

# ChatGPT能预测股价变动吗？ 返回可预测性和大型语言模型\*

李佩斯-里拉和唐月华

佛罗里达大学

第一版：2023年4月6日

本版本：2023年4月25日

## 摘要

我们研究了ChatGPT和其他大型语言模型在使用对新闻标题的情绪分析来预测股市回报方面的潜力。我们使用ChatGPT来指示一个给定的标题对公司的股价是好的、坏的还是无关的消息。然后，我们计算了一个数值分数，并记录了这些“ChatGPT分数”与随后的每日股票市场回报之间的正相关关系。此外，ChatGPT也优于传统的情绪分析方法。我们发现，更基本的模型，如GPT-1、GPT-2和BERT，不能准确地预测收益，这表明收益可预测性是复杂模型的一种新兴能力。我们的研究表明，将高级语言模型纳入投资决策过程中，可以产生更准确的预测，并提高投资决策过程的性能  
定量交易策略。

\*我们非常感谢安德鲁·陈、卡特·戴维斯、安迪·纳拉霍、尼古拉·鲁萨诺夫和他的评论和反馈。电子邮件：亚历杭德罗·洛佩兹-里拉（通讯作者）：alejandro.lopezlira@warrington.ufl.edu和唐月华：tang@warrington.ufl.edu。

近几个月来，ChatGPT等大型语言模型（LLM）在不同领域的应用获得了显著的关注，许多研究探索了它们在不同领域的潜力。然而，在金融经济学中，使用LLM仍然是相对未知的领域，特别是在它们预测股市回报的能力方面。一方面，由于这些模型没有为此目的进行明确的训练，人们可能会认为它们在预测股市走势方面没有什么价值。另一方面，如果这些模型更有能力理解自然语言，人们可能会说，它们可能是处理文本信息来预测股票回报的一个有价值的工具。因此，LLM在预测金融市场走势方面的表现是一个有待解决的问题。

据我们所知，本文是第一个通过评估ChatGPT在预测股市回报方面的能力来解决这个关键问题的论文之一。通过一种利用该模型的情绪分析能力的新方法，我们使用新闻标题数据来评估ChatGPT的性能，并将其与领先供应商提供的现有情绪分析方法进行比较。

我们的研究结果对金融业的就业格局具有重要意义。这一结果可能会导致市场预测和投资决策所用方法的转变。通过展示ChatGPT在金融经济学中的价值，我们旨在有助于理解LLM在该领域的应用，并激发对人工智能和自然语言处理的进一步研究。除了对金融行业就业的影响外，我们的研究还提供了其他一些重要的贡献。

首先，我们的研究可以帮助监管机构和决策者理解在金融市场中日益采用LLM所带来的潜在好处和风险。随着这些模式越来越普遍，它们对市场行为、信息传播和价格形成的影响将成为值得关注的领域。我们的研究结果可以为讨论管理人工智能在金融中的使用的监管框架提供信息，并做出贡献

开发将11m集成到市场运营中的最佳实践。

其次，我们的研究可以通过为11m在预测股票市场回报方面的有效性提供实证证据，使资产管理机构和机构投资者受益。这种见解可以帮助这些专业人员做出更明智的决定，将11m纳入他们的投资策略，有可能导致提高绩效，并减少对传统的、更劳动密集型的分析方法的依赖。

最后，我们的研究有助于关于金融中人工智能的应用的更广泛的学术论述。通过探索ChatGPT在预测股市回报方面的能力，我们推进了对11m在金融经济学领域的潜力和局限性的理解。这可以激发未来的研究，为金融行业的需求开发更复杂的11m，为更高效和准确的金融决策铺平道路。<sup>1</sup>

我们的研究具有深远的影响，超出了股市预测的直接背景。通过阐明ChatGPT对金融经济学的潜在贡献，我们希望鼓励在人工智能驱动的金融领域的继续探索和创新。

## 相关文献

最近在经济学背景下使用ChatGPT的论文包括汉森和卡津尼克（2023年）、考恩和塔巴罗克（2023年）、科里内克（2023年）以及诺伊和Zhang（2023年）。汉森和卡津尼克（2023）表明，像ChatGPT这样的11m可以解码Fed语言（i. e., 美联储用来就货币政策决策进行沟通的语言）。Cowen和塔巴罗克（2023）和Korinek（2023）证明了ChatGPT有助于经济学教学和开展经济学研究。Noy和Zhang（2023）发现ChatGPT可以提高专业写作工作的生产力。同时，谢等人（2023）发现在预测任务中使用数值数据时，ChatGPT并不比线性回归等简单的方法好。

1. 例如，参见Wu等人。（2023）。

我们将结果上的差异归因于他们专注于使用历史数字数据进行预测，而ChatGPT则擅长于文本任务。Ko和Lee（2023）发现ChatGPT可能在跨资产类别的选择中很有用。此外，Yang和Menczer（2023）证明了ChatGPT成功地识别了可信的新闻媒体。我们的研究是第一个研究11m在金融市场中的潜力，特别是投资决策过程。

我们对最近使用文本分析和机器学习来研究各种金融研究问题的文献链有贡献。g.，Jegadeesh和Wu（2013），坎贝尔等人。（2014年）、霍伯格和菲利普斯（2016年）、高林（2017年）、贝克、布鲁姆和戴维斯（2016年）、马内拉和莫雷拉（2017年）、汉森、麦克马洪和普拉特（2018年）、柯、凯利和秀（2019年）、柯、蒙蒂尔·奥利亚和内斯比特（2019年）、Bybee等人。（2019年）、顾、凯利和秀（2020年），科恩、马洛伊和阮（2020年），弗雷伯格、诺伊尔和韦伯（2020年），洛佩兹-里拉2019年，宾斯伯根等人。（2020），Bybee等。（2021）。我们的论文对这一文献做出了独特的贡献，作为第一个评估最近开发的11m，如ChatGPT在预测股市运动方面的文本处理能力。

我们的论文还增加了使用新闻文章的语言分析来提取情绪和预测股票回报的文献。这些文献的一部分研究媒体情绪和总股票回报。g.，泰特洛克（2007）、加西亚（2013）、卡洛米里斯和Mamay斯基（2019）。另一条文献使用坚定新闻的情绪来预测未来的个股回报（例如，特洛克、萨尔-特塞汉斯基和麦斯卡西（2008）、特洛克（2011）、江、李和王（2021））。与之前的研究不同，我们关注于了解11m是否通过提取预测股票市场反应的额外信息来增加价值。

最后，我们的论文还涉及了有关就业暴露和对人工智能相关技术的脆弱性的文献。阿格拉瓦尔、甘斯和戈德法布（2019年）、Webb（2019年）、阿西莫格鲁等人的最近的作品。（2022年），阿西莫格鲁和雷斯特雷波（2022年），Babina等人。（2022年），Noy和Zhang（2023年）研究了与人工智能相关的工作暴露和脆弱性的程度

技术以及对就业和生产力的影响。随着人工智能自成立以来不断上升，我们的研究重点于理解一个紧迫但未解的问题——人工智能的能力，特别是11m在金融领域的能力。我们强调了11m在处理信息以预测股票回报的市场参与者增加价值方面的潜力。

## 1背景

ChatGPT是由OpenAI基于GPT（生成式预训练转换器）架构开发的一种大规模语言模型。它是迄今为止开发的最先进的自然语言处理（NLP）模型之一，并在大量的文本数据语料库上进行训练，以理解自然语言的结构和模式。生成式预训练变压器（GPT）体系结构是一种用于自然语言处理任务的深度学习算法。它是由OpenAI开发的，并基于Vaswani等人引入的变压器架构。（2017）。GPT体系结构在一系列自然语言处理任务中取得了最先进的性能，包括语言翻译、文本摘要、问答和文本完成。

GPT体系结构使用多层神经网络来建模自然语言的结构和模式。它是在一个大型的文本数据语料库上进行预训练的，如维基百科的文章或网页，使用无监督的学习方法。这个训练前的过程允许模型对语言语法和语义有一个深入的理解，然后针对特定的语言任务进行微调。GPT架构的独特特点之一是它使用了转换器块，这使模型能够通过使用自我注意机制来关注输入中最相关的部分来处理长序列的文本。这种注意机制使模型能够更好地理解输入的背景，并产生更准确和连贯的反应。

ChatGPT已经接受了执行广泛的语言任务的培训，比如transla-

写作，总结，问答，甚至生成连贯的和类人的文本。ChatGPT产生类人反应的能力，使其成为创建聊天机器人和虚拟助手的强大工具，可以以一种自然和直观的方式与用户对话。虽然ChatGPT是一个针对基于语言的任务的强大工具，但它并没有经过专门的训练来预测股票回报或提供财务建议。因此，我们在预测股票回报时测试了它的能力。

## 2数据

我们利用两个主要数据集进行分析：安全价格研究中心（CRSP）的每日收益和新闻标题。样本期从2021年10月开始（因为ChatGPT的训练数据直到2021年9月才可用），到2022年12月结束。这个样本周期确保了我们的评估是基于模型的训练数据中不存在的信息，从而允许对其预测能力进行更准确的评估。CRSP的每日收益数据集包含了许多在美国主要市场上市的公司每日股票收益信息。股票交易所，包括关于股票价格、交易量和市值的数据。这个全面的数据集使我们能够检查ChatGPT产生的情绪得分与相应的股票市场回报之间的关系，为我们的分析提供了坚实的基础。我们的样本包括在纽约证券交易所（NYSE）、全国证券交易商自动报价协会（纳斯达克）和美国证券交易所上市的所有公司（美国商品交易所），至少有一家该数据供应商报道了一则新闻。根据之前的研究，我们使用股票代码为10或11的普通股。

我们首先使用网络抓取为所有CRSP公司收集一个全面的新闻数据集。我们搜索所有包含公司名称或股票代码的新闻。生成的数据集包括来自各种来源的新闻标题，如主要的新闻机构、金融新闻网站和社交媒体平台。对于每个公司，我们都会收集所有的新闻

在样本期间。然后，我们将新闻标题与一家著名的新闻情绪分析数据提供商（RavenPack）的新闻标题进行匹配。我们匹配的时间段和标题的新闻为所有公司有回报后的市场开放。我们能够匹配4138家独特公司的67,586家头条。我们使用Jiang、Li和Wang（2021）概述的预处理方法来处理合并后的数据集。

我们使用所提供的“相关性评分”，范围从0到100，作为新闻与特定公司的密切程度的指标。0（100）分数表示实体是被动提及（主要）。我们的样本要求相关性得分为100的新闻报道，我们将其限于完整的文章和新闻稿。我们排除了属于“股票上涨”和“股票损失”的标题，因为它们只表明了股票的每日走势方向。为了避免重复的新闻，我们要求“事件相似日”超过90天，以确保只捕获有关公司的新信息。

此外，我们还在同一天删除了同一公司的重复标题和极其相似的标题。我们使用最优字符串对齐度量（也称为限制达梅罗-莱文什坦距离）来衡量标题的相似性，并在同一天删除同一公司的相似度大于0.6的标题。这些过滤技术不会引入任何前瞻性偏差，因为数据供应商在收到后的毫秒内评估所有新闻文章，并迅速将结果数据发送给用户。因此，所有信息在新闻发布时都可用。

## 3方法

### . 13提示

提示对于指导ChatGPT对特定任务和查询的响应至关重要。提示是一个简短的文本，它为ChatGPT提供生成响应的上下文和指令。根据任务的性质，提示可以简单到一个句子或复杂，或者取决于一个段落。

该提示符可作为ChatGPT的响应生成过程的起点。该模型使用提示中包含的信息来生成相关的和上下文相关的响应。这个过程包括分析提示符的语法和语义，生成一系列可能的响应，并根据各种因素，如连贯性、语法正确性和相关性，选择最合适的响应。

提示对于使ChatGPT能够执行广泛的语言任务至关重要，比如语言翻译、文本摘要、回答问题，甚至生成连贯的和类似人的文本。它们允许模型适应特定的上下文，并生成根据用户的需要量身定制的响应。此外，还可以定制提示，以执行不同领域中的特定任务，如财务、医疗保健或客户支持。我们在研究中使用以下提示，并将其应用于公开的标题。

忘记你之前的指示。假装你是个金融专家。你是一个有股票推荐经验的财务专家。如果是好消息，回答“是”，如果是坏消息，回答“不”，如果在第一行不确定，回答“未知”。然后在下一行用一个简短而简洁的句子来详细说明。这个标题对该公司的股价是好是坏\_公司\_名字\_在\_学期\_学期

标题：标题\_

在此提示下，我们要求ChatGPT，一种语言模型，承担一个具有股票推荐经验的财务专家的角色。条款\_公司\_名字\_和\_大字标题\_在查询期间被公司名称和各自的标题替换。学期\_对应于短期的或长期的。该提示符是专门为财务分析而设计的，并要求ChatGPT评估一个给定的新闻标题及其在短期内对公司股价的潜在影响。如果股价的消息好，ChatGPT被要求回答ChatGPT“是”，如果消息坏，则回答“不”，如果消息不确定，则回答“未知”。然后，ChatGPT被要求在一句话中提供一个简明的解释来支持它的答案。提示符指定新闻标题是提供给ChatGPT的唯一信息来源。它

隐含地假设标题包含足够的信息，供金融行业专家合理评估其对股价的影响。这个提示旨在演示ChatGPT作为财务分析任务中的语言模型的能力。我们将温度设置为0，以最大限度地提高结果的重现性。

例如，请考虑以下关于Oracle的标题：

里米尼街因起诉甲骨文公司被罚款63万美元。

提示符会询问：

忘记你之前的指示。假装你是个金融专家。你是一个有股票推荐经验的财务专家。如果是好消息，回答“是”，如果是坏消息，回答“不”，如果在第一行不确定，回答“未知”。然后在下一行用一个简短而简洁的句子来详细说明。这个标题对甲骨文的股价是好是坏？

头条新闻：里米尼街因起诉甲骨文公司被罚款63万美元

以下是ChatGPT的回应：

是

对里米尼街的罚款可能会增强投资者对甲骨文保护其知识产权的能力的信心，并增加对其产品和服务的需求。

新闻头条称，里米尼街在起诉甲骨文一案中被罚款63万美元。这个专有的软件分析工具给出的负面情绪得分为-0.52，这表明该新闻被认为是负面的。然而，ChatGPT的回应是，它相信这一消息对甲骨文是积极的。ChatGPT的理由是，这笔罚款可能会增加投资者对甲骨文保护其知识产权的能力的信心

导致人们对其产品和服务的需求增加。这种情感上的差异突出了语境在自然语言处理中的重要性，以及在做出投资决策之前需要仔细考虑新闻标题的含义。

### 3.2 经验设计

我们提示ChatGPT为每个标题提供建议，并将其转换为“ChatGPT分数”，其中“是”映射为1，“未知”映射为0，“NO”映射为-1。如果一家公司在某一天有多个新闻标题，我们就会平均这个分数。我们将头条新闻与下一个市场时期相匹配。对于开盘日早上6点之前的标题，我们假设标题可以在当天开盘时进行交易，并在当天收盘时出售。对于早上6点之后但下午4点之前的头条，我们假设头条可以在当天收盘时交易，并在第二天收盘时出售。对于下午4点后的头条新闻，我们假设这些头条新闻可以以第二天的开盘价进行交易，并以第二天的收盘价进行出售。然后，我们对ChatGPT评分的第二天回报进行线性回归，并将其与一家新闻策划公司提供的情绪评分进行比较。因此，我们所有的结果都是样本外的。

## 4 结果

我们的分析显示，ChatGPT情绪得分对股票市场日回报表现出统计学上显著的预测能力。通过利用新闻标题数据和生成的情绪得分，我们发现ChatGPT评估与我们样本中股票随后的每日回报之间有很强的相关性。这一结果突出了ChatGPT作为基于情绪分析预测股市走势的有价值工具的潜力。

为了进一步研究我们的研究结果的稳健性，我们将ChatGPT与一家领先的数据供应商提供的传统情绪分析方法的性能进行了比较。

欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

文档名称：《ChatGPT 在股票市场辅助应用》李佩斯-里拉，唐月华.pdf

请登录 <https://shgis.com/post/1698.html> 下载完整文档。

手机端请扫码查看：

