

结构化字段之外，文本还剩多少 Alpha？ 以 A 股业绩预告为例的可解释因子智能体与增量检验

周梓煜（组长） 张心湑 许艺馨 张修博 刘迅羽 王佳祺

AI 金融研学实训营·AI 篇·主题四

2026 年 7 月

目录

- ① 研究目标
- ② 项目流程
- ③ 关键模块设计
- ④ 平台实现与展示

研究目标

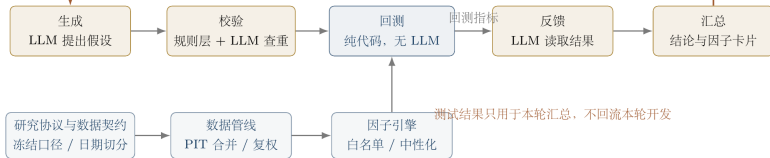
我们做什么

可解释 Alpha 因子生成智能体：LLM 提出带经济含义的候选，校验、计算与回测由本地确定性代码完成——**LLM 提假设，代码当裁判。**

- Alpha 因子：给股票打分的规则（估值、质量、波动等），关注相对表现
- 收窄后的问题：结构化字段已公开关键数字时，公告正文还能提供增量信息吗？
- 场景：A 股业绩预告；对照用结构化字段，文本增量另作处理组检验

项目流程

新协议：下一轮改写方向



组长

- GitHub 仓库管理
- Agent 页面与 workflows 实现

组员

- 研究问题拆解与生成逻辑
- 校验规则与回测指标讨论
- 工单系统汇总问题

流程：生成 → 校验 → 回测 → 反馈 → 汇总。

工单：诊断并补强不完整候选（字段、表达式、窗口、解释等），再进入回测。

生成：划定知识集，输出因子卡片

先定边界：模型只能使用约定好的知识集，不能脱离数据契约自由发挥。

- **行情：**成交量、成交额、市值、行业等
- **财务：**净利润、净资产、EPS、经营现金流，以及关键公告日
- **股票池：**按交易日限定可研究股票；并可接入业绩预告等文本

Harness 式约束

文本先提炼订单景气、成本压力等语义再数值化；输出必须是因子卡片——名称、表达式、所用字段、经济解释、生成来源。目的：可解释、可追溯，而不是凭感觉生成黑箱公式。

校验：回测前的合法性闸门

校验回答的是“能不能进回测”，不是“是否已经有效”。

- **时间与对象**：信息是否有明确时间戳；能否映射到 A 股个股或行业
- **PIT 可得性**：财务字段在交易日 T 只能使用已披露信息，防止前视
- **规则硬闸**：拦截未知字段、未来收益标签、负延迟、重复或高度相关候选
- **语义查重**：必要时由 LLM 辅助；不能替代规则层最终否决权

通过校验 $\Rightarrow$ 允许进入统一回测；不通过则拒绝或打回修改。

回测：纯代码评价因子表现

原则

回测是真正打分的环节：**不调用大语言模型**。数值必须由确定性程序计算，避免主观圆场。

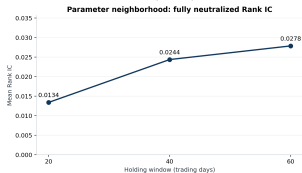
核心指标：

- Rank IC、ICIR：横截面预测力与稳定性
- 分组收益、多空收益：排序后的收益差
- 换手率、成本后多空收益：交易摩擦与可实施性

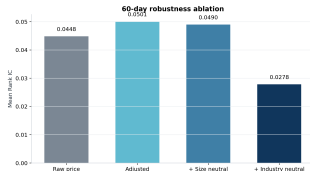
同时考察样本内外表现、信号衰减与持有期匹配；按日期切分训练/测试，禁止随机打乱。

回测证据一览

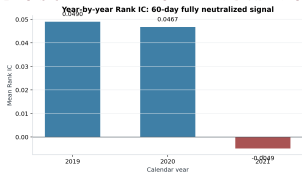
持有期参数邻域 (20/40/60 日 Rank IC)



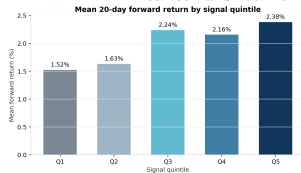
中性化消融 (行业暴露影响)



分年度 Rank IC (2021 为测试年)



五分组收益 (仅作排序描述)



反馈与汇总

反馈

把回测中不好的数字翻译成下一轮改写建议：样本内 Rank IC 接近 0、ICIR 偏低说明方向不稳；换手率到 0.6–0.8 说明交易成本偏高；与基线相关到 0.8 以上说明不够新。

关键约束：只把训练段指标喂回大模型；样本外结果坚决不回传。否则会变成“看答案改题”，样本外评估失去意义。反馈用于辅助理解与改写，不是用测试集继续优化。

汇总

将通过校验与回测的候选收成「因子卡片」：经济解释、所用字段、表达式与回测情况。

这一步确认我们得到的是可用、可追溯的因子交付物，后续才能进入投资组合构建等下一步；负结果同样保留，不粉饰。

为什么做研究面板

定位

面板不是普通展示页，而是把 Alpha 因子研究全流程做成可演示、可复核的平台，让非技术成员也能看懂：一个想法如何进入系统、经过哪些校验、为何得到当前结论。

要解决的问题

- 只有命令行和结果文件时，流程不透明、证据难核对
- 分不清哪些是模型生成的解释，哪些是代码算出的数值
- 正结果容易被展示，负结果容易被丢掉

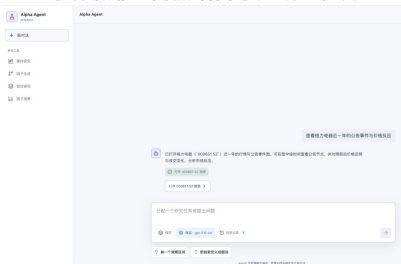
用户需求：顺着流程看完一次研究；看见候选通过/拒绝及原因；确认 Rank IC、ICIR 由确定性代码计算；负结果如实保留。

面板怎么用

使用流程：用自然语言分配研究任务 → Agent 编排打开对应工具页 → 在事件/因子/组合视图中核对结果。

左侧提供事件研究、因子生成、组合研究、因子结果等入口；底部提示：Agent 负责理解与编排，数值始终由确定性代码生成。

对话编排：自然语言下达研究任务



事件研究：公告节点与价格反应对照



扫码进入研究面板



<https://alpha.cihua.run/>

- 离线研究回放：固化轨迹，现场可复核
- 白名单工具：事件 / 因子 / 组合
- 公开界面只读事实记录，不在浏览器现场回测

恳请各位老师
批评指正