arXiv:2603.12500v1 [cs.CE] 12 Mar 2026

TRACE：面向可解释股票走势预测的知识图谱时间规则锚定证据链

Qianggang Ding[1,2], Haochen Shi[1,2], Luis Castejón Lozano[3], Miguel Conner[3], Juan Abia[3], Luis Gallego-Ledesma[3], Joshua Fellowes[3], Gerard Conangla Planes[3], Adam Elwood[4], Bang Liu[1,2,5]

[1] DIRO & Institut Courtois，蒙特利尔大学，蒙特利尔，魁北克省，加拿大[2] Mila—魁北克人工智能研究所，蒙特利尔，魁北克省，加拿大[3] Aily Labs，马德里，西班牙[4] Aily Labs，米兰，意大利[5] 加拿大CIFAR人工智能讲席

{qianggang.ding, haochen.shi, bang.liu}@umontreal.ca

{luis.castejon, miguel.conner, juan.abia, luis.gallego, joshua.fellowes, gerard.conangla, adam.elwood}@ailylabs.com

摘要

我们提出了一种面向知识图谱的时间规则锚定证据链（TRACE），用于实现可解释的股票走势预测。该方法在单一端到端流程中统一了符号关系先验、动态图探索和大语言模型引导的决策。该方法通过规则引导的多跳探索，仅沿允许的关系序列进行搜索；将候选推理链与同期新闻关联，并将有充分依据的证据汇总为可审计的上涨/下跌判断，同时提供连接文本与结构的可读路径。在标准普尔500指数基准测试中，该方法的准确率为55.1%、精确率为55.7%、召回率为71.5%、F1值为60.8%，优于表现强劲的基线方法；在相同评估条件下，其召回率和F1值也高于最佳图基线方法。性能提升源于两点：（i）规则引导的探索将搜索重点放在具有经济意义的模式上，而非任意游走；（ii）基于文本依据的整合有选择地汇总置信度高且有充分依据的假设，而不是对弱信号一概进行汇总。这些设计相辅相成，在不牺牲选择性的前提下提高了敏感度，从而提升预测性能，并提供忠实、可审计且可解释的说明。

CCS 概念

•**信息系统** → **数据挖掘**；信息抽取；检索模型和排序；• **计算方法** → **机器学习**；人工智能。

关键词

时序知识图谱、股票走势预测、规则挖掘、多跳推理、大语言模型

1 引言

预测股价走势长期以来一直是计算金融与数据挖掘中的核心问题。经典方法 [8,11, 25, 29, 48] 依赖于对历史价格和宏观经济指标的时间序列建模，而较新的方法 [9, 30, 44]则纳入从财经新闻和社交媒体中提取的文本情感。尽管已取得显著进展，现有模型在两个关键维度上仍然存在局限。首先，它们通常将异构信息源独立处理，聚合特征而不对公司、事件和产品之间的关系依赖进行建模。其次，基于大语言模型（LLM）的方法虽然在自然语言理解方面很强大，但往往会幻觉出缺乏依据的事实，并且缺乏可解释性，从而在高风险金融领域带来风险。

金融市场本质上具有关系性：单只股票的走势很少仅取决于其自身的基本面或市场情绪，而是受到复杂交互网络的影响。合并、收购、合作伙伴关系、供应链中断、管理层变动或监管裁决等公司行动，会在不同行业和地理区域间传播，不仅影响直接涉事公司，也会影响竞争对手、供应商和客户。例如，人工智能行业中的一项收购可能推高收购方公司的股价，使上游硬件供应商受益，并给竞争对手公司带来竞争压力。同样，针对某家制药公司的监管调查也可能对其合作伙伴和合同制造商产生负面影响。

知识图谱已在 Web 检索领域得到广泛研究 [15, 22, 33], ，尤其是在问答、链接预测和多跳推理方面。现有金融预测模型很少整合显式图推理，因此错失了在统一框架中将非结构化文本（如新闻）与结构化关系（如投资、合作伙伴关系）联系起来的机会。通过将知识图谱推理引入金融预测，我们旨在弥合这一差距，并捕捉市场内在的关系动态。

据我们所知，本文提出了**首个融合大语言模型、用于股票走势预测的时序知识图谱增强框架**。我们的方法并非对现有推理范式的简单组合，而是一种统一方法论，将符号先验、动态图探索和大语言模型引导的决策整合为连贯的预测流程。

我们框架的核心包含以下两个组成部分：**(1) 规则引导的关系先验**。我们从历史知识图谱快照中提取关系模体，以捕捉反复出现的金融交互（例如，锚定于新闻的投资活动，以及与因果事件相关联的合作伙伴关系）。这些挖掘出的模式并非作为独立规则，而是为后续探索界定搜索空间和结构偏置。通过

限定哪些关系序列可以采用，它们确保推理路径反映具有金融意义的依赖关系。**(2) 基于图的推理与文本锚定**。基于这些先验，我们设计了一种引导式探索流程，从每只标准普尔500指数成分股出发，在日期 t遍历金融知识图谱，仅沿着时间有效的关系前进。候选推理路径将股票与事件、交易对手相连，并最终指向新闻或申报文件等TextSource节点。每条路径均通过多因素函数评分，该函数综合考虑规则覆盖、时间新近性、结构多样性，以及大语言模型关系选择器的反馈。此过程生成的证据链既具有语义相关性，又以明确的文本证据为锚点。

这些组件的整合所产生的效果不止是各部分效果的简单相加。规则引导先验能够减少噪声并防止任意探索，而图推理则将这些先验付诸实践，以发现多样化且有新闻依据的证据链。最终预测阶段将路径级信号汇总为二元上涨/下跌预测，每个预测均由明确的推理路径支持。这种统一设计同时实现了 预测准确率和可解释性，解决了金融预测中性能与透明度之间长期存在的权衡问题。

本文主要有三项贡献：

(1) 我们提出了一个名为 TRACE 的统一框架，将可解释股票走势预测嵌入金融时序知识图谱中，并整合符号路径挖掘、大语言模型引导的图探索和经大语言模型验证的决策制定。

(2) 我们设计了一种规则引导且以文本为依据的推理流程，其中挖掘出的关系模体约束可接受的路径，候选推理链则以新闻文章为依据，从而确保经济相关性和时间有效性。

(3) 我们基于2022–2023年标准普尔500指数数据构建了一个大规模金融知识图谱（42K个节点、174K条边），并在其上进行了广泛实验。结果表明，我们的框架优于强大的基线方法，同时能够提供由明确推理路径支持的可解释上涨/下跌预测结果。

2 相关工作

金融预测中的图推理。 传统的时间序列或基于文本的金融预测模型通常将每只股票视为独立对象，忽略了跨行业传播风险与信息的公司间依赖网络 [1, 2, 5, 10, 13, 47, 53] 。

而基于图的建模通过显式捕捉公司、产品和事件等实体之间的关系依赖，为此提供了解决方案 [1, 2, 5, 10, 13, 47, 53]。早期的金融图模型，如 HATS [19] 和 RSR [12],，引入了基于公司与新闻之间共现或文本相关性构建的多关系图。相比之下，后续研究（包括 MDGNN [31] 和 FEEKG [24],）显式建模事件的时序层，以预测股票趋势。另一方面，金融市场中的公司间依赖和宏观事件持续演变，这与基于时序知识图谱（TKG）的建模天然契合 [7, 17, 37]。近来，大语言模型已被用于增强金融图谱的构建和推理能力。例如，FinDKG [20] 利用基于大语言模型的抽取来构建动态金融知识图谱，而 FinGPT [40]或 BloombergGPT [41] 则为金融文本理解提供了强大的语言先验。然而，这些方法在很大程度上依赖潜在嵌入融合或基于文本的表征学习。相比之下，我们的框架在金融知识图谱上执行显式时序推理。**规则学习与神经符号推理**。尽管基于图的建模能够捕捉时序依赖，但仍严重依赖隐式关系嵌入。这类方法对于关系如何驱动预测的可解释性有限。为增强结构可解释性，另一类研究聚焦于显式规则学习与神经符号推理。传统的归纳规则学习框架（如 AMIE [14], 、AnyBURL [28] ）专注于从知识图谱中挖掘频繁关系模式，以提供透明的推理逻辑规则。然而，可微方法将规则学习与基于梯度的优化相结合，从而衔接符号与神经范式 [32, 46]。近期研究则探索将规则归纳与预训练语言模型相结合的混合系统，以增强对知识图谱的推理能力 [34]。与这些研究不同，我们的框架将挖掘出的关系规则视为时序金融知识图谱中的动态搜索先验。我们不采用静态规则或融合嵌入，而是利用规则在符合时间约束的前提下限制可接受的推理路径，从而确保因果有效性。

大语言模型与知识图谱混合推理。在基于规则和神经符号推理取得进展的基础上，越来越多的研究开始探索将大语言模型与结构化知识图谱相结合，以增强事实依据和推理可解释性 [4, 16, 26, 27, 36, 39, 43]。例如，Think-on-Graph (ToG) [27, 36] 将大语言模型用作推理智能体，通过束搜索 [3, 18, 23, 35] 迭代遍历知识图谱，并构建显式关系链，以减少幻觉。图约束推理 (GCR) [26]进一步通过知识图谱引导的字典树约束大语言模型解码，从而提升结构忠实度。另有研究通过将大语言模型推理与知识图谱衍生的知识适配相结合，拓展了这一范式。例如，Chain-of-Knowledge (CoK) [52]通过训练模型在挖掘规则数据集上沿结构化路径进行推理，使大语言模型推理与知识图谱路径相对齐。GenTKG [21] 将这一思路拓展至时间领域，将检索增强生成与时间感知关系推理相结合，从而实现对演化知识图谱的生成式预测。基于这一思路，RAG-KG-IL [49] 在多智能体框架中整合检索增强生成与增量知识图谱学习，以促进异构来源之间的持续知识更新和自适应推理。相比之下，我们的框架在时间金融知识图谱上统一了符号先验、动态图探索和大语言模型引导的决策。

3 方法论

3.1 问题形式化

我们研究使用动态图知识图谱（KGs）进行股票走势预测的任务。令 $G_t = (E_t, R_t, F_t)$ 表示交易日 t的时序知识图谱快照，其中：

- E_t是实体集合，包括文本来源、事件、产品、金融、公司和人物节点；

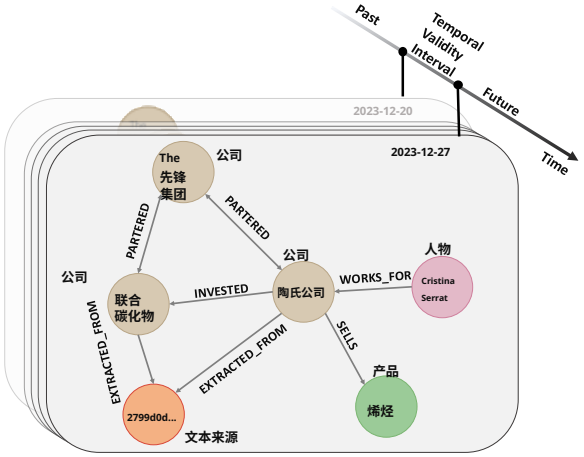

图1：时序金融知识图谱的结构，展示了实体类型、关系类别和时间有效区间。

- R_t 是关系类型集合，包括从.....提取、销售、投资_于、合作、起诉等；
- $F_t \subseteq E_t \times R_t \times E_t$ 表示在时间 t 有效的时序三元组集合 $(e_h,\ r,\ e_t)$。

每只股票 $s \in S$（其中 S 是标普500成分股的集合）均由一个公司节点表示，并通过映射函数与其股票代码关联

$$\pi : \mathcal{E}_t \to \mathcal{T}, \quad \pi(s) = \text{ticker of stock } s.$$

文本来源与截至时点约束。我们用 X 表示 TextSource节点（财经新闻、申报文件、报告）的集合。每个 $x \in X$ 都有关联的发布日期 $\tau(x)$。为避免未来信息泄漏，在构建 G_t 时，只能访问时间戳为 $\le t$ 的边和事实。形式上，三元组 $(e_h,\ r,\ e_t) \in F_t$ 有效，当且仅当

$$\text{valid_from}(e_h,\ r,\ e_t) \le t \wedge (\text{valid_to}(e_h,\ r,\ e_t) \ge t \text{或者 valid_to } = \varnothing).$$

规则。 我们假设存在一个通过自动化规则挖掘方法预先挖掘得到的规则库 $R^* = \{\rho_1, \ldots, \rho_M\}$。每条规则 ρ 都由一个体模式和一个预测头定义：

$$\rho : \underbrace{r_1(X, Z_1) \wedge r_2(Z_1, Z_2) \wedge \cdots \wedge r_L(Z_{L-1}, Y)}_{\text{body}} \Rightarrow \underbrace{\text{predict_direction}(X, \ell)}_{\text{head}},$$

其中，每条规则的 $r_i \in R_t$ 均为 $\ell \in \{\text{UP,DOWN}\}$。这些规则作为标签预测先验，用于高效探索路径并进行评分，其置信度值由历史验证明确给出。

预测任务。 对于评估窗口 $[t_{\text{start}}, t_{\text{end}}]$内的每个交易日 t，以及每只股票 $s \in S$，模型都会预测该股票在下一个交易日是上涨 还是下跌 。

令 $p_{s,t,1}$ 表示股票 s 从第 t 天到第$t+1$ 天（下一个交易日）的前向收益率，根据调整后收盘价计算：

$$p_{s,t,1} = \frac{\text{adj_close}_{s,t+1} - \text{adj_close}_{s,t}}{\text{adj_close}_{s,t}}.$$

真实标签定义如下：

$$y_{s,t,1} = \begin{cases} \text{UP}, & \text{if } p_{s,t,1} > 0, \\ \text{DOWN}, & \text{if } p_{s,t,1} \le 0. \end{cases}$$

目标。 我们的知识图谱推理框架旨在学习一个推理函数

$$f : (\mathcal{G}_t, s, \mathcal{R}^*, h) \mapsto \hat{y}_{s,t,h},$$

该函数输出预测 $\hat{y}_{s,t,h} \in \{\text{UP, DOWN}\}$，以及从 G_t中提取的一组可解释推理路径 $P_{s,t,h}$。

3.2 知识图谱构建

我们的方法论依托于一个动态金融知识图谱（见图1），将异构数据源整合为统一的时序图。该知识图谱旨在捕捉结构化的企业关系和非结构化文本信号（如新闻、申报文件），同时保留时间有效性。

实体类型。 该知识图谱包含多个类别的实体，每个实体都由唯一标识符唯一标识：

- **文本来源** （10，685 个节点）： 金融新闻文章、财报和申报文件，每项均关联一个发布时间戳。
- **事件** （9，174 个节点）： 离散的企业或市场事件，例如合并、诉讼或监管行动。
- **产品** （8，973 个节点）： 公司产品、技术和服务。
- **金融** （5，991 个节点）： 标准普尔500成分股以外的金融工具、指标或资产。
- **公司** （4,865个节点）： 公司实体，包括标准普尔500指数成分股及其交易对手。
- **人物** （2,500个节点）： 与市场事件相关的关键高管、创始人及个人。

关系。 边表示实体之间的类型化关系。每条关系存储为一个时序三元组（$e_h, r, e_t, t_{\text{from}}, t_{\text{to}}$），其中$e_h, e_t \in E$，$r \in R$，且$[t_{\text{from}}, t_{\text{to}}]$为有效期区间。关系集合包括：

- 提取_自（125,276条边）： 将实体（例如公司、事件、产品）锚定至其原始文本来源，从而使结构化节点以文本证据为依据。
- 销售（13,003）和投资_于（11,706）： 刻画产品销售关系与投资活动。
- 导致_上升（8,250）和导致_下降（3,891）： 编码从文本中提取的事件对财务影响的因果关系。
- 合作 (3，933)、剥离 (3，694) 和收购 (1，004)：代表企业重组和合作伙伴关系。
- 任职_于 (3，012)：将人物节点与其关联公司相连。
- 许可 (415)、和解 (153) 和起诉 (78)：描述法律、合同和许可关系。

文本锚定。 每个 TextSource节点 x 都包含元数据，包括标题、出版方、URL 和发布日期 $\tau(x)$。在 x 中提及的实体通过 EXTRACTED_FROM 边相连，使推理路径能够以明确的文本证据为依据。



(a) 知识图谱构建 **(b) 自动化规则挖掘** **(c) 挖掘规则示例**

图2：我们的知识图谱推理框架概览，展示了自动化规则挖掘与时间图推理的集成。

规则集成。 通过自动规则提取挖掘得到的规则库 R^* 被集成到知识图谱推理过程中。虽然规则并未直接嵌入 G_t，但其关系模式会约束搜索轨迹，并在探索过程中引导评分。

3.3 自动化规则挖掘

为指导框架中的高效推理，我们整合了一个利用自动模式提取方法从金融知识图谱中挖掘出的规则库（见图2）。规则挖掘旨在提取反复出现的关系模式，作为知识图谱推理框架的先验，从而提升效率和可解释性。

规则定义。 规则定义为一系列连接实体的关系原子（规则体），并带有一个明确的头部，用于映射到股票走势预测。规则定义请参阅 3.1 节。

挖掘流程。 遵循自动化挖掘框架，我们采用广度优先枚举，在知识图谱中枚举路径以构建候选规则：

(1) **路径采样**。 从知识图谱中的所有三元组 (e_h, r, e_t) 出发，我们构建二跳路径，并递归地将其组合成三跳和四跳路径。(2) **规则实例化**。 将每种路径模式泛化为规则体，方法是用变量替换实体。对于每种体模式，我们都创建上涨和下跌两种头部变体。(3) **支持度与置信度筛选**。 对于头部为 predict_direction(X, ℓ) 的候选规则 ρ，令 x 表示满足体模式的股票-日期实例数量，令 y 表示其中实际股票走势与预测方向 ℓ一致的实例数量。置信度定义如下：

$$\mathrm{conf}(\rho) = \frac{y}{x}.$$

支持度低于频率阈值或 $\mathrm{conf}(\rho) < \tau_{\mathrm{mine}}$ 的规则将被舍弃，其中 $\tau_{\mathrm{mine}} = 0.60$ 是我们的挖掘阈值。

(4) **剪枝**。为避免稀疏性，会移除冗余规则（被更短的规则所涵盖）或过于具体的规则（支持度过低）。

示例。 挖掘得到的规则示例如下：

> 投资_于$(X,\ Z)$ ∧属于_行业$(Z,$ "人工智能"$)\Longrightarrow$ 是 X走势为正的信号，

自动化规则挖掘将原始知识图谱结构转化为紧凑的关系模式库，为我们的知识图谱推理框架提供指引。符号规则与神经推理的协同作用缩小了搜索空间，减少了幻觉，并使预测结果立足于具有金融意义的因果链。

3.4 基于图的探索

我们的框架以基于图的探索为核心，通过规则和大语言模型引导，对金融知识图谱进行迭代遍历，从而发掘用于股票走势预测的可解释推理路径。图3展示了可解释推理路径的示例，算法框架则见算法1。

搜索初始化。 对于每个交易日 $t \in [t_{\mathrm{start}}, t_{\mathrm{end}}]$ 和每只股票 $s \in \mathrm{S}$，我们从对应的TextSource节点开始探索。截至时点约束确保只能访问时间戳为 $\leq t$ 的边和TextSource节点。因此，每条推理路径$\mathrm{P}_{s,\ t}$都以预测时点可获得的信息为依据。

束搜索。我们采用束搜索来探索知识图谱，同时控制组合爆炸。在深度 d，每条部分路径$P = (e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} \cdots \xrightarrow{r_d} e_d)$可通过通向 e_d的邻居节点的候选关系$\{r_{d+1}\}$进行扩展。束搜索保留由评分函数排序的前K条路径（见下文）。搜索持续进行，直至（i）到达TextSource节点，或（ii）达到最大深度 $D_{\max}$。

规则约束的路径扩展。我们的方法论的一项关键创新是在...
...期间融入规则约束的推理

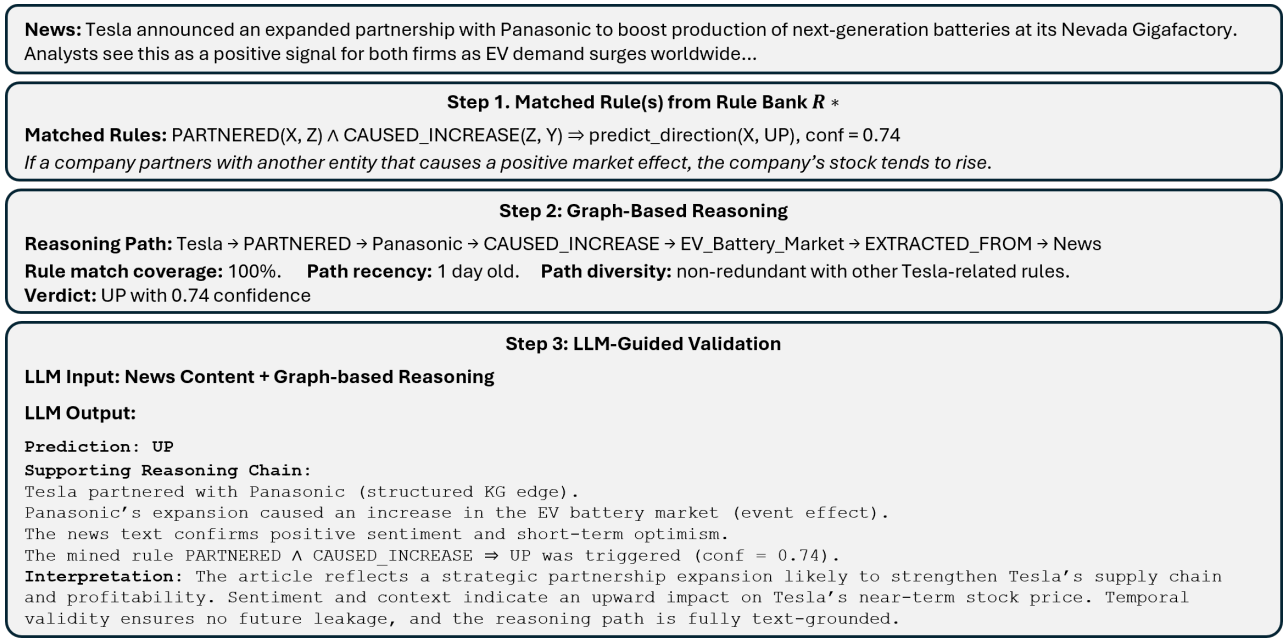

图3：可解释推理路径示例，展示了我们的知识图谱推理框架如何将股票与文本证据关联起来。

图探索。对于扩展过程中的每条部分路径 P，我们都会尝试将其与 R^* 中已知标签预测规则的规则体进行匹配：

(1) **规则匹配：**给定当前部分路径 $P = (e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} \cdots \xrightarrow{r_d} e_d)$，我们找出所有规则 $\rho \in \mathrm{R}^*$，其规则体序列与关系模式 $\{r_1, r_2, \ldots, r_d\}$ 相匹配。(2) **假设生成：**当一条规则的规则体完全实例化，且规则置信度超过我们的假设阈值（$\mathrm{conf}(\rho) > \tau_{\mathrm{hyp}}$ where $\tau_{\mathrm{hyp}} = 0.60$）时，我们将其视为一个标签预测假设。该假设包括：

- 置信度为 $\mathrm{conf}(\rho) > \tau$hyp、预测方向为 ℓ的实例化规则 ρ
- 支持该规则体的完整证据路径 P
- 提供文本证据的关联TextSource节点

这种以规则为依据的方法确保探索并非仅仅遍历图结构，而是主动寻找证据，以验证规则挖掘过程中发现的高置信度预测模式。

关系选择。给定当前路径前缀$\{r_1, \ldots, r_d\}$以及来自 e_d的候选关系集合，我们应用两阶段过滤：

(1) **基于规则的可采纳性：**关系序列必须是至少一个规则体 $\rho \in \mathrm{R}^*$ 的前缀；(2) **基于大语言模型的剪枝：**若启用，大语言模型关系选择器会结合上下文（节点类型、截至日期、邻居统计）评估候选关系，并返回一个带有分数 {0，1，2} 的子集，用以表示相关性。

若大语言模型选择器被禁用，则使用基于邻居频率和时效性的启发式评分器。

路径评分。我们采用逐深度词典序排序。路径 P可被接受，当且仅当(i)其关系序列是R^*中某条规则的前缀，(ii)所有边截至 t均有效，且 (iii) $\mathrm{len}(P) \le D_{\max}$。

在深度 d，为每个可接受的 P计算四个信号，每个信号均按其在深度 d候选中的百分位进行缩放：$\mathrm{hyp}_\tau(P) \in \{0, 1\}$（补全一条置信度为 $\ge \tau$的挖掘规则），$\mathrm{cov}(P) \in [0, 1]$（最佳规则体覆盖率），$\mathrm{rec}(P) \in [0, 1]$（$P$上最新的时间戳），$\mathrm{ahub}(P) \in [0, 1]$（惩罚高度数节点；越高则枢纽偏差越小）。

按字典序键排序

$$K(P) = (-\text{假设}_\tau(P), -\text{覆盖率}(P), -\text{时效性}(P), -\text{枢纽节点}(P), \text{长度}(P), \text{哈希}(P)),$$

按升序排列，其中负号表示“越高越好”，而 $\mathrm{hash}(P)$ 是一种确定性的节点序列哈希，用于稳定地进行平局裁决。此方案优先完成高置信度规则，然后优先选择与规则体最匹配、时间上较近且避开枢纽节点的路径。

提前停止与假设验证。当路径到达一个文本来源节点并形成完整的标签预测假设时，搜索会提前终止。满足以下条件时即会发生：

- 规则体已在当前图快照中完全实例化
- 该规则具有较高的置信度（$\mathrm{conf}(\rho) > 0.6$）
- TextSource 证据支持完整的推理链

形成此类假设后，它会被添加到股票预测的候选解释集合中。如果某个扩展缺乏

支持事实，或无法完成任何高置信度规则，试错机制就会丢弃该分支，并回溯到其他候选项。

3.5 裁决预测

经过图探索后，对于每只股票 s 在日期 t，我们会得到一组候选推理路径 $\mathrm{P}_{s,\,t,\,1}$，用于主要的 1 天预测时域（也支持其他时域 $\mathrm{P}_{s,\,t,\,h}$，其中 $h \in \{3,\ 5,\ 10,\ 30\}$）。每条路径都以一个或多个 TextSource节点 结束，并与挖掘出的规则库 R^{*}中的一部分规则相对应。本阶段的目标是将这些候选路径整合为二元裁决 $\hat{y}_{s,\,t,\,1} \in \{$上涨，下跌$\}$，并为我们的主要关注方向给出置信度分数。

主要裁决生成。最终预测由图探索过程中生成的高置信度基于规则的假设 $\mathrm{H}_{s,\,t,\,h} = \{(P_i,\ \rho_i,\ \mathrm{conf}(\rho_i))\}$的置信度分数组合得出。每个假设都包含一条实例化规则 ρ_i，其置信度为 $\mathrm{conf}(\rho_i) > \tau_{\mathrm{hyp}}$，预测方向为 ℓ_i。组合方式如下：

$$\mathrm{conf}_{\mathrm{UP}}(s,t,h) = \max_{\substack{(P_i,\rho_i)\in\mathcal{H}_{s,t,h}\\ \ell_i=\mathrm{UP}}} \mathrm{conf}(\rho_i),$$

$$\mathrm{conf}_{\mathrm{DOWN}}(s,t,h) = \max_{\substack{(P_i,\rho_i)\in\mathcal{H}_{s,t,h}\\ \ell_i=\mathrm{DOWN}}} \mathrm{conf}(\rho_i).$$

最终预测选择汇总置信度较高的标签：

$$\hat{y}_{s,t,h} = \begin{cases} \mathrm{UP}, & \text{if } \mathrm{conf}_{\mathrm{UP}} \geq \mathrm{conf}_{\mathrm{DOWN}}, \\ \mathrm{DOWN}, & \text{otherwise.} \end{cases}$$

路径评估。 大语言模型（LLM）在严格的提示约束下评估每条候选路径 $P \in \mathrm{P}_{s,\,t,\,1}$，并给出二元标签和置信度（$y_P$，$p_P$）。此评估用于解释排序和可解释性评分。

判定的可解释性。 每个最终决策都会附上对结果有所贡献的排名前M的支持性推理路径。例如，“上涨”的判定可能得到以下内容的支持：

公司 A–投资_于–>初创公司 Z–提取_自–>新闻 A

这与规则模式 ρ 相符：投资_于$(X,\ Z)$ $\wedge$ 提取_自$(Z,\ T)$。判定、底层规则与文本锚点之间的这种明确关联，使金融预测透明且可由人工审计。

4 实验

我们通过全面实验，在真实世界的金融预测任务中评估所提出的时序知识图谱推理方法。评估结果表明，与最先进基线相比，我们的方法不仅预测性能更优，还能生成高质量的可解释说明。

4.1 实验设置与数据集

金融知识图谱构建。 我们为标准普尔500指数成分公司构建了一个时序金融知识图谱（2022年1月至2023年1月），其中包含42,188个实体和174,415条带时间戳的边。参见第3.2节。我们使用全部487家标普500成分股的每日OHLCV数据（187,452条观测），并进行了拆股/股息调整和异常值过滤。根据远期收益定义二元标签：$y_{s,\,t,\,1} = \mathbb{I}[P_{s,t+1} > P_{s,t}]$，由此得到均衡的上涨/下跌分布（51.2%/48.8%）。我们严格执行截至日期约束，并采用逐日预测的滚动前向验证。

4.2 评估指标

我们采用标准分类指标进行评估：准确率、精确率、召回率和 F1分数。对于可解释性，我们衡量路径覆盖率、规则利用率、证据锚定和路径多样性。

4.3 基线方法

我们按照相同的截至日期协议评估具有代表性的基线方法：

- **历史动量**：根据近期价格趋势，使用标准动量信号预测次日方向。回看窗口在验证集上选定，并应用于样本外数据。
- **XGBoost** [38]：使用每日 OHLCV 特征和轻量级技术指标训练的梯度提升树。类别加权和所有超参数均在验证集上调优，且不产生前瞻偏差。
- **新闻情绪（Qwen QwQ-32B）** [30, 45]：仅使用新闻的模型，利用通义千问根据同期文章对每只股票每日的情绪进行评分；评分按日汇总，并通过阈值划分为上涨/下跌。
- **T-GNN** [42]：基于我们的金融知识图谱构建的时序图神经网络，按照相同的数据划分和时间约束跨时间执行消息传递。
- **智能体框架**：*Mars* [6], *StockAgent* [51],和 *Fin-Mem* [50] 均由我们自行实现，以适配分类评测框架。其输出是标准化的、基于规则的智能体风格流程代理方案，纳入其中以求完整；所报告的数值不应被解读为这些原始系统的官方结果。
- **我们的方法**：提出的时序知识图谱推理模型，结合规则引导的多跳探索和路径聚合。

所有方法均采用相同的数据划分、受信息日期限制的特征以及相同的评估协议。在回测评估中，我们为每种方法构建前10名买入并持有组合（等权重，不再平衡），并报告全部股票等权重基准。除非另有说明，训练窗口为2022年1月1日至2023年1月1日，评估窗口为2023年1月1日至2024年1月1日。所有大语言模型实验均使用 Qwen QwQ-32B 模型 [45]。

4.4 主要实验结果

表1展示了各基线方法在1日预测跨度的股票走势预测任务上的性能比较。

预测性能分析。 表1显示，我们的方法取得了最佳的单日性能，准确率为55.1%，精确率为55.7%，召回率为71.5%，F1为60.8%。这相当于较准确率为52.9%的XGBoost提升了2.2个准确率百分点，较准确率为54.4%的FinMem提升了0.7个百分点。这一改进与方法论相一致。首先，规则引导的探索将每次扩展限制为前缀，即

rticle

表 1：1日股票走势预测的性能比较。最佳结果以粗体标出。

方法	准确率	精确率	召回率	F1分数
历史动量	0.517	0.524	0.709	0.603
XGBoost	0.529	0.535	0.678	0.598
新闻情绪（通义千问）	0.524	0.529	0.636	0.594
T-GNN	0.528	0.534	0.685	0.600
Mars	0.539	0.544	0.664	0.598
StockAgent	0.539	0.545	0.648	0.592
FinMem	0.544	0.550	0.647	0.594
TRACE（我们的方法）	**0.551**	**0.557**	**0.715**	**0.608**

表2：详细的消融研究，展示各个组件对整体性能的贡献。

配置	性能指标		
	准确率	精确率	召回率
完整模型	**0.551**	**0.557**	**0.715**
- 时间约束	0.539	0.544	0.534
- 规则挖掘	0.542	0.547	0.537
- 多跳推理	0.532	0.537	0.527
- 路径聚合	0.548	0.553	0.543
- 大语言模型关系选择	0.545	0.550	0.540
随机分类器	0.502	0.501	0.500

挖掘规则体，并且仅当一条高置信度规则在支持性 TextSource 证据的充分支撑下完全落地时，才促进假设的形成，从而将搜索集中于具有经济意义的链条，而非任意的多跳遍历。其次，按深度的字典序排序优先完成高置信度规则，其次是规则体覆盖，同时惩罚枢纽节点，并采用稳定的平局打破方式，从而提升那些实例化强先验、而非高度数捷径的路径。第三，判定阶段将高置信度且已落地的假设聚合为单一标签与置信度，从而实现对异质证据的选择性融合，而非均匀汇聚。这些组成部分共同带来更高的灵敏度，且不损失选择性，体现为在相同评估协议下，相对T-GNN的召回率提升3.0个百分点、F1提升0.8个百分点，以及整体准确率的提升。

5 分析

本节展示了我们框架在可解释性质量、人类专家分析、消融研究、回测分析和反事实分析方面的实验与结果。

5.1 可解释性质量分析

我们方法的一个关键优势是生成支持每项预测的可解释推理路径。

路径质量分析。 我们的框架平均为每项股票预测生成 2.8 条推理路径，其中 85% 的预测至少有一条从目标股票到文本来源的完整路径支持。在发现的路径中，72% 至少匹配挖掘出的规则库中的一种规则模式，表明该探索过程成功利用了已学习的关系模式。

这些路径展现出良好的多样性，每条路径平均包含2.1种不同的关系类型，每次预测平均包含1.8种独特的规则模式。这种多样性表明，该框架捕捉了多种互补信号，而非依赖单一推理模式。

证据锚定。 在考察证据的时间相关性时，我们发现65%的推理路径终止于预测日期前7天内发布的文本来源，86%终止于30天内发布的文本来源，从而确保支撑证据具有时效性，并对1日预测尤为相关。这些路径连接到多样化的信息来源，包括金融新闻文章、公司文件和市场报告，为预测结果提供了广泛基础。

5.2 消融研究

我们通过受控消融实验系统评估关键组件的贡献（见表2）。

组件影响分析。 多跳推理带来最大的性能贡献（提升2.1%），证实复杂的金融关系需要多步图遍历。规则挖掘贡献了1.1%的提升，表明学习到的关系模式能够在基本图结构之外进一步提高预测质量。

时间约束带来1.4%的提升，凸显了在金融预测中遵循信息可用日期的重要性。路径聚合带来0.5%的提升，表明结合多条推理路径比依赖单一路径能产生更稳健的预测结果。

每个组件都对整体性能有正向贡献，验证了我们知识图谱推理框架中的设计选择。

5.3 人类专家分析

路径质量指标。 我们从多个维度评估推理路径质量：

- **语义连贯性**：83%的路径展现出合乎逻辑的金融关系
- **时间一致性**：91%遵循事件的时间顺序
- **规则对齐**：74% 至少匹配一条挖掘规则模式
- **来源多样性**：每次预测平均涉及 2.3 个不同的新闻来源

人工评估方案。 我们邀请金融分析师对随机选取的 100 条预测结果进行了专家评估。评估者间信度（Fleiss' $\kappa = 0.68$）表明，他们对推理路径质量的评价具有良好的一致性。专家认为，73% 的推理路径“具有金融意义”，64% 的推理路径“可用于投资决策”。

表3：前10买入并持有投资组合指标（2023-01-01 至 2024-01-01）。每种方法均构建由其排名前十的股票代码组成的仅做多、等权重投资组合，且不进行再平衡。最佳值以粗体标出。

方法	总回报	夏普比率	年化波动率	最大回撤	胜率
全部等权	16.7%	1.17	14.1%	-12.9%	52.8%
历史动量	8.7%	0.62	15.4%	-17.3%	54.4%
XGBoost	27.0%	1.31	19.9%	-14.8%	53.2%
新闻情绪（通义千问）	9.2%	0.67	14.8%	-12.2%	52.8%
T-GNN	3.5%	0.33	**13.3%**	-12.5%	51.2%
Mars	14.2%	0.98	14.7%	-17.3%	54.8%
StockAgent	26.4%	1.43	17.7%	-15.0%	52.4%
FinMem	13.1%	0.88	15.5%	-15.7%	55.2%
Ours	**41.7%**	**2.00**	18.5%	**-11.7%**	**56.0%**

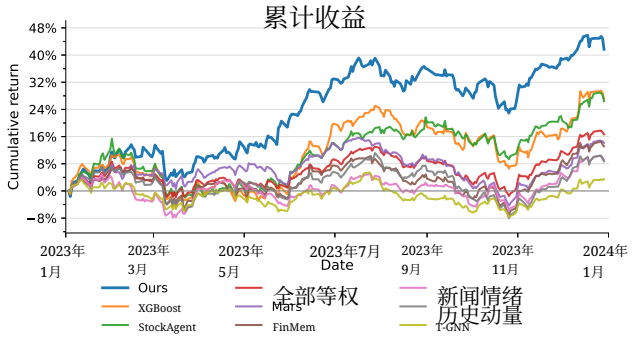

图4：所有方法的前十名买入并持有投资组合的累计收益（2023-01-01 至 2024-01-01）。我们的方法在整个时间范围内均占据优势，并超过等权基准。

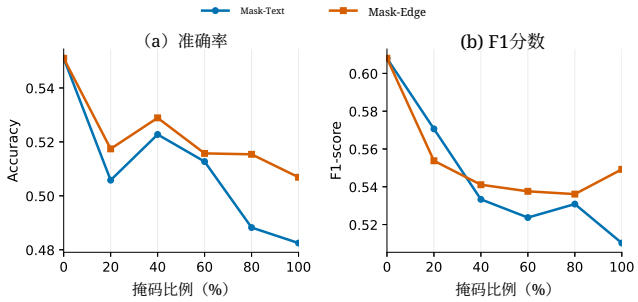

图 5：Mask-Tex 和 Mask-Edge 策略在不同删除比例下的反事实分析。

5.4 对比方法论增强

集成方法分析。 将我们的知识图谱推理框架与顶尖的基于市场数据的基线方法（XGBoost）相结合，准确率达到 56.1%，表明两种信号来源具有互补性。然而，微幅提升（0.8%）伴随着显著的复杂度成本，这支持了我们采用独立框架设计的做法。

5.5 回测分析

设置。 我们评估各方法在 2023-01-01 至 2024-01-01 期间构建的前10只股票买入并持有组合（等权重，不再平衡），并与等权重全股票基准进行比较。价格已进行拆股和股息调整，并严格执行所有截至时点约束。

结果。 见表3和图4，我们的方法实现了 41.7% 的总回报，夏普比率为 2.00，年化波动率为 18.5%，最大回撤为 −11.7%，胜率为 56%。等权基准实现了 16.7% 的总回报（夏普比率为 1.17）。表现较强的基线方法结果较低：XGBoost 为 27.0%（1.31），StockAgent 为 26.4%（1.43），Mars 为 14.2%（0.98），FinMem 为 13.1%（0.88），历史动量为 8.7%（0.62），新闻情绪为 9.2%（0.67），T-GNN 为 3.5%（0.33）。这些结果表明，在相同的截至日期评估条件下，分类准确率的提升能够转化为具有显著经济意义的改进。

5.6 反事实分析

设置。 本分析旨在验证我们的预测判定是否依赖所呈现的证据路径，而非虚假相关性。对于每个经过评估的股票日，我们的框架会返回一条排名靠前的支持路径 $\mathrm{P}^{\star} = e_1 \xrightarrow{r_1} \cdots \xrightarrow{r_{k-1}} e_k$，以及判定 $y \in \{\text{UP}、\text{DOWN}\}$ 的置信度分数 $s \in [0,1]$。我们执行两种反事实删除：（i）**Mask-Text**：移除末端的 TextSource节点及其关联边（EXTRACTED_FROM），从而剥离与该路径相关的新闻证据；（ii）**Mask-Edge**：移除路径上的第一条关键关系（例如 ACQUIRED、PARTNERED、CAUSED_INCREASE），该关系使匹配规则得以完整。随后，我们在修改后的图快照上重新计算判定和置信度。我们确保每个实例最多只进行一次扰动，保持搜索超参数不变，并在受扰动的快照 G_t 上重新计算预测结果。我们报告准确率和 F1；r=0 处的值与未扰动评估相同。

研究结果。 图5显示，随着两种策略的掩码比例提高，两项指标均呈单调下降趋势，这验证了证据路径在判定中的有效性。Mask-Text的性能下降速度快于Mask-Edge，表明终端节点的新闻依据对决策有所贡献。请注意，Mask-Text的掩码比例为 r=100%时，意味着预测仅使用新闻内容，而不使用任何路径。

6 结论

我们提出了一种名为TRACE的时序知识图谱推理框架，将规则引导的多跳探索与基于文本的证据相结合，以预测次日股票走势。在标准普尔500指数任务中，我们的方法准确率达到55.1%、精确率达到55.7%、召回率达到71.5%、F1值达到60.8%，优于表现强劲的机器学习、文本、图和智能体式基线方法。在2023-01-01 至2024-01-01期间进行的前十名买入并持有回测中，该方法实现了41.7%的总回报和2.00的夏普比率，优于等权基准和具有竞争力的基线方法。该框架能够生成路径级解释，但仍受知识图谱覆盖率和漂移的影响；未来工作将拓展至多时域预测、校准和跨市场验证。

A 可复现性细节

A.1 数据、数据划分与截至时点约束

我们研究标准普尔500指数成分股在截至时点设定下的次日股票走势（收盘价到收盘价）。除非另有说明，默认数据划分为：训练集 = 2022-01-01 – 2023-01-01；测试集 = 2023-01-01 – 2024- 01-01。对于任意日期 t，快照 G_t及所有特征仅包含时间戳为 $\le t$ 的事实和文档。金融文本（新闻、申报文件、文字记录）通过股票代码-日期哈希和文本嵌入距离计算，在24小时窗口内去重。特征归一化仅在训练集上拟合，并应用于测试集中的基于市场数据的基线方法。

A.2 模型与搜索设置

我们在截至时点快照 G_t 上进行规则引导搜索，束宽度为 K=8，最大深度为 $D_{\max}$=3，假设阈值为 τ_{hyp}=0.60，且每个实例最多保留10条评分路径（默认采用前M条，证据预算为 M=5）。对于 TextSource 节点和查询片段的稠密表示，文本相关性通过 Qwen3-Embedding- 8B 计算。

我们在两处使用 Qwen QwQ-32B 作为冻结控制器：

（1）LLMRelationSelector：给定来自当前前沿的一组候选单跳扩展，大语言模型会筛选出（i）与新闻片段语义兼容且（ii）与部分匹配的规则前缀一致的扩展。大语言模型能有效减少枢纽节点与语义漂移。（2）LLMValidate：对于每个已完成的（路径，规则）对 (P, ρ)，大语言模型会验证其与新闻内容的标签一致性，并得到（$\hat{\ell}_P, p_P$），其中 $p_P \in [0,1]$ 是用于证据融合的经校准的合理性分数。

A.3 评估与显著性

本节定义了论文中使用的分类 与回测 指标。对于每个（股票代码，日期）实例 $i \in \{1,\dots,N\}$，令 $y_i \in${上涨,下跌} 为次日真实标签（收盘价对收盘价），$\hat{y}_i$ 为模型预测。

分类。 我们报告汇集的（股票代码，交易日）集合上的标准点指标：

$$\text{Accuracy} = \frac{\text{TP}+\text{TN}}{N}, \qquad \text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}},$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}}, \qquad \text{F1} = 2\cdot\frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}}.$$

回测。 令 B 为在测试窗口的开始 时选定的前10、等权重、仅做多组合。对于每个交易日 $t = 1 \dots T$，令 $r_{j,t}$为资产 $j \in$B 从收盘$_{t-1}$到收盘$_t$的简单收益率：

$$r_{j,t} = \frac{P_{j,t} - P_{j,t-1}}{P_{j,t-1}}.$$

每日投资组合收益率（等权重 $w_j = 1/|\mathcal{B}|$）：

$$r_t = \sum_{j\in\mathcal{B}} w_j\, r_{j,t} \quad \text{with } w_j = \tfrac{1}{10}.$$

将每日无风险利率记为 $r^{(f)}{}_t$（除非另有说明，否则设为 0）。

总回报（累计）。

$$\text{TotRet} = \prod_{t=1}^{T}(1+r_t) - 1.$$

年化波动率。 令 $\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$和 $s = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(r_t-\bar{r})^2}$ 为日收益率的样本标准差。则

$$\text{AnnVol} = s\cdot\sqrt{252}.$$

夏普比率（年化）。 使用每日超额 收益率 $\tilde{r}_t = r_t - r^{f}{}_t$，其样本均值和标准差为 $s_{\tilde{r}}$：

$$\text{Sharpe} = \frac{\bar{\tilde{r}}\cdot 252}{s_{\tilde{r}}\cdot\sqrt{252}} = \sqrt{252}\cdot\frac{\bar{\tilde{r}}}{s_{\tilde{r}}}.$$

最大回撤。 令累计权益曲线为 $C_t = \prod_{u=1}^{t}(1+r_u)$，运行峰值为 $H_t = \max_{u\le t} C_u$。在 t 处的回撤为 $DD_t = \frac{C_t}{H_t} - 1$。所报告的统计量为

$$\text{MaxDD} = \min_{t\in\{1,\dots,T\}} DD_t \quad \text{(a non-positive number)}.$$

胜率。

$$\text{WinRate} = \frac{1}{T}\sum_{t=1}^{T}\mathbb{I}[r_t > 0].$$

在测试窗口开始时，每种方法都会构建一个由其排名前十的股票代码组成的等权重仅做多组合，并持有该组合（不再平衡）。除非另有说明，否则不计股息、费用和滑点。

A.4 反事实与掩码比例研究

反事实删除。 (Mask-Text) 删除末端文本来源（表示新闻与事件内容）及其在顶层路径上的 EXTRACTED_FROM 边；(Mask-Edge) 删除第一个补全规则的关系（例如，收购、合作、CAUSED_INCREASE）。然后，我们针对受扰动的G_t重新计算判定及其置信度，并报告 (i) 准确率和 (ii) F1分数。

掩码比例扫描。 对于 $r \in \{0, 20, 40, 60, 80, 100\}$%，随机扰动 r%的实例（每个实例进行一次扰动），并重新评估。

B 伦理与使用注意事项

我们使用公开的带时间戳信息，并严格执行截至时点过滤，以避免前视偏差。模型仅作为辅助工具，不构成财务建议。我们通过规则与新闻路径提供可解释性。回测仅供示例，且对成本和滑点较为敏感。

C 算法流程

我们通过算法1展示支撑所有结果的端到端推理流程，更多技术细节如下。

算法1 规则引导的时序知识图谱推理 输入：带有时间 τ 的新闻 x（x）；知识图谱G_t =（E_t、 R_t、 F_t），包含 $t \geq \tau$（x）；规则 R^*；束搜索 K；最大深度 $D_{\max}$；阈值τ_{hyp}；证据预算 M；融合权重 $\alpha \in [0,1]$输出：$y \in \{\text{UP}, \text{DOWN}\}$，置信度 conf 和证据 P 1: S ← 抽取实体与关系(x)； G_t ←知识图谱落地(S, G_t) 2: Score(P) =⟨𝕀[CompletesHighConfRule($P, R^*, \tau_{\text{hyp}}$)],3: BestRuleCoverage($P, R^*$), 时效性($P$), AntiHubPenalty($P$) $-|P|$⟩4: 对 $d=1$ 到 $D_{\max}$ 执行 5: Cands ←∅6: 对所有 $P \in$ 束搜索中的元素执行 7: 对所有在 t时有效的邻居节点 (Head(P) $\xrightarrow{r}$ v) 执行 8: 如果 IsPrefixOfSomeRule($P \cdot (r, v)$ R^*)，则 9: $P' \leftarrow P \cdot (r, v)$；将 P'加入 Cands 10: 如果 CompletesHighConfRule($P', R^*, \tau_{\text{hyp}}$)，则 11: ($\rho, \ell$, conf ($\rho$)) ← MatchedRule($P', R^*$)；将 ($P', \rho$, conf ($\rho$) ℓ) 加入 H 12: 结束 if 13: 结束 if 14: 结束 for 15: 结束 for 16: 如果 Cands$=\emptyset$，则跳出循环 17: 结束 if 18: Cands ← LLMRelationSelector(Cands)；Beam ← TopK(Cands K, lexi Score) 19: 结束 for 20: 如果 H$=\emptyset$，则 21: 返回 (下跌 0, ∅) 22: 结束 if 23: confUP ← $\max_{(P,\rho,\cdot,\ell)\in\mathcal{H},\, \ell=\text{UP}}$ conf (ρ)；confDOWN ←$\max_{(P,\rho,\cdot,\ell)\in\mathcal{H},\, \ell=\text{DOWN}}$ conf (ρ) 24: 如果 confUP ≥ confDOWN，则 25: y←上涨；conf ←confUP 26: 否则 27: y←下跌；conf ←confDOWN 28: 结束 if 29: E ←∅； 30: 对所有 (P, ρ, conf (ρ) ℓ) ∈H 执行 31: ($\hat{}\ell_P, p_P$) ← LLMValidate(x, P, ρ)；如果 $\hat{}\ell_P = \ell$，则将 (P, ρ, conf (ρ) ℓ, p_P) 加入 E 32: 结束 for 33: E$_y$ ← {$e \in$ E : $\ell = y$}；P ←按(对 E$_y$, M, αconf (ρ) $+(1-\alpha)$ p_P)选取TopM 34: 返回 (y, conf, P)

参考文献

[1] Dogu Araci. 2019. Finbert：使用预训练语言模型进行金融情感分析。*arXiv*预印本 *arXiv:1908.10063* (2019)。[2] Adebiyi A Ariyo、Adewumi O Adewumi 和 Charles K Ayo。2014。使用 ARIMA模型进行股票价格预测。载于*2014*年*UKSim-AMSS*第*16*届计算机建模与仿真国际会议。IEEE，106–112。[3] Farah Atif、Ola El Khatib 和 Djellel Difallah。2023。BeamQA：结合序列到序列预测与束搜索的多跳知识图谱问答。载于第*46*届国际*ACM*信息检索研究与开发*SIGIR*会议论文集（中国台湾台北）（*SIGIR '23*）。计算机协会，纽约，纽约州，美国，781–790。doi:10.1145/3539618.3591698[4] Jinheon Baek、Alham Fikri Aji 和 Amir Saffari。2023。用于零样本知识图谱问答的知识增强语言模型提示。*arXiv*预印本 *arXiv:2306.04136* (2023)。[5] Shaojie Bai、J Zico Kolter 和 Vladlen Koltun。2018。用于序列建模的通用卷积网络和循环网络的实证评估。*arXiv*预印本 *arXiv:1803.01271* (2018)。[6] Jiayi Chen、Jing Li 和 Guiling Wang。2025。MARS：面向风险感知多智能体投资组合管理的元自适应强化学习框架。*arXiv*预印本 *arXiv:2508.01173* (2025)。[7] Shib Sankar Dasgupta、Swayambhu Nath Ray 和 Partha Talukdar。2018。Hyte：基于超平面的时间感知知识图谱嵌入。载于*2018*年自然语言处理实证方法会议论文集。2001–2011。[8] Qianggang Ding、Haochen Shi、Jiadong Guo 和 Bang Liu。2024。Tradexpert：以专家混合大语言模型革新交易。*arXiv*预印本 *arXiv:2411.00782* (2024)。[9] Qianggang Ding、Sifan Wu、Hao Sun、Jiadong Guo 和 Jian Guo。2020。用于股票走势预测的分层多尺度高斯 Transformer。载于*IJCAI*。4640–4646。[10] Xiao Ding、Yue Zhang、Ting Liu 和 Junwen Duan。2015。用于事件驱动的股票预测的深度学习。载于*IJCAI*，第15卷。2327–2333。[11] Fuli Feng、Huimin Chen、Xiangnan He、Ji Ding、Maosong Sun 和 Tat-Seng Chua。2018。通过对抗训练增强股票走势预测。*arXiv*预印本 *arXiv:1810.09936* (2018)。[12] Fuli Feng、Xiangnan He、Xiang Wang、Cheng Luo、Yiqun Liu 和 Tat-Seng Chua。2019。用于股票预测的时序关系排序。*ACM*信息系统汇刊（*TOIS*）37，2（2019），1–30。[13] Thomas Fischer 和 Christopher Krauss。2018。使用长短期记忆网络进行深度学习，以预测金融市场。《欧洲运筹学杂志》270，2（2018），654–669。[14] Luis Galárraga、Christina Teflioudi、Katja Hose 和 Fabian M Suchanek。2015。在本体知识库中进行快速规则挖掘，使用 AMIE

+

. *VLDB*期刊24, 6 (2015), 707–730. [15] Xiao Huang, Jingyuan Zhang, Dingcheng Li, and Ping Li. 2019. 基于知识图谱嵌入的问答。在第十二届 *ACM* 网络搜索与数据挖掘国际会议论文集。105–113. [16] Jinhao Jiang, Kun Zhou, Zican Dong, Keming Ye, Wayne Xin Zhao, and Ji-Rong Wen. 2023. Structgpt：一种利用大语言模型对结构化数据进行推理的通用框架。*arXiv* 预印本 *arXiv:2305.09645* (2023年)。[17] Woojeong Jin, Meng Qu, Xisen Jin, and Xiang Ren. 2019. 循环事件网络：时序知识图谱上的自回归结构推断。*arXiv* 预印本 *arXiv:1904.05530* (2019)。[18] Daniel Jurafsky and James H Martin. 2009. 语音与语言处理：自然语言处理、计算语言学和语音识别导论。第 *2* 版。培生教育国际。[19] Raehyun Kim, Chan Ho So, Minbyul Jeong, Sanghoon Lee, Jinkyu Kim, and Jaewoo Kang. 2019. HATS：一种用于股票走势预测的层次图注意力网络。*arXiv* 预印本 *arXiv:1908.07999* (2019年)。[20] Xiaohui Victor Li and Francesco Sanna Passino. 2024. FinDKG：利用大语言模型构建动态知识图谱，以检测金融市场的全球趋势。在第五届 *ACM* 人工智能金融国际会议论文集。573–581。[21] Ruotong Liao, Xu Jia, Yangzhe Li, Yunpu Ma, and Volker Tresp. 2023. GenTKG：利用大语言模型对时序知识图谱进行生成式预测。*arXiv* 预印本 *arXiv:2310.07793* (2023年)。[22] Xi Victoria Lin, Richard Socher, and Caiming Xiong. 2018. 利用奖励塑造进行多跳知识图谱推理。*arXiv* 预印本 *arXiv:1808.10568* (2018年)。[23] Wei Liu, Weihao Zeng, Keqing He, Yong Jiang, and Junxian He. [无日期]。什么样的数据适合对齐？指令微调中自动数据选择的综合研究。在第十二届国际学习表征会议。[24] Zhenghao Liu, Zhijian Zhang, and Xi Zeng. 2024. 通过知识关联进行风险识别与管理：金融事件演化知识图方法。《专家系统及其应用》252（2024），123999。[25] Wen Long、Zhichen Lu 和 Lingxiao Cui。2019。基于深度学习的股价走势预测特征工程。《基于知识的系统》164（2019），163–173。[26] Linhao Luo、Zicheng Zhao、Gholamreza Haffari、Yuan-Fang Li、Chen Gong 和 Shirui Pan。2024。图约束推理：利用大语言模型在知识图谱上进行忠实推理。arXiv 预印本 arXiv:2410.13080（2024）。[27] Shengjie Ma、Chengjin Xu、Xuhui Jiang、Muzhi Li、Huaren Qu、Cehao Yang、Jiaxin Mao 和 Jian Guo。2024。Think-on-graph 2.0：通过知识引导的检索增强生成，实现深入且忠实的大语言模型推理。arXiv 预印本 arXiv:2407.10805（2024）。[28] Christian Meilicke、Melisachew Wudage Chekol、Manuel Fink 和 Heiner Stuck-enschmidt。2020。用于知识图谱补全的强化式随时自底向上规则学习。arXiv 预印本 arXiv:2004.04412（2020）。[29] David MQ Nelson、Adriano CM Pereira 和 Renato A De Oliveira。2017。利用 LSTM 神经网络预测股票市场价格走势。载于《2017年国际联合神经网络会议》（IJCNN）。Ieee，1419–1426。[30] Thien Hai Nguyen、Kiyoaki Shirai 和 Julien Velcin。2015。基于社交媒体情感分析的股票走势预测。《专家系统及其应用》42，24（2015），9603–9611。[31] Hao Qian、Hongting Zhou、Qian Zhao、Hao Chen、Hongxiang Yao、Jingwei Wang、Ziqi Liu、Fei Yu、Zhiqiang Zhang 和 Jun Zhou。2024。MDGNN：用于全面、动态股票投资预测的多关系动态图神经网络。载于《AAAI人工智能会议论文集》第38卷，14642–14650。[32] Meng Qu、Junkun Chen、Louis-Pascal Xhonneux、Yoshua Bengio 和 Jian Tang。2020。RNNLogic：学习用于知识图谱推理的逻辑规则。arXiv 预印本 arXiv:2010.04029（2020）。[33] Andrea Rossi、Denilson Barbosa、Donatella Firmani、Antonio Matinata 和 Paolo Merialdo。2021。用于链接预测的知识图谱嵌入：比较分析。《ACM数据知识发现汇刊》（TKDD）15，2（2021），1–49。[34] Mohammed Saeed、Naser Ahmadi、Preslav Nakov 和 Paolo Papotti。2021。RuleBERT：向预训练语言模型传授软规则。arXiv 预印本 arXiv:2109.13006（2021）。[35] Hao Sun、Xiao Liu、Yeyun Gong、Anlei Dong、Jingwen Lu、Yan Zhang、Daxin Jiang、Linjun Yang、Rangan Majumder 和 Nan Duan。2023。BeamSearchQA：大语言模型是强大的零样本问答解答器。arXiv 预印本 arXiv:2305.14766（2023）。[36] Jiashuo Sun、Chengjin Xu、Lumingyuan Tang、Saizhuo Wang、Chen Lin、Yeyun Gong、Lionel M Ni、Heung-Yeung Shum 和 Jian Guo。2023。Think-on-graph：大语言模型在知识图谱上的深入且负责任的推理。arXiv 预印本 arXiv:2307.07697（2023）。[37] Rakshit Trivedi、Hanjun Dai、Yichen Wang 和 Le Song。2017。Know-Evolve：针对动态图知识图谱的深度时序推理。载于《国际机器学习会议》。PMLR，3462–3471。[38] Pham Hoang Vuong、Trinh Tan Dat、Tieu Khoi Mai、Pham Hoang Uyen 等。2022。基于 XGBoost 和 LSTM 的股价预测。《计算机系统科学与工程》40，1（2022）。[39] Jianing Wang、Qiushi Sun、Xiang Li 和 Ming Gao。2023。通过知识链提示提升语言模型的推理能力。arXiv 预印本 arXiv:2306.06427（2023）。[40] Neng Wang、Hongyang Yang 和 Christina Dan Wang。2023。FinGPT：面向金融数据集开源大语言模型的指令微调基准测试。arXiv 预印本 arXiv:2310.04793（2023）。[41] Shijie Wu、Ozan Irsoy、Steven Lu、Vadim Dabravolski、Mark Dredze、Sebastian Gehrmann、Prabhanjan Kambadur、David Rosenberg 和 Gideon Mann。2023。BloombergGPT：面向金融的大语言模型。arXiv 预印本 arXiv:2303.17564（2023）。[42] Sheng Xiang、Dawei Cheng、Chencheng Shang、Ying Zhang 和 Yuqi Liang。2022。用于金融时间序列预测的时序异构图神经网络。载于《第31届ACM信息与知识管理国际会议论文集》，3584–3593。[43] Tianbao Xie、Chen Henry Wu、Peng Shi、Ruiqi Zhong、Torsten Scholak、Michihiro Yasunaga、Chien-Sheng Wu、Ming Zhong、Pengcheng Yin、Sida I Wang 等。2022。UnifiedSKG：利用文本到文本语言模型统一并多任务化结构化知识落地。arXiv 预印本 arXiv:2201.05966（2022）。[44] Yumo Xu 和 Shay B Cohen。2018。基于推文和历史价格预测股票走势。载于《计算语言学协会第56届年会论文集（第1卷：长篇论文）》，1970–1979。[45] An Yang、Anfeng Li、Baosong Yang、Beichen Zhang、Binyuan Hui、Bo Zheng、Bowen Yu、Chang Gao、Chengen Huang、Chenxu Lv 等。2025。Qwen3技术报告。arXiv 预印本 arXiv:2505.09388（2025）。[46] Fan Yang、Zhilin Yang 和 William W Cohen。2017。用于知识库推理的逻辑规则可微学习。《神经信息处理系统进展》第30卷（2017）。[47] Yi Yang、Mark Christopher Siy Uy 和 Allen Huang。2020。FinBERT：面向金融传播的预训练语言模型。arXiv 预印本 arXiv:2006.08097

（2020）。[48] Jaemin Yoo、Yejun Soun、Yong-chan Park 和 U Kang。2021。利用具有多层上下文的数据轴Transformer实现准确的多变量股票走势预测。载于第*27*届*ACM SIGKDD*知识发现与数据挖掘会议论文集，2037–2045。[49] Hong Qing Yu 和 Frank McQuade。2025。RAG-KG-IL：一种多智能体混合框架，通过整合检索增强生成与增量知识图谱学习来减少幻觉并增强大语言模型推理。*arXiv* 预印本 *arXiv:2503.13514*（2025）。[50] Yangyang Yu、Haohang Li、Zhi Chen、Yuechen Jiang、Yang Li、Jordan W Suchow、Denghui Zhang 和 Khaldoun Khashanah。2025。FinMem：一种通过分层记忆和角色设计提升性能的大语言模型交易智能体。*IEEE*大数据汇刊（2025）。

[51] Chong Zhang、Xinyi Liu、Zhongmou Zhang、Mingyu Jin、Lingyao Li、Zhenting Wang、Wenyue Hua、Dong Shu、Suiyuan Zhu、Xiaobo Jin 等。2024 年。当人工智能遇上金融（StockAgent）：在模拟真实世界环境中基于大语言模型的股票交易。*arXiv* 预印本 *arXiv:2407.18957*（2024）。[52] Yifei Zhang、Xintao Wang、Jiaqing Liang、Sirui Xia、Lida Chen 和 Yanghua Xiao。2024 年。Chain-of-Knowledge：通过从知识图谱中学习，将知识推理融入大语言模型。*arXiv* 预印本 *arXiv:2407.00653*（2024）。[53] Haoyi Zhou、Shanghang Zhang、Jieqi Peng、Shuai Zhang、Jianxin Li、Hui Xiong 和 Wancai Zhang。2021 年。Informer：超越高效 Transformer 的长序列时间序列预测。在*AAAI*人工智能会议论文集第 35 卷。11106–11115。