(R_p) - R_f}{\sigma_p}.$$

(7)

最大回撤: 是指可能发生的最大亏损幅度, 其值等于策略收益曲线上当前高点到回溯期最低点的回撤幅度的最大值. 它是衡量投资策略风险的重要指标, 最大回撤数值越大, 说明投资者面临的风险越大.

日均基点: 1 个基点等于 0.01%, 用来表示金融工具的百分比变化.

本文计算了两种数据类型的交易信号, 并执行了对应的交易策略. 以 2021 年全年交易日为例, 共计 235 个交易日. 股票投资交易的结果总结在表 6. 对于文本数据, 本文构建的 FTBERT 模型的年化收益率为 21.145%, 夏普比率为 3.115, 最大回撤为 1.199%, 各项指标均优于每日仅投资中证 500 指数股和仅由 GPT-4 打分的结果. 对于传统金融数据, 由 GPT-4 构建的选股因子, 将其用于实际回测的年化收益率为 15.995%, 夏普比率为 0.826. 对于混合文本数据与传统金融数据类型下的交易策略, 根据 GPT-4 构建的选股因子, 每日投资得分最高的前 100 支且财经新闻情绪为正的股票, 结果显示, 市场情绪辅助因子可以显著提高投资策略的效果. 除此之外, 可以发现无论是 GPT-4 文本打分的结果还是构造因子的结果, 最大回撤值均较高, 超过了 10%, 这说明如果根据 GPT-4 的直接结果进行投资, 存在较高的风险性. 而本文将 GPT-4 的结果进行微调, 基于 FTBERT 模型的文本交易策略最大回撤仅有 1.199%、混合数据类型下的交易策略最大回撤是 2.855%, 均明显低于 GPT-4 的最大回撤值, 这说明本文构建的方法以及投资策略均具有优越性.

为了直观地展示本文投资策略相较于基准策略 (中证 500) 的优越性, 在图 4 中绘制了不同策略的累计对数收益率. 在一年测试期内, 基于 FTBERT 模型的文本交易策略和混合数据类型交易策略的累计收益率均优于中证 500 指数股. 对于文本数据交易策略, 其累计对数收益率随时间稳步增长, 对于混合数据类型指导下的交易策略, 其累计对数收益率随时间大幅增长, 远超其余策略, 因此该投资策略具有良好的稳定性和明确的投资价值.

表 6 投资策略判断指标

	策略	年化收益率	夏普比率	最大回撤	日均基点
文本数据	FTBERT 模型	21.145%	3.115	1.199%	8 bps
	GPT-4 打分	11.760%	0.479	14.045%	4 bps
传统金融数据	ChatGPT 构造因子	15.995%	0.826	11.153%	6 bps
文本 + 传统	情绪辅助因子	98.544%	8.328	2.855%	39 bps
基准 (ZZ500)		0.693%	-0.124	5.353%	< 1 bps

5 模型鲁棒性测试

5.1 不同年份对比

为评估本文构建的三种金融投资策略是否受特定时间和市场条件的影响, 本文分别获取不同年份的每日股票得分并执行股票投资策略的回测, 时间为 2020 年 1 月 1 日至 2023 年 7

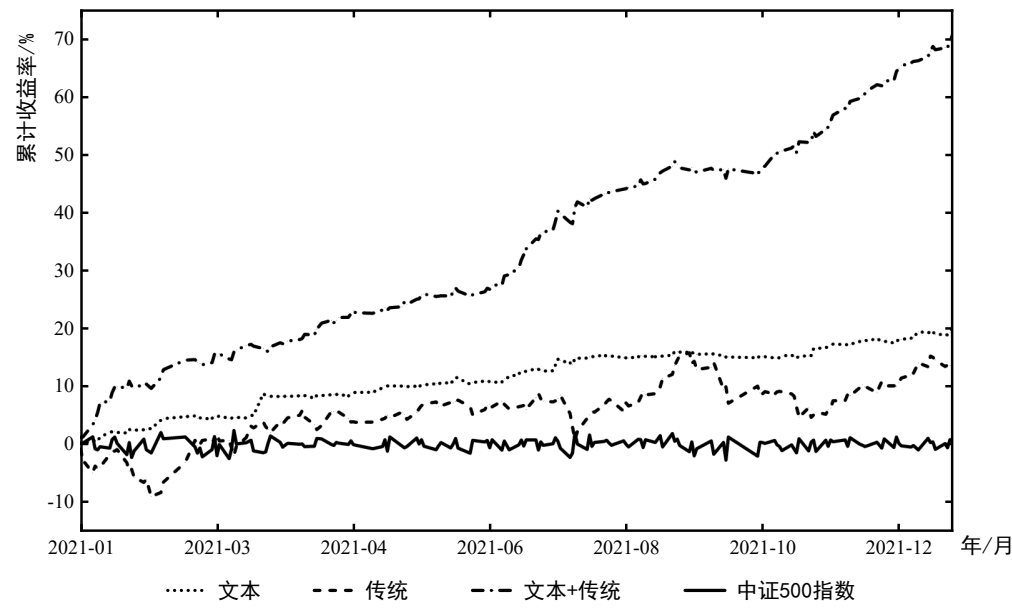


图 4 回测期内累计收益率

月 31 日. 投资回测收益率结果分别展示在表 7 至表 9. 关于文本数据, 虽然各年份年化收益率有差异, 但总体来说, 都远超中证 500 指数股. 充分说明本文构建的基于 FTBERT 模型的文本投资策略不受特定时间的影响.

关于传统金融数据, 经过多次调试指令, 本文构建了适用于传统金融数据的选股指标. 为了验证 GPT-4 构造的选股因子是否存在稳健性, 本文测试了不同年份的投资效果, 结果展示在表 8. 总体而言, 由 GPT-4 构建的因子各项指标优于大盘指数股 (中证 500), 年化收益率稳定在 10% 以上. 需要注意的是 2022 年四项指标表现不佳, 损失集中在第四季度, 可能受到后疫情的影响.

将文本数据与传统金融数据结合, 构建的投资策略结果如表 9 所示. 综合得分在年化收益率、夏普比率以及日基点上优于两种单一策略. 虽然最大回撤略高于 FTBERT, 但风险也控制在 5% 以内, 说明市场情绪的加入有助于捕捉投资交易信号, 能够从中国股票市场中获取股票收益.

表 7 文本数据在不同年份的投资策略结果

年份	年化收益率	夏普比率	最大回撤	日均基点
2020	9.644%	3.469	0.277%	3 bps
2020ZZ500	1.607%	-0.048	10.933%	< 1 bps
2021	21.145%	3.115	1.199%	8 bps
2021ZZ500	0.693%	-0.124	5.353%	< 1 bps
2022	4.238%	0.476	0.915%	1 bps
2022ZZ500	0.127%	-0.104	9.842%	< 1 bps
2023 上半年	9.330%	4.394	0.253%	3 bps
2023ZZ500	1.353%	-0.113	4.274%	< 1 bps

表 8 传统金融数据在不同年份的投资策略结果

年份	年化收益率	夏普比率	最大回撤	日均基点
2020	11.172%	0.298	15.963%	4 bps
2021	15.995%	0.826	11.153%	6 bps
2022	-1.404%	-0.201	22.006%	< 1 bps
2023 上半年	16.874%	1.065	6.227%	6 bps

表 9 综合得分在不同年份的投资策略结果

年份	年化收益率	夏普比率	最大回撤	日均基点
2020	87.783%	7.263	3.446%	35 bps
2021	98.544%	8.328	2.855%	39 bps
2022	62.590%	3.081	3.281%	25 bps
2023 上半年	44.553%	4.785	1.954%	18 bps

5.2 BERT 不同模型的对比

为了确定哪种模型更适用于金融新闻文本的分析, 本节将 FinBERT 模型与其他几种主流的 BERT 变体模型做对比研究, 研究不同模型如何影响数据处理的效果. 本文选用了精确率、准确率、召回率以及 F_1 值作为模型评估指标, 测试结果汇总在表 10 中. 实验数据显示, FinBERT 模型在整体性能上超越了其他比较模型. 值得指出的是, FinBERT 和 RoBERTa 两种模型均适用于处理中文文本数据. 尽管在准确率方面, RoBERTa 略微领先于 FinBERT, 但综合考虑所有评价指标, FinBERT 展现出更加卓越的表现. 此外, 这一发现也间接证实了数据类型对于 BERT 系列模型性能的重要影响, 尤其是在处理针对特定语言 (如中文) 优化的模型时尤为显著. 这些结果不仅展示了 FinBERT 在金融领域中文文本分析中的优势, 也强调了在选择适当的模型时考虑数据特性的重要性.

表 10 不同模型对分类任务的影响分析

模型	准确率	精确率	召回率	F_1
FinBERT	99.159%	99.083%	99.215%	99.146%
RoBERTa	99.051%	99.014%	99.057%	99.035%
DeBERTa	98.770%	98.717%	98.786%	98.751%
BERT	92.557%	92.372%	93.007%	92.505%

5.3 控制交易费用

本文构建的投资交易策略, 以日度频率调整仓位, 属于高频交易. 频繁交易会带来高交易费用, 直接影响净收益, 且对于长期投资效果和风险管理至关重要. 不同市场和交易所往往有自己的费率结构, 包括佣金、交易费和其他相关费用. 本文以 A 股市场为研究对象, 设定投资交易费率为 0.15%. 表 11 展示了控制交易费用对投资策略的影响. 扣除交易费用后, 2021 年的投资策略的年化收益率为 73.897%, 较扣除前降低了 24.647%. 夏普比率由 8.328 降至 6.199. 扣除交易费用前的最大回撤为 2.855%, 扣除后为 3.772%. 虽然交易费用一定程度上削减了部分策略收益, 但 74% 的年化收益率远远高于同期中证 500 指数收益率, 最大回撤在 5% 以内, 属于低风险投资, 证明本文构建的投资策略仍具有较高的参考价值.

表 11 控制交易费用对投资策略的影响

情况	年化收益率	夏普比率	最大回撤	日均基点
2020 年扣除前	87.783%	7.263	3.446%	35 bps
2020 年扣除后	72.203%	5.968	3.831%	29 bps
2021 年扣除前	98.544%	8.328	2.855%	39 bps
2021 年扣除后	73.897%	6.199	3.772%	29 bps
2022 年扣除前	62.590%	3.081	3.281%	25 bps
2022 年扣除后	32.312%	1.504	5.528%	13 bps
2023 年上半年扣除前	44.553%	4.785	1.954%	18 bps
2023 年上半年扣除后	23.343%	2.326	2.853%	9 bps

6 研究结论

本文实现了从财经新闻文本数据和传统金融数据中学习股票交易信号并构建投资策略,通过对中国股票市场的实际数据进行回测,以评估构建的投资策略是否有效. 综合实证研究结果,得出以下几点结论: 第一,本文发现 ChatGPT 能够克服 BERT 等预训练模型需要手动构建标签集的局限,从而显著降低了人工标注的成本. 第二,本文从模型角度出发,基于 GPT 的解码器与 BERT 的编码器,利用 ChatGPT 在人机交互方面的优势以及 BERT 在多重语义处理上的强大能力,构建了股票投资交易策略. 在扣除交易费用后,本文构建策略仍可以获得远超于中证 500 指数的超额收益率. 第三,本文将大语言模型用于金融市场量化投资,为金融科技人员提供了规避潜在风险的指南,并为金融机构在开发和训练自己的大语言模型方面提供了思路.

参 考 文 献

部慧, 解峥, 李佳鸿, 吴俊杰, (2018). 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, 21(4): 86–101.

Bu H, Xie Z, Lee J H, Wu J J, (2018). Investor Sentiment Extracted from Internet Stock Message Boards and Its Effect on Chinese Stock Market[J]. Journal of Management Sciences in China, 21(4): 86–101.

卜湛, 张善凡, 李雪延, 马丹丹, 曹杰, (2023). 基于深度强化学习的自适应股指预测研究 [J]. 管理科学学报, 26(4): 148–174.

Bu Z, Zhang S F, Lee X Y, Ma D D, Cao J, (2023). Adaptive Stock Index Prediction Based on Deep Reinforcement Learning[J]. Journal of Management Sciences in China, 26(4): 148–174.

蔡毅, 唐振鹏, 吴俊传, 杜晓旭, 陈凯杰, (2023). 基于灰狼优化的混频支持向量机在股指预测与投资决策中的应用研究 [J]. 中国管理科学, 31(7): 1–16.

Cai Y, Tang Z P, Wu J C, Du X X, Chen K J, (2023). Research on the Application of GWO-SVR Algorithm in the Prediction of Reverse Mixed Data in Stock Market and Investment Strategy[J]. Chinese Journal of Management Science, 31(7): 1–16.

陈荣达, 林博, 何诚颖, 金骋路, (2019). 互联网金融特征、投资者情绪与互联网理财产品回报 [J]. 经济研究, 54(7): 78–93.

Chen R D, Lin B, He C Y, Jin C L, (2019). Internet Finance Characteristics, Internet Finance Investor Sentiment and the Return of Internet Financial Products[J]. Economic Research Journal, 54(7): 78–93.

谌志群, 鞠婷, (2020). 基于 BERT 和双向 LSTM 的微博评论倾向性分析研究 [J]. 情报理论与实践, 43(8):

173-177.

Chen Z Q, Ju T, (2020). Research on Tendency Analysis of Microblog Comments Based on BERT and BLSTM[J]. Information Studies: Theory & Application, 43(8): 173-177.

何诚颖, 陈锐, 薛冰, 何牧原, (2021). 投资者情绪、有限套利与股价异象 [J]. 经济研究, 56(1): 58-73.

He C Y, Chen R, Xue B, He M Y, (2021). Investor Sentiment, Limited Arbitrage and Stock Price Anomalies[J]. Economic Research Journal, 56(1): 58-73.

洪永淼, 汪寿阳, (2024). ChatGPT 与大模型将对经济学研究范式产生什么影响?[J]. 计量经济学报, 4(1): 1-25.

Hong Y M, Wang S Y, (2024). How Will ChatGPT and Large Models Influence the Research Paradigm in Economics?[J]. China Journal of Econometrics, 4(1): 1-25.

胡楠, 薛付婧, 王昊楠, (2021). 管理者短视主义影响企业长期投资吗? —— 基于文本分析和机器学习 [J]. 管理世界, 37(5): 139-156.

Hu N, Xue F J, Wang H N, (2021). Does Managerial Myopia Affect Long-term Investment? Based on Text Analysis and Machine Learning[J]. Management World, 37(5): 139-156.

蒋锋, 张文雅, (2022). 机器学习方法在经济研究中的应用 [J]. 统计与决策, 38(4): 43-49.

Jiang F, Zhang W Y, (2022). Application of Machine Learning Methods in Economic Research[J]. Statistics and Decision, 38(4): 43-49.

姜富伟, 孟令超, 唐国豪, (2021). 媒体文本情绪与股票回报预测 [J]. 经济学 (季刊), 21(4): 1323-1344.

Jiang F W, Meng L C, Tang G H, (2021). Media Textual Sentiment and Chinese Stock Return Predictability[J]. China Economic Quarterly, 21(4): 1323-1344.

李旭晖, 程威, 唐小雅, 于滔, 陈壮, 等, (2021). 基于多层卷积神经网络的金融事件联合抽取方法 [J]. 图书情报工作, 65(24): 89-99.

Li X H, Cheng W, Tang X Y, Yu T, Chen Z, et al. (2021). A Joint Extraction Method of Financial Events Based on Multi-Layer Convolutional Neural Networks[J]. Library and Information Service, 65(24): 89-99.

林耀虎, 刘善存, 杨海军, (2022). 一种基于机器学习和蜡烛图的股市投资策略研究 [J]. 计量经济学报, 2(1): 126-140.

Lin Y H, Liu S C, Yang H J, (2022). A Novel Stock Investment Strategy Using Fusion of Machine Learning Techniques and Candlestick Charting[J]. China Journal of Econometrics, 2(1): 126-140.

刘景江, 郑畅然, 洪永淼, (2023). 机器学习如何赋能管理学研究?——国内外前沿综述和未来展望 [J]. 管理世界, 39(9): 191-216.

Liu J J, Zheng C R, Hong Y M, (2023). How Can Machine Learning Empower Management Research? A Domestic-foreign Frontier Review and Future Prospects[J]. Management World, 39(9): 191-216.

陆岷峰, 高伦, (2023). 大语言模型发展现状及其在金融领域的应用研究 [J]. 金融科技时代, 31(8): 32-38.

Lu M F, Gao L, (2023). Research on the Development Status and Application of Big Language Models in the Financial Field[J]. FinTech Time, 31(8): 32-38.

马甜, 姜富伟, 唐国豪, (2022). 深度学习与中国股票市场因子投资——基于生成式对抗网络方法 [J]. 经济学 (季刊), 22(3): 819-842.

Ma T, Jiang F W, Tang G H, (2022). Deep Learning and Factor Investing in Chinese Stock Market — Based on Generative Adversarial Networks[J]. China Economic Quarterly, 22(3): 819-842.

欧阳红兵, 黄亢, 闫洪举, (2020). 基于 LSTM 神经网络的金融时间序列预测 [J]. 中国管理科学, 28(4): 27-35.

Ouyang H B, Huang K, Yan H J, (2020). Prediction of Financial Time Series Based on LSTM Neural Network[J]. Chinese Journal of Management Science, 28(4): 27-35.

石善冲, 朱颖楠, 赵志刚, 康凯立, 熊熊, (2018). 基于微信文本挖掘的投资者情绪与股票市场表现 [J]. 系统工

- 程理论与实践, 38(6): 1404–1412.
- Shi S C, Zhu Y N, Zhao Z G, Kang K L, Xiong X, (2018). The Investor Sentiment Mined from WeChat Text and Stock Market Performance[J]. Systems Engineering — Theory & Practice, 38(6): 1404–1412.
- 苏治, 卢曼, 李德轩, (2017). 深度学习的金融实证应用: 动态、贡献与展望 [J]. 金融研究, (5): 111–126.
- Su Z, Lu M, Lee D X, (2017). Deep Learning in Financial Empirical Applications: Dynamics, Contributions and Prospects[J]. Journal of Financial Research, (5): 111–126.
- 汪昌云, 武佳薇, (2015). 媒体语气、投资者情绪与 IPO 定价 [J]. 金融研究, (9): 174–189.
- Wang C Y, Wu J W, (2015). Media Tone, Investor Sentiment and IPO Pricing[J]. Journal of Financial Research, (9): 174–189.
- 汪寿阳, 李明琛, 杨昆, 林文灿, 姜尚荣, 等, (2023). ChatGPT + 金融: 八个值得关注的研究方向与问题 [J]. 管理评论, 35(4): 3–11.
- Wang S Y, Li M C, Yang K, Lin W C, Jiang S R, et al. (2023). ChatGPT + Finance: Eight Noteworthy Research Directions and Issues[J]. Management Review, 35(4): 3–11.
- 王颖洁, 朱久祺, 汪祖民, 白凤波, 弓箭, (2022). 自然语言处理在文本情感分析领域应用综述 [J]. 计算机应用, 42(4): 1011–1020.
- Wang Y J, Zhu J Q, Wang Z M, Bai F B, Gong J, (2022). Review of Applications of Natural Language Processing in Text Sentiment Analysis[J]. Journal of Computer Applications, 42(4): 1011–1020.
- 许海川, 周炜星, (2018). 情绪指数与市场收益: 纳入中国波指 (iVX) 的分析 [J]. 管理科学学报, 21(1): 88–96.
- Xu H C, Zhou W X, (2018). Sentiment Index and Market Return Considering the iVX[J]. Journal of Management Sciences in China, 21(1): 88–96.
- 杨晓兰, 沈翰彬, 祝宇, (2016). 本地偏好、投资者情绪与股票收益率: 来自网络论坛的经验证据 [J]. 金融研究, (12): 143–158.
- Yang X L, Shen H B, Zhu Y, (2016). The Effect of Local Bias in Investor Attention and Investor Sentiment on Stock Markets: Evidence from Online Forum[J]. Journal of Financial Research, (12): 143–158.
- 姚尧之, 王坚强, 刘志峰, (2018). 混频投资者情绪与股票价格行为 [J]. 管理科学学报, 21(2): 104–113.
- Yao Y Z, Wang J Q, Liu Z F, (2018). Mixed-frequency Investor Sentiment and Stock Price Behavior[J]. Journal of Management Sciences in China, 21(2): 104–113.
- 游家兴, 吴静, (2012). 沉默的螺旋: 媒体情绪与资产误定价 [J]. 经济研究, 47(7): 141–152.
- You J X, Wu J, (2012). Spiral of Silence: Media Sentiment and the Asset Mispricing[J]. Economic Research Journal, 47(7): 141–152.
- 张鹏, (2010). 均值——动态方差多阶段投资组合优化研究 [J]. 统计与决策, (6): 67–68.
- Zhang P, (2010). A Research of Mean-dynamic Variance Multi-stage Portfolio Optimization[J]. Statistics and Decision, (6): 67–68.
- 赵朝阳, 朱贵波, 王金桥, (2023). ChatGPT 给语言大模型带来的启示和多模态大模型新的发展思路 [J]. 数据分析与知识发现, 7(3): 26–35.
- Zhao C Y, Zhu G B, Wang J Q, (2023). The Inspiration Brought by ChatGPT to LLM and the New Development Ideas of Multi-modal Large Model[J]. Data Analysis and Knowledge Discovery, 7(3): 26–35.
- 周亮, (2021). 基于随机森林模型的股票多因子投资研究 [J]. 金融理论与实践, (7): 97–103.
- Zhou L, (2021). Research on Stock Multi-Factor Investment Based on Random Forest Model[J]. Financial Theory and Practice, (7): 97–103.
- Black F, Litterman R, (1992). Global Portfolio Optimization[J]. Financial Analysts Journal, 48(5): 28–43.

- Carlini N, Tramèr F, Wallace E, Jagielski M, Herbert-Voss A, et al. (2021). Extracting Training Data from Large Language Models[C]// 30th USENIX Security Symposium (USENIX Security 21): 181–198.
- Devlin J, Chang M W, Lee K, Toutanova K, (2019). 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: 4171–4186.
- Gutiérrez-Fandiño A, Alonso M N I, Kolm P, Armengol-Estapé J, (2022). FinEAS: Financial Embedding Analysis of Sentiment[J]. The Journal of Financial Data Science, (3): 45–53.
- Jiang J W, Kelly B, Xiu D C, (2023). (Re-)Imag(in)ing Price Trends[J]. The Journal of Finance, 78(6): 3193–3249.
- Ko H, Lee J, (2023). Can ChatGPT Improve Investment Decision? From a Portfolio Management Perspective[J]. Finance Research Letters, 2024: 105433.
- Konno H, Yamazaki H, (1991). Mean-Absolute Deviation Portfolio Optimization Models and Its Applications to Tokyo Stock Market[J]. Management Science, 5: 519–531.
- Lai T Y, (1991). Portfolio Selection with Skewness: A Multiple Objective Approach[J]. Review of Quantitative Finance and Accounting, 1: 293–305.
- Leippold M, Wang Q, Zhou W, (2022). Machine Learning in the Chinese Stock Market[J]. Journal of Financial Economics, 145(2): 64–82.
- Li D, Ng W L, (2000). Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation[J]. Mathematical Finance, 10(3): 387–406.
- Liu Z, Huang D, Huang K, Li Z, Zhao J, (2020). FinBERT: A Pre-trained Financial Language Representation Model for Financial Text Mining[C]// International Joint Conference on Artificial Intelligence: 4513–4519.
- Lopez-Lira A, Tang Y, (2023). Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models[J]. Available at SSRN. <https://ssrn.com/abstract=4412788>.
- Mansini R, Speranza M G, (1999). Heuristic Algorithms for the Portfolio Selection Problem with Minimum Transaction lots[J]. European Journal of Operational Research, 114(2): 219–233.
- Markowitz H M, (1952). Portfolio Selection[J]. The Journal of Finance, 7(1): 77–91.
- Peters M E, Neumann M, Iyyer M, Gardner M, Clark C, et al. (2018). Deep Contextualized Word Representations[C]// Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: 2227–2237.
- Rockafellar T, Uryasev S, (2000). Optimization of Conditional Value-at-Risk[J]. Journal of Risk, 2: 21–41.
- Ross S A, (1976). The Arbitrage Theory of Capital Asset Pricing[J]. Journal of Economic Theory, 11(3): 341–360.
- Sahamkhadam M, Stephan A, Ostermark R, (2022). Copula-based Black-Litterman Portfolio Optimization[J]. European Journal of Operational Research, 297(3): 1055–1070.
- Sharpe W F, (1963). A Simplified Model for Portfolio Analysis[J]. Management Science, 9(2): 277–293.
- Sun J, Fang W, Wu X, (2011). Solving the Multi-stage Portfolio Optimization Problem with a Novel Particle Swarm Optimization[J]. Expert Systems with Applications, 38(6): 6727–6735.
- Tetlock P C, (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market[J]. Journal of Finance, 62(3): 1139–1168.
- Veldkamp L L, (2006). Media Frenzies in Markets for Financial Information[J]. American Economic Review, 96(3): 577–601.