

尊敬的各位老师:

大家好!我是商学院商务分析专业的付舟行,我的毕业论文题目是《基于自然语言处理的投资者情绪对股市行情的影响研究》。下面我将简要介绍本论文的主要内容。

随着互联网技术的快速发展,社交媒体平台上的投资者言论成为了市场情绪的重要载体。投资者情绪作为一种非理性因素,对资产定价具有不可忽视的影响。本研究旨在探究网络平台上的投资者言论所蕴含的情绪信息是否能有效预测股票股票收益率变动,并分析不同行业间的异质性特征。

在研究方法上,本文采用了多学科交叉的研究范式。首先,我利用 Python 开发网络爬虫从东方财富网股吧获取了 8 个行业 111 只股票在 2025 年 2 月的评论数据,共计 15 万余条帖子和 22 万余条评论。然后,应用 StructBERT 和 RoBERTa 深度学习模型对文本进行情感分析,构建了包括情感得分、情感极性和情感强度等在内的多维情绪指标体系。通过 akshare 接口获取股票市场交易数据,将情绪指标与股价数据匹配构建面板数据集。最后,运用计量经济学方法,通过面板数据固定效应模型分析情绪指标对股票收益率的影响。

研究发现包括三个主要结论:第一,投资者情绪对股票未来一日收益率具有显著的正向预测作用,实证结果表明,平均情绪得分每提高一个百分点,未来一日股票收益率平均提高约 0.014-0.017 个百分点。第二,情绪的预测作用存在明显的行业异质性,对计算机、银行、医药生物和房地产等行业的影响更为显著,而对电子、食品饮料等行业影响较弱。第三,情绪的影响具有明显的时效性,主要体现在短期(一日),随着时间延长逐渐减弱甚至出现反转,符合"情绪过度反应-修正"理论。

本研究的创新点主要体现在两个方面:一是技术方法层面,采用了预训练语言模型对金融文本进行深度语义理解,提高了对投资者评论情感的识别准确性;二是研究视角层面,创新性地构建了多维情绪指标体系,全面刻画投资者情绪的不同特征。

本研究的结果对投资者、监管机构和上市公司均具有实践指导意义。投资者可将情绪指标作为短期交易决策的重要参考;监管机构可构建情绪监测系统,及时识别市场异常;上市公司则应重视舆情管理,维护良好的投资者关系。

限于数据和方法,本研究仍存在一些局限性。未来研究可从扩展样本期、整合多源数据、改进情感分析模型以及探究情绪与公司基本面的交互作用等方面进行深化。

感谢各位老师的聆听,欢迎批评指正!