

FinDPO: 通过大型语言模型的偏好优化进行算法交易的金融情感分析

Giorgos Iacovides
giorgos.iacovides20@imperial.ac.uk
Imperial College London
London, UK

Wuyang Zhou
wuyang.zhou19@imperial.ac.uk
Imperial College London
London, UK

Danilo Mandic
d.mandic@imperial.ac.uk
Imperial College London
London, UK

Abstract

在线金融相关文本数据中表达的观点对交易决策和市场动态产生了越来越深远的影响。这一趋势凸显了情感分析作为量化此类观点的性质和强度的工具的重要作用。随着生成式 AI (GenAI) 的快速发展, 经过监督微调 (SFT) 的大型语言模型 (LLMs) 已成为金融情感分析的实际标准。然而, SFT 模式可能导致对训练数据的记忆, 并且往往无法推广到未见样本。在金融领域中, 这一限制尤为关键, 因为模型必须适应以前未观察到的事件和金融的微妙、特定领域的语言。为此, 我们引入了 FinDPO, 这是第一个基于直接偏好优化 (DPO) 的人类偏好对齐后训练的金融专用 LLM 框架。所提议的 FinDPO 在标准情感分类基准测试上实现了最先进的表现, 平均超出现有的监督微调模型 11%。独特的是, FinDPO 框架通过一种新颖的 “logit-to-score” 转换, 使微调的因果 LLM 能够集成到实际投资组合策略中, 将离散的情感预测转化为连续的、可排名的情感得分 (概率)。通过这种方式, 模拟表明 FinDPO 是第一个保持可观年收益 67% 和强劲的风险调整表现的基于情感的方法, 如夏普比率为 2.0 所示, 即使在 5 个基点 (bps) 的实际交易成本下也是如此。

Keywords

Large language models, sentiment analysis, direct preference optimization, algorithmic trading, portfolio construction.

1 介绍

财务情感分析是指对未标记文本数据中的意见进行量化, 旨在分类总体观点是积极的、消极的还是中立的, 这被称为价态。鉴于与财务情感相关的文本信息量迅速增加及其易获取的特性, 这对金融市场产生重大影响是毋庸置疑的。这再加上算法交易在量化金融中的关键作用, 突出了对基于如此庞大模态数据流训练的可靠且可操作的人工智能模型的需求。情感分析特别关注生成式人工智能 (GenAI), 因为它能够识别来自新闻文章、收益电话、财务报告及其他非数字化数据的情感。生成式人工智能处理大规模信息来源的能力使得捕捉宏观趋势成为可能, 承诺在股票价格预测 [13] 和开发有效且稳健的交易策略 [19] 上具备竞争优势。尽管具有不可否认的概念优势, 金融文本多样、细致入微且特定于领域的性质在准确和可操作的情感提取方面带来了重大挑战。这突显了情境感知情感分析的必要性, 并强调了在金融应用中使用自然语言处理 (NLP) 的非平凡障碍。大语言模型 (LLMs) 已经成为应对此类挑战的有力工具, 尤其是在通过后期训练技术在标注的金融数据集上进行微调时。确实, 监督微调 (SFT) 方法, 特别是指令微调, 利用 LLMs 理解和生成类似人类文本的非凡能力, 已成为增强预训练 LLMs 在金融情感分类任务中表现的事实标准。尽管最初取得了成功, 最近的研究表明, SFT 范式可能导致训练数据的记忆化, 同时难以推广到未见样本。这一限制在金融领域尤为关键, 因为能够推广到先前未见事件对于

强大的算法交易策略至关重要。它也与人类金融分析师的行为形成鲜明对比, 人类分析师能够从有限的先前信息中进行推断, 并应用领域知识来评估新且意外的市场状况下的情感。为了解决这些问题, 我们着手回答: 为此, 我们引入了第一个用于金融情感分析的后训练人类偏好对齐技术, 基于直接偏好优化 (DPO)。这使我们能够利用人类对齐方法, 这些方法在其他自然语言处理任务 (如翻译和摘要) 中已显示出对分布外输入的改进泛化能力, 在金融领域中进行应用。特别是, 我们的解决方案称为 FinDPO, 将人类偏好优化与预训练的大型语言模型 (具体为 Llama-3-8B Instruct) 结合, 用于金融情感分析, 从而能够在一个小规模的专业、带标签且公开可用的金融新闻数据集上进行有效训练。通过这种方式, FinDPO 通过参数高效的 DPO 实现了增强金融情感分析的最终目标, 同时也将计算资源需求最小化。

因此, 这项工作的主要贡献是:

大型语言模型中的后训练技术。尽管在大量文本语料上训练的 LLM 具有显著的通用能力, 但其在下游任务上的表现可以通过后训练微调技术显著提升。后训练方法通常分为两类: i) 监督微调 (SFT), 包括指令微调, 它使用标记的输入-输出对来调整预训练的 LLM, 以及 ii) 偏好优化。尽管有成功, 指令微调在捕获细微用户偏好方面仍然有限。因此, 引入了人类偏好优化, 最初通过从人类反馈中进行强化学习 (RLHF)。这种方法通常涉及首先在一组人类偏好数据集上训练奖励模型, 然后使用强化学习算法微调语言模型以最大化该奖励, 最常用的是 REINFORCE、近端策略优化 (PPO) 或其变体。虽然 RLHF 与指令微调相比展现了更优越的性能, 但由于奖励建模的挑战和 RL 优化过程的复杂性, 它往往计算成本高且不稳定。为了缓解这些挑战, 直接偏好优化 (DPO) 被提出作为一种更简单、无 RL 以及更稳定的从偏好数据训练语言模型的替代方案。重要的是, 尽管 DPO 隐式地优化与 RLHF 相同的目标 (在 KL 散度约束下最大化奖励), 但它通过一种直接且更易处理的方式实现这一目标, 保持实现的简单性和训练的稳定性。

采用大型语言模型的情感分析。基于变压器的模型用于金融情感分析的开始由 FinBERT [1] 开启, FinBERT 是一种在金融文本上微调的 BERT 变体, 在金融领域的情感分类任务中显示了良好的表现。然而, FinBERT 存在一些限制, 例如对数值的不敏感, 并且由于其相对较小的规模 (1.1 亿参数), 其分类准确性在句子复杂度增加时会降低 [4]。最近, Instruct-FinGPT [18] 和 FinGPT [15] 模型通过采用指令微调来增强模型性能, 分别利用 Llama-7B 和 Llama-2-13B 模型作为其基础模型。虽然这些模型代表了一种向更强大且更具普遍适用性的大型语言模型的转变, 但 FinGPT 并没有专门针对金融情感分析进行优化。此外, 这两种模型都仅限于预测情感极性 (即正面、负面或中性), 但缺乏量化情感类别强度的能力, 而情感强度是投资组合构建的一个关键参数。与之相比, FinLlama 结合了 SFT 与 Llama-2-7B 模型, 并在 LLM 的输出端引入了一个分类

头，以产生连续的情感分数。虽然这种修改使得情感信号能够直接集成到投资组合构建工作流程中，但也将模型的主要目标从下一个标记预测转移到分类上。这限制了依赖于语言模型生成能力的更高级后训练技术的适用性。为了解决这些限制，我们提出了 FinDPO，这是第一个基于后训练人类偏好对齐而非 SFT 的金融特定情感分析 LLM。我们模型的动机是偏好优化在性能和泛化能力上相对于 SFT 表现优越，尤其是在涉及捕捉微妙人类偏好且以细致入微的语言表达的任务中，这一特性也是我们金融情感分类任务的基础。给定数据集，DPO 的目标是最小化

$$\mathcal{L}_{\text{DPO}} = -\mathbb{E}_{(\tilde{x}, y^w, y^l)} \left[\log \sigma \left(\beta \cdot \left(\log \frac{\pi_{\theta}(y^w|\tilde{x})}{\pi_{\text{ref}}(y^w|\tilde{x})} - \log \frac{\pi_{\theta}(y^l|\tilde{x})}{\pi_{\text{ref}}(y^l|\tilde{x})} \right) \right) \right], \quad (1)$$

其中 σ 代表非线性 sigmoid 函数， β 控制与基础参考策略 π_{ref} 的偏离，即初始参考 LLM。

DPO 的典型工作流程概述如下：

- (1) 对于每个提示， $\tilde{x}$ ，采样完成项 $y_1, y_2 \sim \pi_{\text{ref}}(\cdot|\tilde{x})$ 。
- (2) 收集表示更偏好 y_1 或 y_2 的人类偏好标签，从而生成偏好对的离线数据集 $\mathcal{D} = \{(\tilde{x}_i, y_i^w, y_i^l)\}_{i=1}^N$ 。
- (3) 优化策略（可训练）模型 π_{θ} ，以最小化 DPO 损失，给定参考（冻结）模型 π_{ref} 、数据集 $\mathcal{D}$ 和所需的超参数 β 。

从直观上看，DPO 更新策略模型 π_{θ} 的参数，以增加首选响应 y^w 的可能性，并降低次选响应 y^l 的可能性。这些更新的幅度由 π_{θ} 与人类偏好不一致的程度调节，并由 β 进行缩放。这种动态加权稳定了训练过程，防止模型崩溃。

我们的工作旨在利用通用大型语言模型的表达能力和上下文理解，并将其调整适应于情感分析应用。通过应用人类偏好对齐，使用 DPO，对 Llama-3-8B-Instruct 进行调整，这个过程指定了我们的 π_{ref} 模型，并在特定的金融语料库上进行对齐。我们从两方面评估我们方法的有效性。首先，我们使用标准的分类为中心的基准，诸如加权 F1 分数，来评估模型相对于其他特定于金融领域的大型语言模型的表现。为了严谨起见，我们的模型进一步通过一组与现实世界的投资组合构建紧密相关的基准进行评估——这是情感分析的最终目标。

预训练的大型语言模型提供了一系列能力，比如推理、翻译、总结和文本生成，然而，当直接应用于特定领域任务如金融情感分析时，通常会遇到困难。这种局限性在金融领域尤为明显，那里复杂的语言、投机性的叙述和金融新闻文章的篇幅构成了对通用模型的显著挑战。为了应对这些挑战，我们的工作重新审视大型语言模型的第一原则，并利用人类偏好对齐来调整它们以进行金融情感分析。更具体地说，我们利用 DPO，它直接作用于从真值情感标签和模型预测构建的成对偏好上，使模型能够学习金融情感的细粒度区分，同时保留其生成能力。通过这个过程，所提出的 FinDPO 模型生成三个类别的情感预测：正面、负面或中性。

1.0.1 训练数据集。 训练数据由三个公开可用的标注金融新闻数据集组成，来自 HuggingFace: Financial PhraseBank (FPB) 数据集 [12]，Twitter Financial News 数据集 [16]，以及 GPT 标注的 Financial News 数据集 [11]。总共得到 32,970 个标注样本，其中 80% 用于训练，其余 20% 用于测试。以下是数据集的总结。

- 金融短语库 (FPB) 数据集。FPB 数据集由 4,840 个样本组成，这些样本是从财经新闻文章中随机提取的。为了确保高质量的标注，这些样本由具有金融和商业背

景的 16 名专家进行标注。每个样本都被标注为三种标签之一：正面、负面和中性。

- Twitter 金融新闻情感 (TFNS)。TFNS 数据集包含 11930 条来自金融领域的推文。每条推文都被标注为正面、负面或中性。
- GPT 标记的金融新闻 (NWGI)。NWGI 数据集包括 16,200 篇金融新闻文章。每篇文章都标注了以下五个标签之一：强烈负面、轻微负面、中性、轻微正面和强烈正面。为了使这个数据集与其他数据集的三类结构保持一致，强烈和轻微负面的分类被合并为一个负面类别，同样地，强烈和轻微正面的类别被合并为一个正面类别。

由于 DPO 需要偏好对而不是类别标签，我们将这三个数据集转换为合成的成对偏好数据。每个原始文本输入 x_i 首先被转换为与指令调优模型兼容的指令格式 $\tilde{x}_i$ 。对于每个样本，我们将优选的响应 y_i^w 设置为真实情感标签。为了获得不优选的响应 y_i^l ，我们使用 $\tilde{x}_i$ 提示参考模型 π_{ref} 生成情感预测。如果预测与真实情感标签匹配，我们随机抽取一个不同的错误标签作为 y_i^l 以避免偏见选择不优选响应。如果预测错误，我们使用预测的标签作为 y_i^l ，以便引导模型远离其自身错误并强化正确的情感预测。这产生了一个偏好对数据集 $\mathcal{D} = \{(\tilde{x}_i, y_i^w, y_i^l)\}_{i=1}^N$ ，用于 DPO 训练。

1.0.2 模型训练。 我们提出的 FinDPO 模型首先用 Llama-3-8B-Instruct 模型进行了初始化，该模型作为基础参考策略 π_{ref} ，随后进行了 5 轮的 DPO 对齐。训练过程中采用了 AdamW 优化器 [9]，因为它有效地将权重衰减与优化步骤解耦，从而实现更稳定的收敛。初始学习率被故意设置得较小，因为 Llama-3 模型已经在大量数据集上进行了预训练，同时预热比例和权重衰减作为关键的正则化技术，用于防止过拟合，这一点在我们有限的 DPO 训练数据集大小下尤为重要。

此外，低秩适应 (LoRA) [5] 被整合到 DPO 训练中，采用秩 $r = 16$ 、缩放因子 $\alpha = 16$ ，以及 0.05 的 dropout 率，以便在保持高端性能的同时最小化可训练参数的数量。通过 LoRA 的实现，可训练参数数量设定为 41.9M，仅占基础模型总参数的 0.52%。这使得我们的训练过程可以在单块 A100 (40 GB) GPU 上实现，从而避免了过多计算资源的需求。整个训练过程在 A100 (40 GB) GPU 上用 4.5 小时完成。

整个训练流程，包括将原始训练数据转换为偏好对数据集 $\mathcal{D}$ ，以及基础参考策略的后续 DPO 对齐 π_{ref} ，在 $\mathcal{D}$ 上产生我们的 FinDPO 模型，详见图 ??。

1.1 情感驱动的投资组合构建框架

一旦我们的 FinDPO 模型建立起来，我们就按照图 1 所示的框架进行。该框架的目标是使用金融特定的真实指标评估 FinDPO 相对于其他已建立的情感分析方法的表现，提供一个更实用的模型效用评估，以便构建投资组合。

数据收集。 为了支持适当的多空 (L/S) 投资组合的构建，收集了文本和市场数据。对于文本数据，从 2015 年 2 月到 2021 年 6 月期间，从包括路透社、The Motley Fool 和 MarketWatch 在内的权威在线来源收集了 204,017 篇金融新闻文章。这些来源因其可靠性、声誉、低偏倚以及对主要公司的关注而被选择。在相同期间，相应的金融市场数据是从雅虎财经获取的，包括我们投资宇宙 (SP&P500) 中 500 家公司的每日股票回报数据。这导致每家公司获得 1,672 天的股票回报数据。

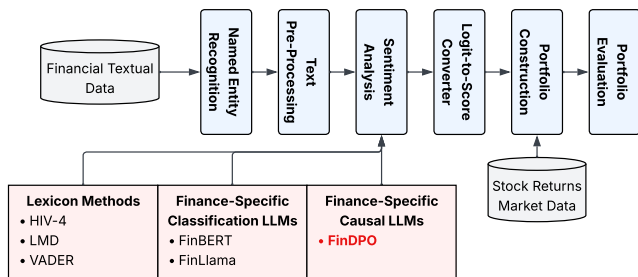


Figure 1: 我们情感驱动的投资组合构建的建议框架。‘logit-to-score’ 转换器仅对于金融特定的因果 LLMs 是必需的。

实体识别。为了确保新闻文章能够准确关联到正确的组织实体，每篇文章必须至少与一个相关的股票 [3] 相关联。此步骤减少了将不相关的文章连接到特定股票 [2] 的可能性。为了增强鲁棒性，我们的研究采用了 BERT-base-NER 模型 [8]，该模型能够识别四种类型的实体：位置、组织、人员和杂项，并为每个识别出的实体提供一个置信度评分。对于语料库中的每篇文章，如果关联的实体的公司置信度评分超过 98 %，则保留该文章；否则，将其丢弃。表格 1 显示了应用 NER 过滤到初始抓取的数据集时在保留文章数量方面的结果。观察到该过程使文章总数减少了 24.1 %。对于 MarketWatch 和 The Motley Fool，效果类似，皆减少了大约 25 %，而路透则经历了大约 6.3 % 的小幅降低。

Table 1: 在 NER 过滤前后每个来源获得的文章总数，括号中显示了减少的百分比。

News Source	No. of articles pre processing	No. of articles post NER filtering
MarketWatch	309,187	236,214 (23.6 %)
Reuters	38,141	35,741 (6.3 %)
The Motley Fool	205,270	147,413 (28.2 %)
Total	552,598	419,368 (24.1 %)

文本预处理：每篇新闻文章都使用词袋方法表示，随后执行以下步骤：a) 分词，b) 停用词移除，c) 词形还原，d) (小写) 大小写规范化，e) 特征选择。在特征选择步骤中，每个词的出现频率被用作主要特征，特别是在基于词典的方法中。情感分析。总共使用了五种情感分析方法作为基准，与我们的 FinDPO 模型进行比较。这些方法包括基于词典的方法和 SFT 分类 LLMs，其中包括现有的 SOTA FinLlama [7]。在基于词典的方法中，使用 pysentiment2 Python 库实现了 LMD [10] 和 HIV-4 [14]，而使用 NLTK 库实现了 VADER [6]。对于 SFT 分类模型，我们评估了通过 HuggingFace 获得并通过 Transformers 库使用的 FinBERT [1] 和 FinLlama [7]。

尽管在标准分类聚焦的基准测试中表现最强（如表 ?? 所示），但经过微调的因果 LLMs 在本质上仅限于生成离散的情感标签（即，正面、负面、中性），并未捕捉情感强度。然而，理解情感的表达强度对于投资组合构建至关重要，因为在投资组合中，资产必须根据情感的幅度进行排名，而不仅仅是简单分类为某些类别。为了克服这一限制，展示使用因果 LLMs 进行投资组合构建的可行性，我们在 FinDPO 框架中提出了一种新的 ‘logit-to-score’ 转换器，它利用模型的内部表示实现情感强度排名。与我们的 FinDPO 模型相关，先前的研究表明，偏好对齐的大型语言模型（LLMs）可能表现出过度自信的现象。我们在我们的模型中也观察到了这一效应，其中它始终将预测

的情感类别的概率分配为 1.0，而其他两个情感类别的概率为 0.0。为了缓解这一现象并生成更有意义的概率情感分数，我们采用了温度缩放（一种事后校准方法）。温度参数是在训练集上通过最小化负对数似然（NLL）进行优化的。此校准是独立于用于投资组合构建的财经文章语料库进行的，以确保没有数据泄漏。所考虑的方法在每个公司的语料库内的每篇文章上进行了评估。如果在同一天为某个公司发布了多篇文章，则计算该天的平均情感值。这里，表示第 t 天的平均情感值，表示在同一个第 t 天为某个公司发布的新闻文章数量，而表示特定第 t 天的第 i 篇新闻文章的情感强度。对每个公司的每日情感输出进行了合并，以得出作为投资组合构建阶段参数的最终情感数据。

投资组合构建。一旦为每家公司定义了每种方法的情感值，就构建了多空投资组合。我们使用情感作为参数来确定哪些公司应处于多头或空头头寸，以最大化这两种头寸的回报。多空投资组合是通过以下程序构建的：投资组合评估。使用所提模型构建的投资组合的表现与使用其他情绪方法构建的投资组合进行对比评估。我们考虑了：

- 收益率通过累计回报率， r_{cum} ，与年化回报率， R_p ，定义为

$$r_{cum} = \sum_{i=1}^N r_{daily}(i) \tag{2}$$

$$R_p = \frac{1}{N} \sum_{i=1}^N \eta_{log}(i) \cdot 252 \tag{3}$$

- 风险调整后的绩效，通过年化 Sharpe, S_a ，Sortino, S_o ，和 Calmar, C_r ，比率定义如下

$$S_a = \frac{R_p - R_f}{\sigma_p} \tag{4}$$

$$S_o = \frac{\bar{r}_{simple} - R_f}{\sigma_d} \cdot \sqrt{252} \tag{5}$$

$$C_r = \frac{(1 + \bar{r}_{simple})^{252} - 1}{MDD} \tag{6}$$

在这里， N 是交易日总数，共计 1,672 天， $\eta_{log}(i)$ 表示每日对数收益率， $\bar{r}_{simple}$ 表示平均每日简单收益率。符号 R_f 代表年化无风险收益率， σ_p 是年化波动率， σ_d 是每日收益的下行偏差，而 MDD 是最大回撤。常数 252 对应于日历年中的工作日数。无风险收益率 R_f 通常代表 10 年期国债收益率；然而，由于分析期间其长期低收益率 [17] 的影响，通常采用了 % 为 0 的收益率，并在我们的分析中采用了这一标准。

2 实验结果

我们通过机器学习和财务业绩评估验证了我们提出的 FinDPO 框架的有效性。机器学习评估侧重于标准情绪分类基准以衡量预测性能。财务评估则通过将其输出整合到算法交易策略中，在理想化和现实市场条件下分析投资组合表现来检验 FinDPO 的实际应用效用。我们首先评估了我们的 FinDPO 模型在用于训练的三个数据集的测试拆分上的性能，这些数据集被认为是金融情感分析的标准基准，使用加权 F1 分数。我们将模型与传统的基于词典的方法，以及 FinBERT 和 Instruct-FinGPT 进行了对比。此外，我们还将 FinDPO 与金融情感分析领域的最先进模型进行比较，即 FinLlama 和 FinGPT v3.3。为了进一步评估 DPO 相对于指令调优的有效性——后者是改进通用 LLMs

中金融情感分类的最突出方法（如 FinGPT v3.3 所示）——我们使用相同的基础模型，将 FinDPO 与其指令调优的对应版本进行比较，我们称之为 FinSFT。表 ?? 显示，FinDPO 在所有三个基准数据集上都取得了最高的平均性能，具有 0.846 的加权 F1 分数，比金融情感分类中最先进的模型 FinGPT v3.3 高出 11%。值得注意的是，FinSFT 的性能与 FinGPT v3.3 相当，尽管基于较新的 Llama-3-8B 模型而不是 Llama-2-13B，这表明指令调优的模型已经达到了性能的上限。这一比较强调，FinDPO 的优越性能并不是由于更强大的基础模型，而是因为使用了我们特定于情感的 DPO 框架，该框架比指令调优在金融情感分类的背景中更有效地使模型输出与真实标签对齐。

Remark 1. 引入的 DPO 不仅鼓励大语言模型增加首选情感标签的可能性，还明确惩罚不受青睐标签的选择，与单纯优化正确预测的 SFT 方法相比，这导致其对未见样本具有更好的泛化能力。

2.1 通过真实世界的金融指标进行评估

我们接下来评估了根据第 1.1 节描述构建的六个投资组合的表现，而未考虑交易成本。这些结果在图 2 的左侧面板和表 2 中进行了说明。请注意，我们提出的方法 FinDPO 显著优于现有的基于情感的投资组合构建领域的最先进模型 FinLlama，实现了累积收益 187% 的提升。关键是，FinDPO 还展示了更高的风险调整表现，夏普比率增加了 1.1，表明更为有利的收益与波动率权衡。此外，FinDPO 达到了最高的 Sortino 和 Calmar 比率（分别为 6.05 和 11.94），反映了改善的下行风险控制和更强的回报与回撤关系，这两者都是风险规避投资者的关键考量项。

需要注意的是，在这种未考虑交易成本的理想评估环境中，包括词典法在内的所有方法均表现得相对较强。这突出了现有许多文献的关键局限性，其中基于情感的投资组合策略在无摩擦假设下进行评估。这种环境可能导致过于乐观的结论，而在现实交易环境中，交易成本可能会显著影响表现。

Table 2: 基于情感的投资组合在无交易成本情况下的绩效指标。对于每个指标，最佳绩效用粗体显示，次佳绩效用下划线标出。

Method	Cumulative Return (%)	Annualized Return (%)	Sharpe	Sortino	Calmar
S & P 500	83.12	11.34	0.62	0.81	0.41
HIV-4	90.07	12.88	0.81	1.25	0.67
VADER	82.81	11.76	0.75	1.21	0.34
LMD	139.88	20.62	1.26	1.96	1.17
FinBERT	199.19	29.64	1.65	2.39	1.24
FinLlama	<u>260.74</u>	<u>39.47</u>	<u>2.33</u>	<u>3.48</u>	<u>3.30</u>
FinDPO (Ours)	747.10	111.78	3.41	6.05	11.94

2.2 将交易成本纳入投资组合构建

我们进一步评估了交易成本对投资组合表现的影响，以更好地反映实际交易条件。在现实情况下，这些成本可能会显著降低回报，尤其是对每日再平衡策略等高周转率的策略。鉴于我们的可投资范围由相对流动性强的工具构成，特别是 S & P 500 成分股，我们将交易成本参数 k 从 1 至 5 个基点 (bps) 变化，

这与美国股票中常见的 2–5 个基点范围一致。为了体现交易成本，我们使用标准线性模型调整了每日投资组合回报，公式为

$$r_t = R_t - k \cdot \text{Turnover}_t, \quad (7)$$

其中， R_t 是第 t 天未经调整的投资组合回报，而

$$\text{Turnover}_t = \sum_i |w_t^{(i)} - w_{t-1}^{(i)}| \quad (8)$$

表示所有资产的投资组合权重总绝对变化，涵盖多头和空头头寸。

当考虑交易成本时，FinDPO 与所有其他基于情感的投资组合构建方法之间出现了明显的性能差距。即使在高达 5 个基点的高交易成本情况下，FinDPO 仍然是唯一一种能够持续提供显著正回报和稳健风险调整后表现的方法，实现了年度化回报 67%，夏普比率为 2.0。相比之下，所有其他方法，包括 FinLlama，在绝对性能和风险调整后性能上都表现出明显退化，在 4–5 个基点的现实交易成本水平上，所有方法均呈现出非常低至极为负的累积回报。这些方法在交易成本下无法维持盈利能力，突显出它们对换手率的敏感性以及在现实交易环境中缺乏稳健性。相反，FinDPO 在所有交易成本水平上表现出强劲和稳定的性能，强调了其提取稳健交易信号的有效性及其对市场摩擦的抵抗力。这些结果表明，FinDPO 不仅在理想化条件下高度盈利，而且在实际中可部署，即使在具有挑战性的市场条件下也能维持高年度化回报和强健的风险调整后绩效。我们介绍了一种创新的金融情感分析方法，这种方法基于直接偏好优化 (DPO)，这是一种基于人类偏好优化的后训练对齐技术。通过利用预训练的 LLM 的表达能力和上下文理解并将其适应于金融情感分析，这种方法使模型能够更加适应金融领域的细微语言，同时也减少了资源使用和计算开销。相比于 SFT 方法（包括指令调优），偏好对齐技术的强泛化能力使我们提出的 FinDPO 模型在加权 F1 得分上平均比最先进的 FinGPT v3.3 模型高出 11 个百分点。独特的是，我们的框架使得经过微调的因果 LLM 在机器学习基准测试中表现优异，能够被整合到投资组合优化策略中。这使我们能够使用超越当前文献中现有的传统分类评估的真实世界财务指标来评价我们的 FinDPO 模型。总的来说，这些优势使 FinDPO 在我们所知的范围内，成为首个即使在高交易成本下也能维持显著正收益和强风险调整表现的基于情感的方法，因此使其在实际交易环境中稳健且可实用。免责声明：此处所述内容不构成财务建议，也不推荐进行实际资金交易。请使用常识并始终首先咨询专业人士再进行交易或投资。

References

- [1] Dogu Araci. 2019. FinBERT: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063* (2019).
- [2] Jacob Boudoukh, Ronen Feldman, Shimon Kogan, and Matthew Richardson. 2013. *Which news moves stock prices? A textual analysis*. Technical Report. National Bureau of Economic Research.
- [3] Ronen Feldman, Benjamin Rosenfeld, Roy Bar-Haim, and Moshe Fresko. 2011. The stock sonar—sentiment analysis of stocks based on a hybrid approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 25. 1642–1647.
- [4] Sandro Gössi, Ziwei Chen, Wonseong Kim, Bernhard Bernreiter, and Siegfried Handschuh. 2023. Finbert-fomc: Fine-tuned finbert model with sentiment focus method for enhancing sentiment analysis of fomc minutes. In *Proceedings of the Fourth ACM International Conference on AI in Finance*. 357–364.
- [5] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. LoRA: Low-Rank Adaptation of Large Language Models. In *The Tenth International Conference on Learning Representations*.
- [6] Clayton Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media*, Vol. 8. 216–225.

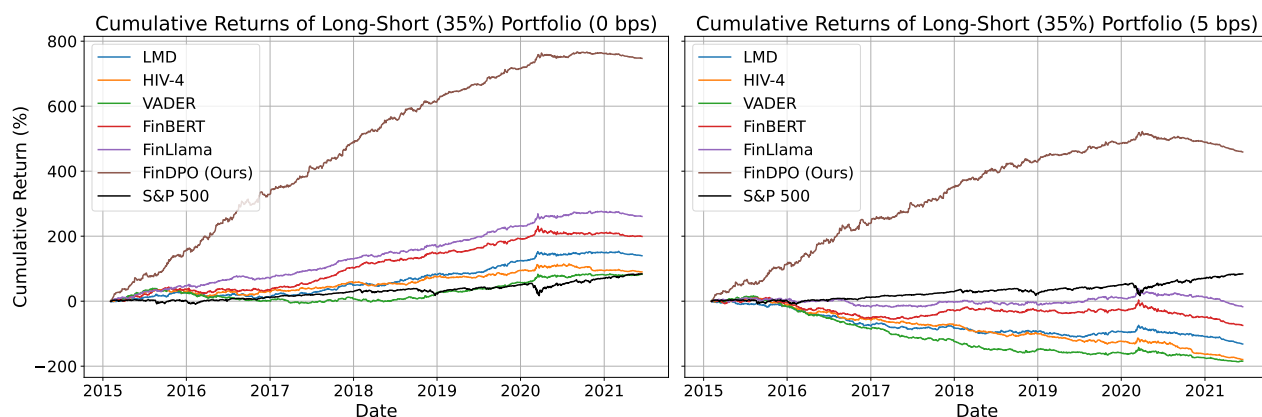


Figure 2: 基于情绪的多空组合在 0 和 5 个基点交易成本下的累计收益，与 S & P 500 基准相比。

- [7] Giorgos Iacovides, Thanos Konstantinidis, Mingxue Xu, and Danilo Mandic. 2024. FinLlama: LLM-Based Financial Sentiment Analysis for Algorithmic Trading. In *Proceedings of the 5th ACM International Conference on AI in Finance*. 134–141.
- [8] D. S. Lim. 2021. BERT-base-NER. <https://huggingface.co/dslim/bert-base-NER>.
- [9] Ilya Loshchilov, Frank Hutter, et al. 2017. Fixing weight decay regularization in adam. *arXiv preprint arXiv:1711.05101* 5 (2017), 5.
- [10] Tim Loughran and Bill McDonald. 2011. When Is a Liability NOT a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance* 66 (02 2011), 35 – 65.
- [11] Neural Magic. 2022. Twitter Financial News Sentiment. <http://precog.iitd.edu.in/people/anupama>
- [12] Pekka Malo, Ankur Sinha, Pekka Korhonen, Jyrki Wallenius, and Pyry Takala. 2014. Good debt or bad debt: Detecting semantic orientations in economic texts. *Journal of the Association for Information Science and Technology* 65, 4 (2014), 782–796.
- [13] Yangtuo Peng and Hui Jiang. 2016. Leverage Financial News to Predict Stock Price Movements Using Word Embeddings and Deep Neural Networks. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, 374–379.
- [14] P. J. Stone, D. C. Dunphy, M. S. Smith, and D. M. Ogilvie. 1966. *The General Inquirer: A Computer Approach to Content Analysis*. MIT Press.
- [15] Neng Wang, Hongyang Yang, and Christina Dan Wang. 2023. FinGPT: Instruction Tuning Benchmark for Open-Source Large Language Models in Financial Datasets. *NeurIPS Workshop on Instruction Tuning and Instruction Following* (2023).
- [16] Oliver Wang. 2023. News with GPT instructions. https://huggingface.co/datasets/oliverwang15/news_with_gpt_instructions
- [17] Yahoo Finance. 2023. Treasury yield 10 years historical data. <https://finance.yahoo.com/quote/%5ETNX/history>
- [18] Boyu Zhang, Hongyang Yang, and Xiao-Yang Liu. 2023. Instruct-FinGPT: Financial Sentiment Analysis by Instruction Tuning of General-Purpose Large Language Models. *ArXiv abs/2306.12659* (2023). <https://api.semanticscholar.org/CorpusID:259224880>
- [19] Wenbin Zhang and Steven Skiena. 2010. Trading Strategies to Exploit Blog and News Sentiment. *Proceedings of the International AAAI Conference on Web and Social Media* 4, 1 (May 2010), 375–378.