ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 40,No. 5,Sep. 2025,pp. 1193－1206
DOI:10. 16337/j. 1004-9037. 2025. 05. 007

http://sjcj. nuaa. edu. cn
E-mail:sjcj@nuaa. edu. cn
Tel/Fax:＋86-025-84892742

大语言模型指导的多模态时序-语义预测框架

叶诗敏[1]，刘非菲[1]，张　岩[2]

(1. 苏州工学院商学院,苏州 215500;2. 厦门大学人工智能研究院,厦门 361005)

摘　要：多模态预测任务通常需要同时对文本、图像与结构化数值等异构数据进行建模,以在复杂环境中实现稳健的时序建模、跨模态语义对齐与可解释推理。传统单模态或弱融合方法难以在语义对齐、信息互补与跨源推理方面取得一致性,且深度模型的黑箱特性限制了结果的可解释性。与此同时,大语言模型(Large language model, LLM)在语义理解、指令跟随与推理方面展现出强大能力,但其与时序建模、跨模态对齐及实时知识整合之间仍存在鸿沟。因此,提出LLM指导的多模态时序-语义预测框架,通过将变分推理的时序建模与LLM的语义分析相结合,构建“时序-语义-决策”的协同机制:时序模块利用递归潜变量与注意力机制提取历史行为模式;语义模块利用领域化语言模型与多模态编码器提炼高层语义与解释;两者在可学习融合器中联合优化,并提供不确定性标注与可解释报告。在StockNet、CMIN-US和CMIN-CN数据集上的实验表明,本文方法准确率达63.54%,较最优基线提升5.31个百分点,马修斯相关系数(Matthews correlation coefficient, MCC)提升至0.223。本文研究为多模态时序预测提供了统一范式,并在金融科技领域展现出应用潜力。

关键词：多模态;大语言模型;人工智能;预训练模型;时间序列预测
中图分类号：TP391　　**文献标志码**:A

Large Language Model-Guided Multi-modal Time Series-Semantic Prediction Framework

YE Shimin[1], LIU Feifei[1], ZHANG Yan[2]

(1. College of Business, Suzhou University of Technology, Suzhou 215500,China; 2. Institute of Artificial Intelligence, Xiamen University, Xiamen 361005,China)

Abstract: Multi-modal prediction tasks typically require the simultaneous modeling of heterogeneous data, including text, images and structured numerical information, to achieve robust inference and explainable decision-making in complex environments. Traditional uni-modal or weak fusion methods struggle to consistently address semantic alignment, information complementation and cross-source reasoning, while the inherent black-box nature of deep models limits the result interpretability. Meanwhile, the large language model (LLM) has demonstrated strong capabilities in semantic understanding, instruction following, and reasoning, yet a gap remains in their performance for time series modeling, cross-modal alignment, and real-time knowledge integration. To address these challenges, this paper proposes a LLM-guided multi-modal time series-semantic prediction framework. By combining variational inference-based time series modeling with LLM -driven semantic analysis, the approach establishes a collaborative

收稿日期:2025-06-15;**修订日期**:2025-08-10

"temporal-semantic-decision" mechanism: The temporal module extracts historical behavior patterns using recurrent latent variables and attention mechanisms; the semantic module distills high-level semantics and interpretations through domain-specific language models and multi-modal encoders; and both components are jointly optimized via a learnable fusion module, which also provides uncertainty annotations and explainable reports. Experiments on the StockNet, CMIN-US, and CMIN-CN datasets demonstrate that the approach achieves an accuracy of 63.54%, an improvement of 5.31 percentage points over the best baseline and an Matthews correlation coefficient (MCC) elevated to 0.223. This study offers a unified paradigm for multi-modal time series prediction and underscores its promising application in the field of financial technology.

Key words: multi-modal; large language model (LLM); artificial intelligence; pre-trained model; time series prediction

引　言

多模态智能系统正从"单一模态感知"走向"跨模态理解与协同决策"。在实际任务中,文本、图像与结构化数值等数据源同时存在,且在时间尺度、噪声水平与信息密度上差异显著。如何在动态环境中实现稳健的时序建模、跨模态语义对齐与可解释推理,已成为通用人工智能的重要课题。时序预测是一类经典的问题,在学术界和工业界都有着广泛的研究和应用。许多问题都可抽象为时间序列问题,例如市场价格、天气变化等。传统方法大致分为3类:(1) 统计时序模型,如依赖平稳性假设的自回归差分移动平均模型(Auto-regressive moving average model,ARIMA),这种模型在面对非平稳、突变与异质噪声时表现受限;(2) 经典机器学习模型,如支持向量机(Support vector machin, SVM)和随机森林,这种模型需要大量人工特征工程,难以捕捉跨模态关联;(3) 深度学习方法,如长短期记忆网络(Long short-term memory, LSTM)以及图神经网络(Graph neural network, GNN),它们在时间依赖建模上更为有效,但仍存在跨模态融合与可解释性不足等问题。

在此背景下,大语言模型(Large language model, LLM)凭借在语义理解、指令跟随与多步推理方面的能力,为多模态任务提供了"语言为接口"的统一范式。然而,LLM与时序建模、跨模态对齐及与实时知识的整合之间仍存在鸿沟:一方面,缺乏针对目标领域的大规模高质量训练数据与严格的时序对齐机制,容易引发幻觉和不稳定的推断;另一方面,如何将LLM的语义推理与数值/图像的时序行为建模进行"结构化对接",仍需系统性的算法与工程化设计。现有算法面临如下两种局限。(1) 多源异构数据的割裂性。任务目标往往受多重因素驱动:文本传递事件与语义线索(如政策变动、事故通告、舆情波动),时序图像反映历史形态与结构性模式,结构化指标刻画系统状态与基本事实。不同模态在时间密度、采样频率与噪声结构上差异明显,导致在特征对齐与语义融合过程中容易产生信息损失或引入噪声。现有研究常聚焦单一模态或松散耦合的双模态,尚未在统一表示空间中实现稳定对齐。(2) 可解释性不足影响决策信任度。在复杂预测任务中,用户不仅需要结果,更需要理解模型给出结论的逻辑与依据。然而,多模态深度模型通常是"黑箱",内部决策机制难以外显。这不仅降低了模型在高可靠场景中的可采纳性,也使风险管理与合规评估更具挑战。现有方法,如注意力权重、显著性图等,在复杂场景下常面临解释稳定性与一致性不足的问题。同时,跨模态推理链条较长,缺少结构化、不确定性标注与人机协同的解释框架。

针对上述挑战,本文提出大语言模型指导的多模态时序-语义(LLM-guided multi-model time series-semantic, LMTSS)预测框架,核心创新在于构建"时序-语义-决策"三元协同机制。(1) 多模态感知层

面,为了有效地整合和利用这些不同的数据类型,使用多模态大语言模型[1]中预训练的视觉编码器,用于处理和综合来自多种模式的信息,从而全面并细致地了解时序数据。(2)决策机制层面,使用双通路协同架构,时序通路基于变分递归网络捕捉价格波动规律,这种设计可以防止高度结构化的时序信号被文本模态的语义噪声干扰,反之亦然,从而实现更稳健的特征学习。时序通路基于变分递归网络捕捉历史变化规律,语义通路依托大语言模型生成语义驱动信号。二者通过权重融合共同决策,同时使用偏差检测模块,即当语义预测与时序预测分歧超过阈值时,将其判定为高不确定性样本。(3)应用生态层面,实现端到端的智能决策闭环。输入多源异构数据后,系统输出两类结果:未来时段变化概率和可视化分析报告(含决策建议依据)。本文旨在借助大语言模型的多模态能力,构建一个通用的多模态时序-语义预测框架,充分整合多源信息,实现高精度、高可解释性的时序预测。与现有研究相比,系统实验与消融实验结果验证了所提方法在多种市场情境下的优越性,并探讨了其在实际投资决策中的应用前景。

1 研究现状

近年来,大语言模型与多模态学习的结合推动了人工智能从单模态任务向跨模态推理、生成与决策的全方位能力演进。这种技术融合不仅在通用人工智能(General artificial intelligence, AGI)研究中产生了深远的影响,更为核心的数字经济领域带来了范式变革,其强大的语义理解与推理能力为构建下一代智能决策系统提供了新的技术基座。本节将回顾大语言模型的发展及其在时序分析中的迁移应用,探讨多模态大语言模型(Multi-modal large language model, MLLM)如何赋能复杂的时序预测与决策任务。

1.1 大模型与多模态

大规模预训练模型,简称“大模型”,在人工智能领域取得了显著突破,其典型代表包括GPT系列[2]、BERT[3]和T5[4]等。这类模型的显著特征在于参数规模巨大(通常达到数十亿甚至数千亿级)、网络结构深度高、训练数据覆盖面广,并通过庞大的容量与极深的网络结构在特征抽取、语义建模与推理能力方面表现出显著优势。与传统的浅层机器学习模型相比,大模型在任务性能随规模增长的过程中呈现出一种临界点现象——当参数规模与训练数据量超过某一阈值后,模型会表现出原本未被显式设计或预期的“涌现能力”[5],如零样本推理、跨任务迁移和复杂语义理解等。这一特性使大模型在处理多样化、复杂度高的数据分布时展现出卓越的泛化与推理性能[6]。

大模型的核心训练范式通常包括3个阶段:首先在海量无标注数据上进行预训练,通过语言建模目标(如自回归或掩码预测)学习通用语义表示与深层统计关系;其次利用领域特定的小规模标注数据进行微调,将模型知识迁移至目标任务;最后可在部分应用中结合人类反馈强化学习进行对齐优化[7],使模型输出在准确性、相关性与安全性上更贴近人类需求。大模型凝聚了大数据精华,形成一个“隐式知识库”,是构建人工智能应用的关键载体。这类模型通常在巨量无标注数据上进行预训练,从中学习通用特征与规则。如图1所示,在应用开发阶段,开发者可选择对大模型进行微调(使用下游特定任务的小规模标注数据进行二次训练)。这种“先通用预训练、再特定适配”的方法使得大模型具备了迁移学习的天然优势。当目标领域数据稀缺时,模型可依托预训练阶段形成的庞大“隐式知识库”快速适应领域任务[8]。这种能力不仅推动了自然语言处理领域的范式转变,也广泛提升了计算机视觉、语音识别、推荐系统等任务的性能。值得注意的是,这一训练策略在金融等数据密集型

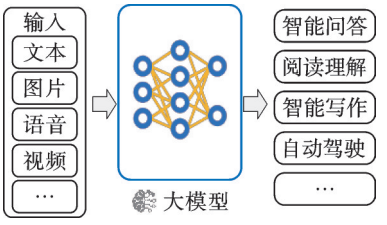

图1 大模型统一不同下游任务的能力

Fig.1 The ability of large models to unify different downstream tasks

领域同样展现出潜力。例如在FinBERT模型[9]中,通用语言模型经过领域化微调后,可以高效处理金融新闻、政策文件、行业研报等专业语料,实现信息抽取、情感分析和事件驱动推理等功能,从而为后续的市场分析和风险评估提供高质量的语义输入。

然而,大模型的高性能通常以高昂的计算与存储代价为前提,表现为预训练阶段需依赖高性能计算集群与分布式并行训练框架,推理阶段的计算开销亦不可忽视。因此,其实际部署多依托云计算平台或本地高性能计算设施,并在部分应用场景中探索参数高效化,如低秩适(Low-rank adaptation, LoRA)、前缀微调(Prefix-tuning)与模型蒸馏等轻量化技术,以降低资源消耗并提高部署灵活性。总体而言,大模型的发展体现出"规模-能力"之间的强相关性,在技术路线上经历了从单一任务到多任务、从单模态到跨模态的演进。这一趋势为后续的多模态学习奠定了坚实的技术基础,也为金融等领域引入跨源信息融合和深度语义推理提供了可能。

1.2 多模态大模型与跨模态方法

多模态学习旨在同时利用来自不同模态(如文本、图像、音频、结构化数据等)的信息,通过联合建模提升对复杂任务的理解与推理能力[10]。人类的感知与认知过程天然具备多模态特性,而传统单模态模型无法有效捕捉不同信息源之间的互补性与语义关联,从而在任务表现上存在瓶颈[11]。随着Transformer架构的普及与大规模预训练方法的成熟,多模态大模型逐渐成为跨模态人工智能研究的主流方向[12]。

在架构设计上,多模态大模型可按信息交互策略分为3类:基于单流编码的融合模型、基于双流交互的理解模型以及基于编码-解码结构的生成模型[13]。单流方法将不同模态的输入通过统一的嵌入空间映射后,直接送入共享的Transformer编码器中进行联合建模;双流方法则为不同模态分别配置独立的编码器,并通过跨注意力或门控机制实现信息交互;编码-解码范式则常用于生成任务,如文本生成图像、图像生成文本等,其中编码器负责融合多模态信息,解码器则根据任务需求生成目标模态输出[14]。在生成式多模态模型方面,GPT系列在文本生成的成功经验直接启发了跨模态的扩展研究。例如,OpenAI发布的DALL·E[15]模型在拥有120亿参数的GPT-3架构基础上,通过对图像-文本对的预训练,实现了根据文本描述生成相应图像的能力;Google的Imagen模型则进一步在文本-图像生成的质量与可控性上取得突破。这类模型展现了将语言模型的生成优势迁移至视觉领域的可行性,同时凸显了多模态大模型在跨领域任务中的广阔应用前景。本文提出的LMTSS框架属于带有后期融合机制的双流模型,其时序通路与语义通路分别建模,最终在决策层进行融合,兼具模态特异性与交互性。

近年来,各行业纷纷着手构建面向特定领域的专用大模型。在众多应用场景中,金融行业凭借其数据流通规模庞大与数字化基础完善等优势,被普遍视为生成式大模型落地的高潜力领域。在该类场景中,信息来源往往包括新闻与公告等非结构化文本、K线图与技术指标图等图像数据以及财务报表与宏观经济指标等结构化数据。传统方法通常独立处理各模态数据,缺乏对它们之间隐含的语义关联与因果关系的统一建模能力,从而限制了预测与决策的上限。多模态大模型能够通过统一的表示学习框架,将文本、图像与数值数据映射至共享语义空间,从而在时序预测、风险预警、策略生成等任务中实现深度的信息整合。例如,在量化交易任务中,文本模态可解析新闻中的事件逻辑[16];图像模态可识别技术图表中的形态模式;结构化模态则可量化市场微观结构特征[17]。这种多模态融合能力同样适用于气象分析或工业运维等广泛领域。通过多模态融合,模型能够生成对市场动态更为全面的表征。

此类复杂时序预测与决策任务是多模态大模型展现其价值的典型场景。已有研究探索了多模态数据在金融预测中的组合方式,例如将新闻文本与历史价格序列共同输入时间依赖性模型以捕捉短期波动趋势,或通过引入情绪分析模块揭示媒体报道与市场回报之间的非线性关联。这些尝试不仅验证了多模态融合的有效性,也揭示了其在高维、动态和噪声较大的时序任务中所面临的共同挑战,包括模态间语义鸿沟、信息密度差异、训练成本高昂以及可解释性不足等。

总体而言,多模态大模型的发展趋势体现为3个方向:(1)向更大规模、更高参数量的架构演进,以提升跨模态信息对齐与生成的能力;(2)探索高效的融合机制和轻量化推理方法,以适应资源受限的部署环境;(3)针对特定领域开展定制化的多模态预训练与微调,形成能够理解专业语境与跨源数据关系的专用模型。这一趋势为后续在更广泛的时序预测与决策任务中引入多模态大模型奠定了理论与技术基础。

2 LMTSS预测框架

2.1 整体概述

本文构建了LMTSS预测框架,用于预测金融领域的时间序列走势,其主要由时序预测子模块与大语言模型子模块组成。如图2所示,时序模块基于历史股票价格与社交媒体数据进行建模,通过引入递归连续潜变量应对市场波动的随机性,并且假设时序的运动预测可以受益于学习其历史运动。模型构建过程中,使用了带有时间依赖性的多个输入,通过时间序列特征提取和递归神经网络来实现对时间信息的建模,从而进行时序预测。通用大语言模型的接入,使LMTSS拥有了在金融上下文中的自然语言理解能力,对多源信息进行解释、关联分析和任务指令解析,从而生成高层次语义特征与决策依据。

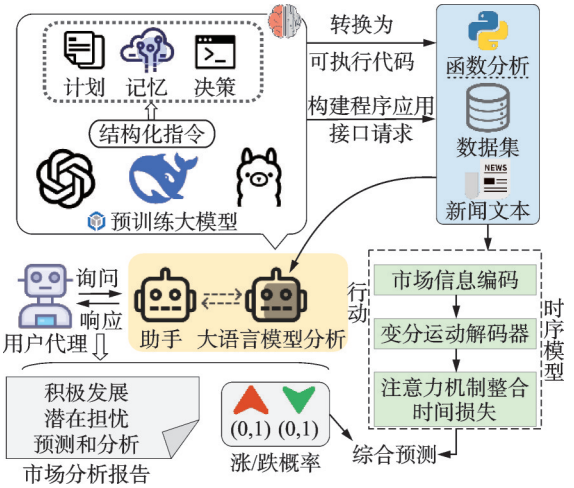

图2 LMTSS的整体框架

Fig.2 Overall framework of LMTSS

在输入层面,框架可接收多种模态的数据,包括文本(如描述性说明、日志、评论等)、图像(如技术图形、时序可视化等)以及结构化数值(如传感器读数、统计指标等)。时序模块通过递归神经网络及其变体,结合时间注意力机制,对多模态输入进行统一的时间序列特征建模,以捕捉长期依赖关系与潜在动态模式。大语言模型模块则对经过编码的多模态信息进行语义聚合与逻辑推理,并可根据任务需求输出分类结果、趋势预测或概率性评估。为提升模型的泛化能力与稳健性,框架引入了多模态特征融合机制,将语义推理结果与时序预测结果在决策层进行联合优化。该融合策略支持多任务学习与不确定性建模,可在不同任务和领域中灵活适配,从而为多模态数据驱动的智能预测提供统一的技术路径。

2.2 大语言模型技术基础和文本编码

大语言模型以生成式预训练变换器为代表,其技术基础建立在深度学习与自然语言处理的交叉融合之上。模型核心采用基于自回归语言建模的Transformer架构,通过对海量文本的无监督预训练学习,掌握语言规律与语义知识,并在多种任务中展现出卓越的生成与理解能力。在训练流程方面,GPT系列通常经历大规模预训练、有监督微调和基于人类反馈的强化学习对齐(Reinforcement learning from human feedback, RLHF)三个阶段。预训练阶段通过自回归建模掌握通用语言模式和丰富的语义关联;有监督微调阶段在领域标注数据上优化参数,使模型能够适配特定任务需求;RLHF阶段则引入人类偏好信号,通过奖励模型与强化学习优化生成策略,显著提升生成文本的流畅性、相关性与安全性。本文所采用的大语言模型方法建立在生成式预训练Transformer架构之上,其核心思想是以自回归语言建模为目标函数,在给定历史文本序列的条件下预测下一个词元。

在文本处理过程中，输入端首先将金融新闻、公告、研报、财报等原始文本进行分词和Token化，并映射为高维嵌入向量，结合位置编码生成模型可处理的输入表示。Transformer的多层解码器堆叠在此基础上，通过掩码自注意力机制捕捉金融文本中跨句、跨段的依赖关系，例如将“利率上调”与“股价波动”建立关联；多头注意力机制则可同时聚焦不同层面的信息，如宏观经济背景、公司财务指标与行业情绪；前馈网络进一步提炼与评估相关的关键特征，形成具有高信息密度的语义表示。输出端通过线性映射和Softmax运算，将内部表示转化为预测结果或自然语言解释，使模型不仅能生成与金融任务相关的答案，还能对市场走势、风险因素等进行推断。

对于一般时序模型的文本编码，参考StockNet模型[16]。该模型结合文本和价格信号，利用神经变分推理来处理模型中的后验推断问题，并采用混合目标函数来灵活地捕捉预测依赖关系。单个样本可预测一系列走势$\boldsymbol{y}=[y_1, y_2, \cdots, y_T]$，其中$y_T$为主要预测目标，$\boldsymbol{y}^*=[y_1, y_2, \cdots, y_{T-1}]$为时间辅助目标。辅助目标与主要目标结合，通过多任务学习提升预测精度。具体而言，模型在编码器中通过历史市场信息的随机变量进行信息编码，并通过变分运动解码器递归地解码由潜在因素驱动的运动变化。为了更好地捕捉预测的时间依赖性，模型使用时间辅助机制来进一步强化学习过程。

2.3 用于市场评估的大语言模型

本文旨在探讨如何借助现有的多模态大型语言模型，构建一个高效、可扩展的价格预测系统，结合自然语言文本、结构化财务指标及技术图形等多源数据，实现对股市走势的精准预测与语义解释。本文方法部分在FinRobot[1]的多模态数据处理流程与任务指令格式基础上进行重构设计，同时结合StockNet类时序建模思想，引入概率加权融合机制与不一致性检测模块，从而形成LMTSS的整体方法框架。与FinRobot相比，LMTSS的核心创新在于将时序建模与语义推理结构化对接，实现跨模态的一致性约束并提升可解释性。图3是基于大语言模型的金融应用开源人工智能智能体平台的整体框架结构示意图。

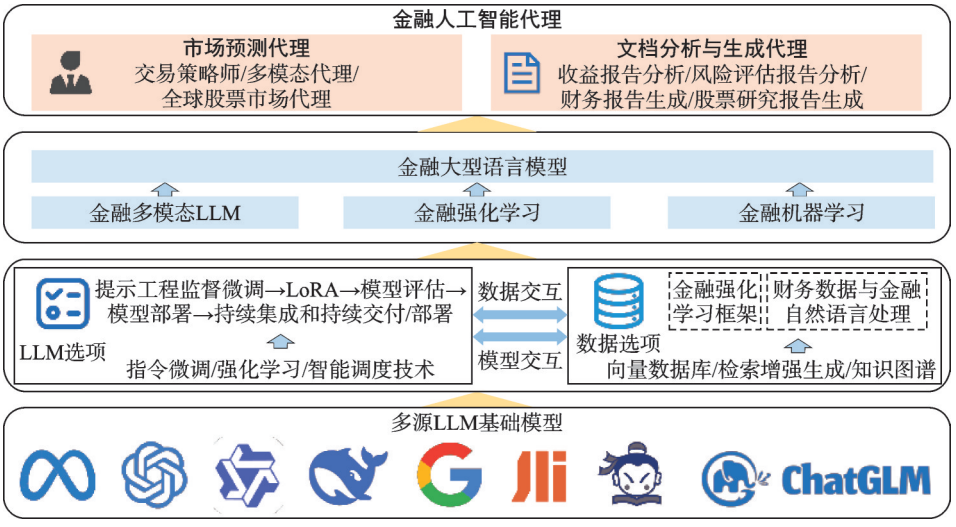

图3 基于大语言模型的金融应用开源人工智能智能体平台的整体框架

Fig.3 Overall framework of an open-source artificial intelligence agent platform for financial applications based on large language models

市场预测器集成多源全球市场数据，利用多样的数据源收集信息并做出预测决策。具体而言，市场预测器收集关于多方面公司信息的数据，例如最近的新闻、最新的基本财务数据和目标市场价格等，

包括美国股市、中国股市、加密市场及其他可能的扩展。根据多任务指令调优框架,市场预测器遵循复杂的提示格式。具体而言,市场预测器收集多方面公司信息的数据,然后进行提示工程,以“任务指令与公司信息(公司概述+最近价格+最近新闻+最新基本财务数据)”的结构格式化指令性提示。图4展示了市场预测器提示模板的一个示例。

指令:您是一位经验丰富的股票市场分析师。您的任务是根据公司在过去几周内的相关新闻和季度财务公司列出公司的积极发展和潜在问题,然后将它们与您对整体金融经济市场判断的看法相结合,提供预测并分析下周的股票价格变化。您的答案格式应如下:
[积极发展]:
1. ……
[潜在问题]:
1. ……
[预测和分析]:
……
信息:
a. 公司简介
b. 股价变化
c. 最新新闻信息
d. 最近的基本金融
指令:基于2024-04-19之前的所有信息,首先分析AAPL的积极发展和潜在问题。分别提出2~4个最重要的因素,并保持简洁。然后对下周的AAPL价格变动进行预测(2024-04-22至2024-04-26)。提供摘要分析以支持预测。

图4 市场预测提示模板

Fig.4 Market forecaster prompt template

FinRobot原生设计中已集成多模态编码路径,包括文图融合、结构化数值向量建模,在此基础上加以改造,用于捕捉金融模态之间的语义依赖与潜在因果关系。首先,对于文本模态的处理,采用了预训练的金融语言模型FinBERT作为基础编码器,以保留财经文本中蕴含的语义特征。设原始文本输入为序列$[\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_n]$,通过Transformer模型的多层自注意力机制,对输入序列中所有词元的初始嵌入进行动态聚合与交互,从而生成蕴含全局上下文信息的语义向量表示。其次,对于图像模态,使用对比语言-图像预训练(Contrastive language-image pre-training, CLIP)视觉编码器将图像转化为特征向量,与文本数据结合处理,利用预训练对图像数据进行初步处理。通过对金融图表(如K线图、折线图等)的图像进行标注和描述,模型可以学习到图像中可视模式与文本解释之间的复杂关系。

为了有效地整合和利用这些不同的数据类型,金融多模态LLM集成过程的数学表达式为

$$F(\boldsymbol{x}_{\mathrm{t}},\boldsymbol{x}_{\mathrm{g}},\boldsymbol{x}_{\mathrm{h}})=L(T(\boldsymbol{x}_{\mathrm{t}}),G(\boldsymbol{x}_{\mathrm{g}}),H(\boldsymbol{x}_{\mathrm{h}})) \tag{1}$$

式中:$F(\boldsymbol{x}_{\mathrm{t}},\boldsymbol{x}_{\mathrm{g}},\boldsymbol{x}_{\mathrm{h}})$表示模型的输出,$\boldsymbol{x}_{\mathrm{t}}$、$\boldsymbol{x}_{\mathrm{g}}$和$\boldsymbol{x}_{\mathrm{h}}$分别表示文本、图形和表格数据的输入。函数$T$、$G$和$H$将这些输入转换为统一的嵌入空间。然后,大语言模型$L(\cdot)$合成这些嵌入以产生连贯可靠的输出,从而提高金融分析的准确性和可靠性。

此外,为了增强模型的泛化能力,引入了多模态对比学习目标函数。在训练过程中,模型不仅优化分类损失,还通过最大化同一时间点不同模态之间的语义一致性来约束特征空间,表达式为

$$\mathcal{L}_{\mathrm{contrastive}}=-\ln\frac{\exp(\mathrm{sim}(h_{\mathrm{T}},h_{\mathrm{I}})/\tau)}{\sum_{j=1}^{N}\exp(\mathrm{sim}(h_{\mathrm{T}},h_{\mathrm{I}}^{j})/\tau)} \tag{2}$$

式中:$\mathrm{sim}(\cdot,\cdot)$表示余弦相似度函数,$\tau$为温度系数,$N$为batch内负样本数。

最终的损失函数为多任务联合形式,即

$$\mathcal{L}=\mathcal{L}_{\mathrm{CE}}+\lambda\mathcal{L}_{\mathrm{contrastive}} \tag{3}$$

式中:$\mathcal{L}_{\mathrm{CE}}$为交叉熵分类损失,$\lambda$为调节对比损失权重的超参数。

2.4 大模型与时序模型的联合预测

本文提出的LMTSS框架结合了语言模型在语义理解与推理方面的优势,以及时序模型在历史价格模式提取方面的专长,实现了文本驱动的市场理解与结构化数据驱动的价格行为建模的有机统一。

基于时序建模的价格趋势预测模块,使用近年来较为成熟的基于门控循环单元与注意力机制的时序建模框架,主要用于从历史交易数据中建模价格行为与市场反应模式。本文采用其核心架构用于处理公司在过去k天内的价格序列、成交量序列与技术指标序列,作为结构化输入特征,并从编码后的市

场信息中递归地推断并解码潜在驱动因子和走势。生成模型中后验推断难以直接求解，于是借鉴变分自编码器思路，用神经网络拟合潜在分布，通过神经近似和重参数化规避求解难题[18-19]。

大语言模型指导的市场理解模块基于FinRobot类模型中“市场分析师”的能力进行构建。FinRobot原系统以多模态预训练语言模型为核心，配合知识库查询与工具调用机制，用于处理包括投资问答、公司研报生成、财务摘要、市场情绪分析等任务。以“市场分析师”功能为主干，重构其方法体系以适配股票走势预测任务。具体而言，定义如下任务目标：给定某支股票在时间t的多模态信息输入集合：$\mathcal{X}_t=\{T_t, \boldsymbol{A}_t, \boldsymbol{B}_t\}$。其中，$T_t$表示公司相关文本（包括新闻、公告、社交媒体内容等），$\boldsymbol{A}_t$表示技术分析图，$\boldsymbol{B}_t$表示结构化财务数据，目标是预测未来时间窗口$[t+1, t+W]$内该股票的收益率方向$y_t\in\{0,1\}$，即上涨或下跌标签。本文只将相关文本作为输入。

本文对输入大语言模型的指令进行了调整，基于图3示例的结构加入概率预测。例如，基于2024-04-19之前的所有信息，首先分析AAPL公司的积极发展和潜在问题，分别提出2～4个最重要的因素，并保持简洁，然后对AAPL在2024-04-22至2024-04-26这一周的股票价格变动趋势进行预测，明确给出上涨概率（0～1之间的小数），并提供摘要分析解释该概率的逻辑依据。

为了充分利用大语言模型的语义预测与时序模型的结构行为预测，本文出于简洁性、可解释性及易于调试的考虑，采用线性融合策略。未来工作中可探索注意力机制或神经网络融合器等更复杂的融合方式以进一步提升性能。TS模块表示时序模块，最终的预测概率表达式为

$$p_t^{\text{final}}=\alpha\cdot p_t^{\text{LLM}}+(1-\alpha)\cdot p_t^{\text{TS}} \tag{4}$$

式中：融合权重$\alpha\in[0,1]$可以根据任务需求手动设置或在验证集上自动调优，p_t^{LLM}和p_t^{TS}分别表示大语言模型和时序模型的预测结果，取值范围为$(0,1)$。本文使用手动设置的静态值，在实验部分展示不同α值对模型性能的影响。

此外，为增强模型鲁棒性与可解释性，测试阶段引入了不一致性判定，对两种模型输出差异显著的样本进行不确定性标记，可表示为：$\Delta t=|p_t^{\text{LLM}}-p_t^{\text{TS}}|$。在测试中，设置不一致性阈值$\varepsilon=0.3$，即当LLM与TS模块的概率输出差异超过30%时该样本标记为高不确定性，大模型估计结果失效，该值为本项目经验值，但阈值需依据具体任务和数据调整的思路与预警系统及质量控制等领域的研究实践保持一致。

3 实验结果

3.1 实验设置

3.1.1 度量标准与数据集

在3个基准数据集上评估本文方法，以符合当前的最先进技术：(1) StockNet[16]包括来自9个行业的87只股票，并附有2014年1月1日至2016年1月1日的股票相关推文和历史价格数据，涉及美国股市；(2) CMIN-US[20]包含美国股市中排名前110的股票及其推文和2018年1月1日至2021年12月31日的历史价格数据；(3) CMIN-CN[20]由CSI300指数中的300只股票组成，附带其推文和2018年1月1日至2021年12月31日的历史价格数据，涉及中国股市。统计量如表1所示，每只股票都由包含推文和股票价格的时间序列数据表示。数据集按时间划分为训练集、开发集和测试集，训练后的模型在测试集上预测未来某个时间点股价相对于当前是“上涨”还是“下跌”，是一个二分类任务。

根据之前研究，将准确率（Accuracy, ACC）、马修斯相关系数（Matthews correlation coefficient, MCC）和F_1分数作为评估指标[21-22]。准确率、马修斯相关系数和F_1分数分别定义为

$$\text{ACC}=\frac{\text{tp}+\text{tn}}{\text{tp}+\text{tn}+\text{fp}+\text{fn}} \tag{5}$$

$$\mathrm{MCC}=\frac{\mathrm{tp}\cdot\mathrm{tn}-\mathrm{fp}\cdot\mathrm{fn}}{\sqrt{(\mathrm{tp}+\mathrm{fp})(\mathrm{tp}+\mathrm{fn})(\mathrm{tn}+\mathrm{fp})(\mathrm{tn}+\mathrm{fn})}} \tag{6}$$

$$F_1=\frac{2\times\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}} \tag{7}$$

式中:tp、fn、fp、tn为混淆矩阵的元素,Precision为精准率,Recall为召回率。

表 1 基准数据集统计量

Table 1 Statistics of benchmark datasets

数据集	StockNet	CMIN-CN	CMIN-US
数据类型	—	时间序列与文本	—
数据来源	—	股票价格序列和新闻推文	—
数据大小	19 318	83 553	198 781
股票市场	美国	中国	美国
股票数量	87	300	110
数据范围	2014-01-01至2016-01-01	2018-01-01至2021-12-31	2018-01-01至2021-12-31

3.1.2 训练设置与基线模型

设定模型的输入窗口为5个交易日,采用批量大小为32的样本随机打乱训练。为控制模型复杂度,单条文本的最大Token数限制为30,每个交易日最多处理40条消息,超出部分将被截断。词嵌入维度为50(控制内存消耗,适配11 GB GPU);消息嵌入层隐藏维度100,变分走势解码器隐藏维度150。所有权重矩阵用fan-in技巧初始化,偏置项初始化为0;使用Adam优化器[23],初始学习率0.001。输入dropout率为0.3(正则化潜在变量)[24];采用TensorFlow框架[25]构建计算图,超参数在验证集上调优。

(1) 大语言模型。文本编码器FinBERT使用预训练参数进行微调,ViT采用公开的ImageNet预训练权重。模型使用Adam优化器,初始学习率设为1e−4,batch size为64,训练周期为30轮。超参数λ设置为0.3,温度系数τ=0.05。

(2) 基于时序的基准模型。Random[26]:随机猜测股票涨跌的简单预测器;ARIMA[27]:仅用价格信号的自回归积分滑动平均模型;RANDFOREST[28]:使用Word2vec文本表示的判别式随机森林;TSLDA[29]:联合学习主题和情感的生成式主题模型;HAN[30]:带层次注意力的判别式深度神经网络CMIN[31]:一种端到端的深度神经网络,建模金融文本数据与因果关系增强的股票相关性之间的多模态性。

(3) 基于情感的基准模型。EDT[32]:利用基于BERT的双层事件检测框架对影响股票价格的企业事件进行分类;FinGPT[33]:对大型语言模型进行指令微调,以增强金融情感分析能力;GPT-4-turbo[2]:建立在GPT-4的基础上,提高了效率和性能,特别是在处理较长文本时;GPT-4[2]:相比于之前的版本,实现了显著的前向飞跃,具有更快的处理速度和更好的长文本输入处理能力;GPT-3.5-turbo[2]:以其全面的理解能力和类人文本生成能力而闻名,优化了多任务处理;RoBERTa[34]:应用Financial RoBERTa2进行英语文本的情感分析,并使用针对中文文本微调的情感特定RoBERTa模型;FinBERT[9]:一种经过微调的BERT模型,用于金融情感分类,提供英语和中文版本。

3.2 实验结果与分析

由于股票预测是一项具有挑战性的任务,而其预测性能的提升通常会带来收益,因此在二元股票价格的变动预测中,56%的准确率通常被认为是令人满意的结果[29]。表2展示了基于时序的基准方法和本文方法的性能,其中RANDFOREST、TSLDA和CMIN方法的F_1值未在相关文献中报告。由表2

可以看出,本文方法表现最佳,准确率达到63.54%,MCC为0.193,F_1为60.3。相较于使用的时序方法StockNet,本文方法在ACC、MCC和F_1值关键指标上均优于对比方法,体现出更强的预测性能,尤其是在ACC上提高5.31个百分点。

表2展示了本文方法及各种基于情感的基准方法在3个数据集上的结果。本文方法在所有基准数据集上的表现均优于当前的先进方法,显示出其更强的泛化能力与稳定性。此外,基于情感的方法在不同方法之间表现不一,例如GPT-4和GPT-4-turbo等大语言模型在识别文本情感方面表现出色。尽管EDT方法在其特定数据集上达到了最高的准确率,但其较低的MCC表明方法性能不平衡。由于基于时序的方法考虑了文本的整体而没有进行详细分析,它们的性能与基于情感的方法相当,后者可能缺乏时间上下文,但包含了详细的文本信息。而本文方法不仅能从文本数据中识别出影响股价的重要因素,还能结合关系和时间信息,从而增强了其过滤无关内容的能力,可以提供更全面的分析。

表2 本文方法与时序基线方法的比较结果

Table 2 Comparison results of the time series baseline methods and the proposed method

方法	ACC/%	MCC	F_1
RAND	50.89	−0.002	50.2
ARIMA	51.39	−0.021	51.3
RANDFOREST	53.08	0.012	—
TSLDA	54.07	0.065	—
HAN	57.64	0.051	57.2
StockNet	58.23	0.081	57.5
MAN-SF	60.80	0.195	60.5
CMIN	62.69	0.209	—
本文方法	63.54	0.193	60.3

表3 本文方法与基于情感方法的比较结果

Table 3 Comparison results of the methods based on emotion and the proposed method

方法	StockNet		CMIN-CN		CMIN-US	
	ACC/%	MCC	ACC/%	MCC	ACC/%	MCC
EDT (ACL2021)	40.31	−0.066	49.86	−0.004	40.00	0.021
FinGPT	54.91	0.083	59.98	0.182	55.78	0.120
GPT-4-turbo	53.56	0.060	64.61	0.284	56.94	0.135
GPT-4	53.88	0.062	62.18	0.260	56.96	0.136
GPT-3.5-turbo	52.31	0.044	56.10	0.156	56.68	0.124
RoBERTa	54.46	0.088	57.75	0.138	52.24	0.064
FinBERT	55.42	0.111	58.26	0.158	55.98	0.121
本文方法	63.54	0.193	65.05	0.223	57.80	0.191

此外,在表2中本文方法的MCC略低于CMIN和MAN-SF,表明在部分类别上的预测平衡性仍有提升空间。同理,在表3中本文方法在CMIN-CN上的MCC低于GPT-4-turbo,说明通用LLM在特定语言或市场数据的建模上仍具优势。

3.3 消融实验

3.3.1 模块消融实验

为了进一步验证本文方法中各关键模块对整体性能的贡献,设计了消融实验,依次剔除某一模块后重新训练模型,观察预测性能变化。具体消融设置包括以下3个模块:LLM大语言模型模块,TS时序模块和Fusion融合模块。StockNet数据集上的消融实验结果如表4所示,其中对应空格打勾则表示使用,反之不使用。从消融实验结果可以观察到:LLM模块和TS模块的缺失会显著降低模型整体表现,完整方法在所有数据集上表现最优,验证了各模块协同设计的有效性。

表4 StockNet数据集上的模块消融实验

Table 4 Module ablation experiments on the StockNet dataset

LLM	TS	Fusion	ACC/%	MCC
√			54.91	0.083
	√		58.23	0.081
√	√	√	63.54	0.193

3.3.2 融合权重消融实验

为探究大LLM语义通路与TS通路在最终决策中的相对贡献,进行融合权重α的消融实验。表5展示了α在0、0.2、0.5、0.8时的ACC与MCC结果,可以看到模型在$\alpha=0.5$时表现最优,模型的性能随着α的变化呈现出明显的先升后降的趋势。当α过小或过大时,模型会因忽视文本语义或视觉特征而导致性能受损。最终选择在验证集上取得最高准确率的$\alpha=0.5$作为模型的配置。这一结果表明,对等的权重分配在本任务中能够最有效地实现模态间的信息互补。尽管更精细的搜索算法(如贝叶斯优化)可能发现更优解,但当前基于网格搜索的简单策略已在性能和计算成本间取得了良好平衡,并清晰地揭示了模态权重的影响规律。

表5 StockNet数据集上α消融实验结果

Table 5 Experimental results of α ablation on the StockNet dataset

α值	0	0.2	0.5	0.8
ACC/%	54.91	56.83	63.54	63.65
MCC	0.083	0.091	0.193	0.154

3.4 市场分析验证结果

本节通过示例应用验证了本文方法的实际有效性。市场报告旨在综合最新的市场新闻和金融数据,提供关于公司最新成就和潜在问题的全面见解,以及对股价走势的预测。图5分别展示了StockNet、CMIN-CN、CMIN-US数据集中的案例,目标股票为APPLE、平安银行和Google,括号内为股票代码。示例市场分析报告展示了本文方法在整合多源数据方面的能力,能够生成具有洞察力的判断,并

分析对象:APPLE (AAPL)
分析时间范围:2015年12月31日

[正面发展]
1. iPhone 6s 销售预期:有观点认为市场对iPhone 6s的销售情况过于担忧,但实际上安卓阵营在触觉反馈方面的进展可能不会过于激进,这可能为苹果的销售表现提供支撑。
2. OLED 显示屏谈判:苹果正在讨论新的iPhone显示屏,这可能提升其产品竞争力,尤其是在显示技术上的创新可能吸引消费者。

[潜在担忧]
1. 股价表现不佳:苹果股价在2015年的表现可能不佳,接近自2008年以来的最差表现,这可能引发投资者对长期增长的担忧。2. 税务纠纷:苹果与意大利的税务纠纷可能影响其国际业务扩展和声誉,尽管最终达成和解,但潜在的税务风险仍需关注。

[股票价格预测]
基于当前信息,预计苹果股票在接下来一周可能会有轻微下跌,幅度在1%至2%之间。这一预测主要考虑到市场对iPhone 6s销售的担忧以及潜在的税务问题,尽管OLED显示屏的进展可能在一定程度上缓解这些压力。

(a) StockNet

分析对象:平安银行(000001)
分析时间:2021年12月31日

[正面发展]
1. 科技创新:平安银行的信用卡"A+"新核心系统获得央行和证监会的认可,并荣获金融科技发展奖一等奖。这表明公司在科技创新方面取得了显著进展,有助于提升其在行业中的竞争力。
2. 智能化升级:平安银行连续两年蝉联亚太最佳手机银行APP奖,显示出其在移动金融服务领域的领先地位,进一步增强了用户体验和市场认可度。

[潜在担忧]
1. 供应链金融风控:平安银行通过"卫星上天"项目解决了供应链金融风控难题,但该技术的成熟度和广泛应用仍需时间验证。
2. 个人不良贷款批转业务:岁末回顾显示,个人不良贷款批转业务可能存在一定的风险和不确定性,需关注其对银行资产质量的影响。

[股票价格预测]
综合考虑科技创新带来的增长潜力和市场对供应链金融风控技术的关注,预计平安银行股票价格在下周将呈现小幅上涨趋势,涨幅约为1%~2%。

(b) CMIT-CN

分析对象:Google (GOOG)
分析时间范围:2021年12月31日

[正面发展]
1. 行业增长与技术创新。Big Tech公司在2021年实现了显著增长,为全球经济贡献了2.5万亿美元的价值。作为科技巨头之一,Google在技术创新和市场扩展方面表现突出,尤其是在人工智能、云计算和自动驾驶等领域。2. 市场需求与投资信心。多篇新闻提到投资者对科技股的长期信心,尤其是Nasdaq指数在2021年的表现强劲,表明市场对科技公司的未来增长持乐观态度。Google作为Nasdaq的重要成分股,受到投资者的关注。3. 竞争格局与潜在机会。私空投资领域的快速发展为科技公司带来了新的增长点,尤其是SpaceX等公司的成功吸引了大量投资。Google在太空技术领域的布局可能在未来带来新的收入来源。

[潜在担忧]
1. 监管压力与反垄断风险。Meta等科技公司面临的反垄断调查和监管压力成为2022年的主要风险之一。Google作为另一家科技巨头,也可能面临类似的监管审查,这可能对其股价产生负面影响。2. 竞争加剧与技术追赶。特斯拉的竞争对手计划在2024年前推出完全自动驾驶汽车,这表明科技行业的竞争正在加剧。Google的自动驾驶技术(Waymo)虽然领先,但需要持续投入以保持竞争优势。3. 市场波动与宏观经济因素。2021年加密货币市场的波动对整体科技股产生了影响。2022年全球经济的不确定性可能进一步加剧市场的波动,影响Google的股价表现。

[股票价格预测]
基于上述因素,预计Google股票在接下来一周可能会呈现小幅波动。考虑到市场对科技股的整体乐观情绪以及Google在技术创新和市场扩展方面的优势,预计其股价可能会上涨1%~2%。然而,监管压力和竞争加剧可能对股价产生一定的抑制作用,因此整体预测为上涨1%~2%。

(c) CMIT-US

图5 StockNet、CMIN-CN、CMIN-US数据集中公司的市场预测

Fig.5 Market forecasts for companies in the StockNet, CMIN-CN, and CMIN-US datasets

辅助投资者理解股票走势的潜在驱动因素。此外,市场预测工具还提供了关于股票未来轨迹的建议,强调了其基于分析数据提供可操作指导的能力。

图6分别展示了StockNet、CMIN-CN、CMIN-US数据集中涨跌概率预测的案例,通过修改输入大语言模型的指令来实现,并和一般时序模型共同对股票价格涨跌概率进行预测。

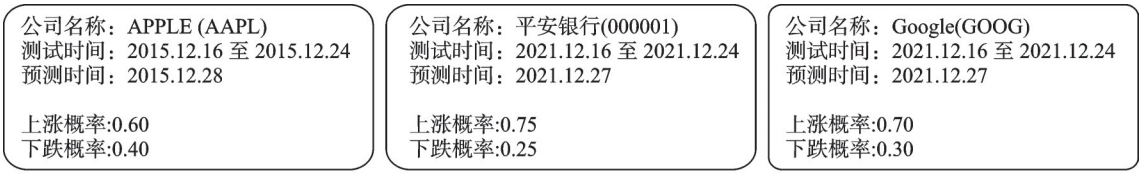

图6 StockNet、CMIN-CN和CMIN-US数据集涨跌概率预测

Fig.6 Probability prediction of rise and fall in the StockNet, CMIN-CN and CMIN-US datasets

4 结束语

本文提出了一个大语言模型指导的多模态时序-语义预测框架。该框架通过“时序-语义-决策”的三元协同机制,实现了跨模态一致性约束、动态权重融合与可解释推理,从而在复杂异构场景下提供稳健的时序建模、跨模态语义对齐与可解释推理。实验表明,本文方法在真实世界数据上取得了优于基线的方法性能,并能生成自然语言解释,提升了决策透明度与可信度。未来工作将从3个方面推进:(1)轻量化与高效推理:通过知识蒸馏与参数高效微调,实现低延迟与低成本部署;(2)隐私与安全:引入联邦学习与鲁棒训练,降低数据泄露与对抗性风险;(3)跨域迁移与可扩展性:验证在更多时序多模态场景下的泛化能力,并完善工具调用与外部知识整合的工程化支撑。

参考文献:

[1] YANG H, ZHANG B Y, WANG N, et al. FinRobot: An open-source AI agent platform for financial applications using large language models[EB/OL]. (2024-05-23). https://arxiv.org/pdf/2405.14767.pdf.

[2] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: ACM, 2020: 1877-1901.

[3] DEVLIN J, CHANG M W, KENTON L, et al. 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, Volume 1 (Long and Short Papers). Minnesota, USA: ACL, 2019: 4171-4186.

[4] RAFFEL C, SHAZEER N, ROBERTS A, et al. Exploring the limits of transfer learning with a unified text-to-text transformer[J]. Journal of Machine Learning Research, 2020, 21(1): 5485-5551.

[5] WEI J, TAY Y, RISHI B, et al. Emergent abilities of large language models[EB/OL]. (2022-06-15). https://arxiv.org/pdf/2206.07682.pdf.

[6] 赵睿卓, 曲紫畅, 陈国英, 等. 大语言模型评估技术研究进展[J]. 数据采集与处理, 2024, 39(3): 502-523.
ZHAO Ruizhuo, QU Zichang, CHEN Guoying, et al. Research progress in evaluation techniques for large language models[J]. Journal of Data Acquisition and Processing, 2024, 39(3): 502-523.

[7] OUYANG L, WU J, XU J, et al. Training language models to follow instructions with human feedback[C]//Proceedings of the 36th International Conference on Neural Information Processing Systems. New Orleans, USA: ACM, 2022: 27730-27744.

[8] PETERS M, NEUMANN M, IYYER M, et al. Deep contextualized word representations[C]//Proceedings of the 2018 Conference of the North American Chapter Ofthe Association for Computational Linguistics: Human LanguageTechnologies, Volume 1 (Long Papers). New Orleans, USA: ACL, 2018: 2227-2237.

[9] YANG Y, MARK C S U Y, ALLEN H. FinBERT: A pretrained language model for financial communications[EB/OL]. (2020-07-09). https://arxiv.org/pdf/2006.08097.pdf.

[10] BALTRUŠAITIS T, AHUJA C, MORENCY L P. Multimodal machine learning: A survey and taxonomy[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(2): 423-443.

[11] NGIAM J, KHOSLA A, KIM M, et al. Multimodal deep learning[C]//Proceedings of the International Conference on Machine Learning. Washington, USA: [s.n.], 2011.

[12] ALAYRAC J B, DONAHUE J, LUC P, et al. Flamingo: A visual language model for few-shot learning[C]//Proceedings of the Advances in Neural Information Processing Systems. New Orleans, USA: [s.n.], 2022, 35: 23716-23736.

[13] JIASEN L, DHRUV B, DEVI P, et al. ViLBERT: Pretraining task-agnostic visiolinguistic representations for vision-and-language tasks[C]//Proceedings of the Advances in Neural Information Processing Systems. New Orleans, USA: [s.n.], 2019: 13-23.

[14] JAEMIN C, JIE L, HAO T, et al. Unifying vision-and-language tasks via text generation[C]//Proceedings of the International Conference on Machine Learning. Vancouver, Canada: PMLR, 2021: 1931-1942.

[15] ADITYA R, MIKHAIL P, GABRIEL G, et al. Zero-shot text-to-image generation[C]//Proceedings of the International Conference on Machine Learning. Vancouver, Canada: PMLR, 2021, 139: 8821-8831.

[16] XU Y, COHEN S B. Stock movement prediction from tweets and historical prices[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Melbourne, Australia: ACL, 2018: 1970-1979.

[17] SAWHNEY R, MATHUR P, MANGAL A, et al. Multimodal multi-task financial risk forecasting[C]//Proceedings of the 28th ACM International Conference on Multimedia. Seattle, USA: ACM, 2020: 456-465.

[18] DIEDERIK P K, MAX W. Auto-encoding variational bayes[C]//Proceedings of the International Conference on Learning Representations. Banff, Canada: [s.n.], 2014.

[19] REZENDE D J, MOHAMED S, WIERSTRA D, et al. Stochastic backpropagation and approximate inference in deep generative models[C]//Proceedings of the 31st International Conference on International Conference on Machine Learning-Volume 32. Beijing, China: ACM, 2014: 1278-1286.

[20] LUO D, LIAO W, LI S, et al. Causality-guided multi-memory interaction network for multivariate stock price movement prediction[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Toronto, Canada: ACL, 2023: 12164-12176.

[21] XIE B Y, REBECCA J P, LEON W, et al. Semantic frames to predict stock price movement[C]//Proceedings of the 51st Annual Meeting of the Association for Computational linguistics. Sofia, Bulgaria: ACL, 2013: 873-883.

[22] DING X, ZHANG Y, LIU T, et al. Deep learning for event-driven stock prediction[C]//Proceedings of the 24th International Joint Conference on Artificial Intelligence. Buenos Aires, Argentina: ACM, 2015: 2327-2333.

[23] KINGMA D P, BA J. Adam: A method for stochastic optimization [EB/OL]. (2014-12-22). https://arxiv.org/pdf/1412.6980.pdf.

[24] BOWMAN S R, VILNIS L, VINYALS O, et al. Generating sentences from a continuous space[C]//Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning. Berlin, Germany: [s.n.], 2016: 10-21.

[25] ABADI M, AGARWAL A, BARHAM P, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems[EB/OL]. (2016-03-16). https://arxiv.org/pdf/1603.04467.pdf.

[26] ROBERT G B. Smoothing, forecasting and prediction of discrete time series[M]. New York, USA: Dover Publications, 2004.

[27] BROWN, ROBERT G. Smoothing, forecasting and prediction of discrete time series[M]. New York, USA: Courier Corporation, 2004.

[28] PAGOLU V S, REDDY K N, PANDA G, et al. Sentiment analysis of Twitter data for predicting stock market movements [C]//Proceedings of the 2016 International Conference on Signal Processing, Communication, Power and Embedded System. Paralakhemundi, India: IEEE, 2016: 1345-1350.

[29] NGUYEN T H, SHIRAI K. Topic modeling based sentiment analysis on social media for stock market prediction[C]//

Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Beijing, China: ACL, 2015: 1354-1364.

[30] HU Z, LIU W, BIAN J, et al. Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction[C]//Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. California, USA: ACM, 2018: 261-269.

[31] LUO D, LIAO W H, LI S Q, et al. Causality-guided multi-memory interaction network for multivariate stock price movement prediction[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Toronto, Canada: ACL, 2023.

[32] ZHOU Z H, MA L Q, LIU H. Trade the event: Corporate events detection for news-based event-driven trading[EB/OL]. (2021-05-28). https://arxiv.org/pdf/2105.12825.pdf.

[33] ZHANG B Y, YANG H Y, LIU X Y. Instruct-FinGPT: Financial sentiment analysis by instruction tuning of general-purpose large language models[EB/OL]. (2023-06-22). https://arxiv.org/pdf/2306.12659.pdf.

[34] ZHANG J X, GAN R Y, WANG J J, et al. Fengshenbang1.0: Being the foundation of Chinese cognitive intelligence[EB/OL]. (2023-03-30). https://arxiv.org/pdf/2209.02970.pdf.

作者简介:

叶诗敏(1992-),女,硕士,研究方向:工商管理、市场经济,E-mail: yeshimin@126.com。

刘非菲(1979-),**通信作者**,女,博士,副教授,研究方向:财务管理,E-mail: liuff@szut.edu.cn。

张岩(1987-),男,博士,工程师,研究方向:计算机视觉研究、大语言模型研究。

(编辑:张黄群,王婕)