

房地产行业深度报告

地产+AI 工具系列报告之四：从地产投研到交易——OpenClaw 的跨界实践（如何训练一只会交易能风控的“龙虾”）

增持（维持）

投资要点

■ 先回答一个问题，为什么一个地产团队要做量化交易实践？三个原因，从近到远：**第一**，验证 OpenClaw 框架的通用性。东吴地产前三篇报告证明了 OpenClaw 在地产行业能跑通——做日报、搭选股平台、提效投研。但“在一个行业跑通”和“框架本身通用”之间还有距离。交易是一个和地产投研完全不同的领域：实时性要求高（盘中秒级决策）、容错空间小（真金白银亏损）、外部依赖多（行情 API、券商接口、LLM 服务）。如果 OpenClaw 在这么“硬”的场景也能跑通，那它在任何垂直领域的可行性就不需要再被质疑。**第二**，探索 AI Agent 在金融决策中的能力边界。地产投研场景里，AI 犯错的代价是“报告写错了，改一下”——可逆、低成本。但在交易场景里，AI 犯错的代价是“下单亏钱了”——不可逆、高成本。这逼迫我们设计一套比“辅助场景”严密得多的安全机制：十道风控防线、双模型交叉验证、幻觉检测、熔断冷却。这些机制的设计经验，对任何想用 AI Agent 做高自主度决策的场景（不只是交易，也包括信贷审批、合规检查、自动报价）都有参考价值。**第三**，给卖方研究提供一个“AI Agent 落地”的范本。卖方研究谈论 AI，往往停留在“模型能力评测”或“提示词工程”层面。本篇展示的是一个完整的生产系统——从数据采集到决策到执行到风控到监控，全链路闭环。读者看到的不是“AI 能做什么”的展望，而是“AI 已经在做什么”的实录。

■ **系统架构闭环设计：全自动选股—评分—交易—风控—监控流水线**：构建“调度层—选股层—评分层—决策层—执行层—观测层”六层架构，实现美股日频全自动交易。系统具备定时调度、多路选股、三引擎量化评分、双 LLM 交叉验证、信号生成、券商自动下单、十道风控防线与全链路可观测能力。

■ **选股体系稳健：四路入池+保底机制，杜绝数据源失效风险**：采用 Finviz 筛选、异动扫描、SEED_TICKERS 种子池、杠杆 ETF 专属池四路构建候选池，并配置原子写入、最小阈值、Top180 截断与行业集中度监控，确保候选池稳定、高质量、低噪音。

■ **评分机制科学：三引擎专业化打分+AI 交叉验证，降低单一模型偏差**：针对成长股、杠杆 ETF、小盘股分别设计专用评分引擎；引入 MiniMax 与 GLM 双模型独立判别并交叉融合，结合量化、AI、反向指标加权形成最终得分，提升决策可靠性与可解释性。

■ **风控防线严密：在设置止损硬约束之外，采用三层动态止盈止损机制——ATR 自适应（止损宽度联动个股波动率，高波动放宽、低波动收紧）、峰值阶梯锁利（盈利越大回撤容忍越紧，逐级锁利）、VIX 联动（恐慌时收紧保命、平静时放宽让利润跑），配合 48 小时冷却期、行业分散、动态预算管理，实现进攻性与安全性平衡。**

■ **策略价值明确：简单有效、低门槛可复现、纪律性远超人工**：不以复杂因子与高频交易取胜，而依靠机器纪律性克服人性弱点；回测显示夏普比率 1.24、最大回撤-19%，适合追求稳健收益、厌恶大幅回撤的个人与小型机构使用。

■ **风险提示**：AI 评级仅供参考，不构成投资建议；市场风险；流动性风险；底层资产风险；模型风险。

2026 年 05 月 13 日

证券分析师 姜好幸

执业证书：S0600525110001

jianghx@dwzq.com.cn

证券分析师 刘汪

执业证书：S0600526030001

liuwang@dwzq.com.cn

行业走势



相关研究

《地产+AI 工具系列报告之三：基于多模型联合决策的 C-REITs 智能评级与跟踪分析体系》

2026-04-14

《基于 OpenClaw 的房地产股票投研生产力提升实践》

2026-03-11

《基于量化分析与大模型的房地产股票 AI 选股系统》

2026-03-05

内容目录

1. 引言: OpenClaw 与 InvestmentClaw 体系	8
2. 写给新读者的话	9
2.1. 术语解释.....	9
2.2. 致东吴地产 AI 系列报告的读者.....	12
2.2.1. 亮点: 并非“又一个量化系统”.....	12
2.2.2. 为什么要做这项工作?	12
2.2.3. 与大型量化机构的差异.....	13
2.2.4. 为什么“简单”但仍然有价值?	14
3. 系统架构	15
3.1. 概述.....	15
3.2. 完整架构图.....	15
3.3. 分层职责.....	17
3.4. 分层间的数据流与接口契约.....	17
3.5. 两套代码仓库的关系.....	18
3.6. 故障隔离与降级策略.....	19
3.7. 可观测性设计.....	20
4. 选股池体系	21
4.1. 四路选股入池.....	21
4.2. 每路选股的技术细节与失败处理.....	21
4.2.1. Finviz 筛选:	21
4.2.2. 异动扫描:	21
4.2.3. SEED_TICKERS:	22
4.2.4. 杠杆 ETF 专属池:	22
4.2.5. 四路合并的去重逻辑:	22
4.3. Finviz4 条筛选策略.....	23
4.4. 策略参数的选择逻辑与回测验证.....	23
4.5. SEED_TICKERS 稳定种子.....	24
4.6. SEED_TICKERS 的完整列表与更新机制.....	24
4.7. 候选池写入保护.....	26
4.8. Top180 截断.....	27
4.9. 候选池质量监控.....	27
5. 三引擎评分	29
5.1. 为什么采用“三引擎”.....	29
5.2. 每个引擎的评分构成 (“三个评委的打分标准”)	29
5.3. 各引擎核心因子详解.....	30
5.4. 评分拉伸 (“提升分数区分度”)	33
5.5. 后置调节规则.....	33
5.6. 波动率过滤.....	34
5.7. 因子数据链路.....	35
5.7.1. 第一步: 外部数据源.....	35
5.7.2. 第二步: 数据获取层.....	36
5.7.3. 第三步: 清洗与计算.....	36

5.7.4. 数据延迟对评分的影响:	36
6. AI 多模型判别: 让两个 AILLM 独立打分, 然后交叉验证	37
6.1. 流程概述.....	37
6.2. 模型选择理由与成本分析.....	38
6.2.1. 为什么是“双模型”而非单模型	38
6.2.2. 推理能力与独立性.....	38
6.2.3. 成本对比: 国产推理模型 vs 海外前沿模型	39
6.2.4. 线路稳定性与可用性.....	40
6.2.5. 推理模型的工程价值.....	41
6.2.6. 未来升级路线.....	41
6.3. InvestmentClaw 的模型配置 (独立体系)	42
6.4. LLM 评分执行流程.....	42
6.4.1. Step1 Prompt 组装.....	42
6.4.2. Step2 并发双模型调用	43
6.4.3. Step3 交叉验证融合 (_fuse_llm_scores)	43
6.4.4. Step4 注入 final_score.....	44
6.4.5. 熔断与冷却.....	44
6.5. 幻觉处理.....	45
6.6. LLM 熔断与跨批次冷却.....	46
6.7. final_score 下限保护.....	46
6.8. LLM 评分的局限性.....	46
6.8.1. 局限性 1: 训练数据的时效性.....	46
6.8.2. 局限性 2: 无法获取非公开信息.....	47
6.8.3. 局限性 3: 对复杂财务工程的理解有限.....	47
6.8.4. 局限性 4: 情绪化与一致性.....	47
6.8.5. 局限性 5: 成本约束下的信息压缩.....	48
6.8.6. 总结: AI 评分的正确用法	48
7. 信号生成: 从分数到实际行动	48
7.1. Action 分级.....	48
7.2. 买入候选 TopN 筛选.....	49
7.3. 信号终调规则.....	50
7.3.1. 现金比例.....	50
7.3.2. 单日买入笔数上限.....	50
7.3.3. 行业集中度.....	50
7.3.4. 单票红线.....	50
7.3.5. 杠杆 ETF 合计红线	51
7.3.6. 持仓数上限.....	51
7.4. Action 状态机与冲突处理.....	51
7.5. 信号生成完整实例 (“看系统怎么过一天”)	52
7.5.1. Step1 选股 (21:00SGT)	53
7.5.2. Step2 三引擎评分 (21:05SGT)	53
7.5.3. Step3 波动率过滤	53
7.5.4. Step4 AI 评分 (21:10SGT)	54
7.5.5. Step5 final_score 计算.....	54

7.5.6. Step6 买入候选筛选	54
7.5.7. Step7 三道过滤	55
7.5.8. Step8 信号终调 (“主编审稿”)	55
7.5.9. Step9 最终信号	55
8. 交易执行: 从信号生成到成交落地	56
8.1. 单笔交易流程 (以买入为例)	56
8.2. 券商终端的 API 超时保护	58
8.3. 仓位权重计算	58
8.4. cmd_trade 完整执行流程	59
8.4.1. Step1 启动自检	59
8.4.2. Step2 预算计算 (核心动态逻辑)	60
8.4.3. Step3 账户+持仓快照	60
8.4.4. Step4 三引擎评分	61
8.4.5. Step5 生成买卖候选	61
8.4.6. Step6 卖单执行	61
8.4.7. Step7 买单执行	62
8.4.8. Step8 汇总推送	62
8.5. cmd_monitor 盘中监控循环	62
8.6. 其余 CLI 命令速查	63
9. 风控体系: 风险约束的十道防线	63
9.1. 风控参数全景	63
9.2. 参数读取方式	65
9.3. 日亏上限含浮亏	65
9.4. 持仓生命周期管理	65
9.5. 止损卖出后冷却, 而非永久纳入黑名单	66
9.6. 三桶预算与杠杆 ETF 合计校验	67
9.6.1. 三桶预算模型	67
9.6.2. RiskManager.can_trade() 十道红线	68
9.6.3. 日预算 N 动态计算	69
9.6.4. 持仓数自适应	70
10. 盘中监控与守护	72
10.1. 监控架构与 InvestmentClaw 的交互角色	72
10.2. 三层可观测性的互补关系	72
10.3. InvestmentClaw 的“主动模式”	73
10.4. 监控架构	73
10.5. Gateway 守护	74
10.6. 告警节流与冷却	74
10.7. Profile 感知	74
10.8. Gateway 热重载	74
11. Telegram 通知	74
11.1. 通道设计	74
11.2. 速率控制	75
11.3. 消息类型总览	75
12. 文件与路径	75

12.1. 核心目录结构.....	75
12.2. 工作区路径解析.....	76
13. Cron 调度	76
13.1. 一天的时间表（北京时间）	76
13.2. 完整 crontab.....	76
13.3. 调度设计依据.....	77
14. 研究框架：因子、回测、调参	78
14.1. 研究流程.....	78
14.2. 32 因子库.....	78
14.3. 3 年因子 IC 排名.....	79
14.4. 近 90 天 vs 历史 3 年 IC 对比.....	80
14.5. 回测引擎.....	80
14.6. 回测核心结论.....	80
14.7. CPCV 交叉验证.....	81
14.8. OptunaBayesian 调参.....	81
15. 总结：从 AI 辅助到 AI 自主：本系统的核心价值与启示	82
15.1. 全文回顾.....	82
15.1.1. 架构层面（第 2 章）	82
15.1.2. AI 多模型判别（第 5 章）	82
15.1.3. 从信号到成交（第 6-7 章）	82
15.1.4. 风控体系（第 8 章）	82
15.1.5. 运行保障（第 9-12 章）	82
15.1.6. 研究框架（第 13 章）	83
15.2. 核心方法论贡献.....	83
15.2.1. 分层解耦的系统设计范式。	83
15.2.2. 量化 + AI 的互补融合范式。	83
15.2.3. “主动但不失控”的自主决策范式。	83
15.2.4. 可观测性作为系统的基础能力。	84
15.3. 策略定位与核心优势.....	84
15.3.1. 定位明确：不拼信息差和速度差，聚焦“认知差”	84
15.3.2. 以简单换取稳健.....	84
15.3.3. 可解释本身就是竞争力.....	84
15.3.4. 低门槛可复现.....	84
15.4. 对卖方研究 AI 应用的三点启示.....	85
16. 风险提示	85

图表目录

图 1: 美股量化交易系统架构.....	16
图 2: 数据链路全景.....	35
图 3: AI 评分流程图.....	37
图 4: 交易执行流程图.....	57
图 5: 持仓状态示意图.....	66
图 6: Gateway 守护流程图.....	73
表 1: 体系层次.....	8
表 2: InvestmentClaw 与大型量化机构对比.....	13
表 3: 美股量化交易系统架构分层职责.....	17
表 4: 多层降级策略.....	19
表 5: 三层观测的互补关系:	20
表 6: 四路选股入池.....	21
表 7: Finviz4 条筛选策略.....	23
表 8: 单因子回测.....	24
表 9: SEED_TICKERS 列表 (数据截至 2026 年 4 月 30 日)	25
表 10: 每日质量仪表盘.....	28
表 11: 候选池行业分布监控示例.....	28
表 12: 引擎评分构成.....	29
表 13: 实际评分示例.....	30
表 14: growth 引擎因子 (4 大类, 共 12 个因子)	30
表 15: leverage 引擎因子 (3 大类, 共 8 个因子)	31
表 16: smallcap 引擎因子 (3 大类, 共 7 个因子)	32
表 17: 因子数据来源与延迟.....	32
表 18: 波动率过滤示例.....	34
表 19: 回测数据.....	35
表 20: 外部数据源.....	35
表 21: 数据延迟对评分的影响.....	36
表 22: 模型成本对比.....	39
表 23: 模型单次费用.....	39
表 24: 实际工作量折算.....	40
表 25: 未来升级路线.....	41
表 26: InvestmentClaw 的当前配置.....	42
表 27: 交叉验证融合.....	43
表 28: 交叉验证融合实例.....	44
表 29: Final-score 实例.....	44
表 30: 六种行动建议.....	48
表 31: 实例三引擎评分.....	53
表 32: 实例 AI 评分.....	54
表 33: 实例 final-score 计算.....	54
表 34: 实例最终信号.....	55

表 35:	券商终端超时保护.....	58
表 36:	仓位权重计算.....	59
表 37:	CLI 命令速查.....	63
表 38:	风控参数类别.....	64
表 39:	系统局限性.....	65
表 40:	三桶预算模型.....	67
表 41:	三桶预算模型实例.....	68
表 42:	分差门槛.....	71
表 43:	最弱评分门槛.....	71
表 44:	每日换仓次数上限.....	71
表 45:	InvestmentClaw 三层可观测性设计.....	72
表 46:	InvestmentClaw 主动模式.....	73
表 47:	Cron 调度一天时间表.....	76
表 48:	32 因子类别.....	79
表 49:	3 年因子 IC 排名.....	79
表 50:	近 90 天 vs 历史 3 年 IC 对比.....	80
表 51:	回测表现.....	80
表 52:	CPCV 交叉验证.....	81
表 53:	TOP3 稳健参数.....	81

1. 引言：OpenClaw 与 InvestmentClaw 体系

1) OpenClaw 是什么？

OpenClaw 是目前最火的开源 AI Agent 框架，提供完整的 agent 基础设施——包括任务调度、工具调用、记忆管理、多模型路由、可观测性等核心能力。你可以把它理解为 AI Agent 领域的“操作系统”，开发者基于它可以快速构建各种垂直领域的智能体。

2) InvestmentClaw 是什么？

本报告描述的量化交易系统，是基于 OpenClaw 框架构建的垂直领域 agent，名字叫 InvestmentClaw。它专注于美股全自动交易，是 OpenClaw 生态中的一个具体应用实例。

表1：体系层次

层级	名称	角色	类比
底层框架	OpenClaw	开源 AI Agent 基础设施	Android 操作系统
上层应用	InvestmentClaw	基于 OpenClaw 构建的美股交易 agent	某个具体的 App
使用者	用户	机构投资者	手机用户

数据来源：东吴证券研究所整理

3) 为什么选 OpenClaw？

OpenClaw 框架为 InvestmentClaw 提供了以下核心能力：

- Agent 生命周期管理：启动、调度、重启、健康检查——InvestmentClaw 作为其中一个 agent 实例运行
- 多模型路由：支持 MiniMax、GLM、DeepSeek 等多种 LLM 的灵活切换和熔断保护
- 记忆系统：短期记忆（当日交易状态）+ 长期记忆（历史持仓、归因报告）
- 工具调用：调用 yfinance、Finviz、券商终端 API 等外部工具获取数据
- 可观测性：Telegram 通知、结构化日志、运行时监控
- Gateway 架构：统一的 API 网关，管理所有外部服务的连接和认证

4) InvestmentClaw 的核心人格（来自 IDENTITY.md / SOUL.md）

InvestmentClaw 不是通用聊天机器人，它有明确的交易人格：

- 进攻型：系统发现机会信号后即刻响应，重视 alpha 收益，减少冗余确认。
- 先查数据，再说话：回答交易问题之前，必须确认三件事——当前持仓、可用现金、今天是不是交易时段。不凭记忆、不凭直觉。
- 结论明确：对 NVDA 等标的形成看多判断时直接表达，对 TSLA 等交易 setup 较弱的标的明确提示风险。模糊表述会降低交易执行效率。
- 指令清晰即直接执行：当用户发出“买 NVDA 50 股”等明确指令时，只要不触及风控红线，系统将直接执行，并以简洁方式反馈成交结果。
- 唯一可拒绝执行的场景：触及风控红线（单票 > 30%、杠杆 ETF 合计 > 50%、要求 margin）。此时系统将直接拒绝，并说明拒绝原因。

5) InvestmentClaw 不负责什么？

InvestmentClaw 拥有独立 profile 与独立 Telegram bot，与日常通用助手完全隔离：如果用户在 InvestmentClaw 窗口提出非交易类问题，系统将提示当前窗口与问题场景不匹配。

2. 写给新读者的话

2.1. 术语解释

三引擎

系统将股票划分为三类：成长股（如 AAPL）、杠杆 ETF（如 TQQQ）和小盘股（如部分 AI 概念股）。不同类别标的分别由对应评分引擎处理：growth 引擎关注基本面（ROE、营收增速），leverage 引擎关注大盘趋势（RSI、均线），smallcap 引擎关注市场情绪（成交量异动）。该设计反映出不同类型资产的定价逻辑差异。详见章节 5.1。

SEED-TICKERS

系统每日从 Finviz、异动扫描等渠道收集候选股票。若外部渠道同时故障，候选数据可能中断。SEED-TICKERS 是约 40 只核心标的（Mega Cap 大公司 + 主流 ETF + 防御品种）的固定列表，相当于基础备选池，可在外部数据源不可用时维持最低覆盖。详见章节 4.5。

Finviz 选股

Finviz 是一个免费的美国股票筛选网站 (finviz.com)，用户可在该网站设置筛选条件（如“市值 >100 亿”、“ROE >15%”），并获取符合条件的股票列表。系统每日自动访问 Finviz，并基于 4 条预设策略筛选股票。详见章节 4.3。

异动扫描

系统实时监控 5 类日内异常事件：创 52 周新高、成交量暴增、单日大涨、单日大跌、轧空（做空者被迫买回）。这些事件往往意味着有重大消息（如财报超预期、行业政策变化），可能带来交易机会。详见 章节 4.2。

final_score

公式: $\text{final_score} = 0.6 \times \text{quant_score} + 0.2 \times \text{llm_fused} + 0.2 \times \text{inverse_score}$

量化评分基于历史数据，稳定性相对更高，因此权重最大；AI 模型存在“幻觉”风险，故仅占 20%；反向指标用于风险校正，避免高估风险。详见章节 6.4。

IC / IC 驱动权重

IC (Information Coefficient) 是量化领域的核心概念，衡量某个指标（如“20 日均线斜率”）与未来收益的相关性。IC=0.06 表示这个指标有 6% 的预测能力。系统根据 IC 大小给不同指标分配权重：IC 高的指标权重高，IC 低甚至为负的指标权重归零。详见 章节 14.3。

Regime

系统将市场状态划分为四类：BULL-STRONG（强势牛市）、BULL-WEAK（弱势牛市）、BEAR-WEAK（弱势熊市）和 BEAR-STRONG（强势熊市）。不同市场状态对应不同配置策略：强势牛市提升成长股配置权重，弱势熊市提升防御品种权重。详见 章节 14.5。

夏普比率

夏普比率 = 年化收益 / 年化波动。策略 A 年化赚 30%、最大亏 20%（夏普 ≈ 1.5），策略 B 年化赚 20%、最大亏 5%（夏普 ≈ 4）。策略 B 的夏普更高，因为它用更小的波动换来了接近的收益。夏普 >1 就是“划算”，夏普 < 0.5 就是“不划算”。详见章节 14.6。

CPCV

CPCV (Combinatorial Purged Cross-Validation) 是量化策略中用于检验过拟合的方法。CPCV 将历史数据划分为多段，并在不同时间段分别测试策略表现。若策略在多个时间段均表现稳定，则说明其预测能力并非单纯来自对历史样本的拟合。详见 章节 14.7。

守护进程

系统包含多个后台服务（如 Gateway 连接券商、定时选股任务），守护进程每 5 分钟检查一次服务状态。如发现 Gateway 异常，系统会自动重启并发送 Telegram 告警，实现持续性运行监控。详见章节 10.5。

VIX（恐慌指数）

VIX 是芝加哥期权交易所编制的指数，基于标普 500 期权价格计算。VIX<15 表示市场很平静（大家很淡定），VIX>25 表示市场在恐慌（大家很害怕）。系统根据 VIX 调整每天的交易次数：平静日少交易，恐慌日多交易（因为波动大机会多）。详见章节 9.6.3。

移动止损

移动止损 (Trailing Stop) 是“动态版止损”。比如 100 美元买入，涨到 130 美元 (+30%)，然后回落到 124.8 美元（从最高点回落 4%），触发移动止损卖出。最终盈利 24.8%。这样既让利润奔跑，又防止利润全部回吐。详见章节 8.4.5。

杠杆 ETF

杠杆 ETF（如 TQQQ）用衍生品放大指数的涨跌幅。纳斯达克 100 涨 1%，TQQQ 涨约 3%；纳斯达克 100 跌 1%，TQQQ 跌约 3%。收益和风险都被放大了。适合在牛市中放大收益，但在熊市中也会放大亏损。详见章节 5.1。

黑名单

blacklist.txt 中列出的股票系统不会买入。例如，若某股票被 SEC 调查财务造假，或用户出于合规、偏好等原因希望规避该标的，可将其纳入黑名单，以降低道德风险或法律风险。

冷却期

某股票触发止损卖出后，系统会在 48 小时内暂停重新买入该股票。该机制并非永久限制；48 小时后若评分仍然较高，系统可重新评估买入。其目的在于避免止损后过早重新建仓。详见章节 9.5。

三桶预算

系统将交易按来源划分为三个相互隔离的“桶”： - instruction 桶：用户手动下达的订单（如“买 NVDA 50 股”），不受每日交易次数限制 - spontaneous 桶：系统自发触发的买入订单，受每日预算 N 限制（2-15 笔） - risk 桶：止损/止盈/强平订单，不受普通交易次数限制，优先保障风险控制

这样设计是为了防止“系统自己过度自发交易把次数用完，导致该止损的时候无法执行止损”。详见章节 9.6.1。

持仓生命周期

每笔持仓经历 5 个状态： - open: 刚下单，还没完全成交 - held: 已成交，正常持有中 - trailing: 涨到一定程度，开始追踪移动止损 - closing: 触发卖出，正在执行 - closed: 完全卖出，结束

状态持久化到 SQLite 数据库，进程重启后可自动恢复。避免“判断已持仓实际未成交”或“判断已卖出实际挂单未成交”等状态不一致问题。详见 章节 9.4。

2.2. 致东吴地产 AI 系列报告的读者

如果读者从东吴地产前三篇 AI Agent 系列报告一路阅读至此，或许会自然产生一个疑问：一个地产研究团队，为什么要开发并讨论一套美股量化交易系统？

2.2.1. 亮点：并非“又一个量化系统”

本系统的核心亮点并不在于收益曲线本身。坦率而言，本系统并不试图与 Two Sigma 等顶级量化机构在回撤控制或策略容量上直接比较。其真正价值在于以下三方面能力的叠加：

① AI Agent 从“辅助”走向“自主”的完整落地。

在前三篇报告中，AI Agent 更多承担的是“辅助工具”角色，帮助地产研究员撰写日报、筛选股票、整理数据，但最终决策权始终保留在人手中。本篇所讨论的 InvestmentClaw 则进一步转向“协同搭档”：系统可在盘中自主完成选股、评分、下单和止损，人只负责设定风控边界，而不再逐笔审批。这标志着 AI Agent 能力边界的重要跃迁：从 Copilot 迈向 Autopilot。

② 双 AI 交叉验证与熔断机制的工程化实践。

采用两个推理模型独立评分并进行融合，本身并不是全新的概念；但将“幻觉检测—方向分歧处理—单模型降权—双模型失败兜底—熔断冷却”这一完整链路工程化，并部署在生产环境中持续运行，是本系统相较于一般 AI 投研工具的核心实践价值。第五章中的每一节，均对应一个真实工程问题及其可复用的解决方案。

③ 全链路可观测、可审计、可回溯。

从候选池构建，到三引擎评分、AI 评分、信号生成、交易执行和风控管理，系统每一步均保留结构化日志。任意交易日的任意一笔操作，均可完整还原其决策链条。这并非“黑箱模型给出的交易建议”，而是在白箱机制下运行的自主决策系统——自主，但并非不可解释。

2.2.2. 为什么要做这项工作？

我们认为，开展这一工作的原因主要有三点，由近及远分别如下：

第一，验证 OpenClaw 框架的通用性。

我们前三篇报告已经证明，OpenClaw 能够在地产行业场景中跑通，包括日报生成、选股平台搭建和投研效率提升等。但“在一个行业中跑通”与“框架本身具备通用性”之间仍存在距离。美股交易是一个与地产投研显著不同的应用场景：其对实时性要求更高，需要盘中秒级响应；容错空间更小，错误决策可能直接导致资金损失；外部依赖也更加复杂，包括行情 API、券商接口和 LLM 服务等。如果 OpenClaw 能够在这一类更“硬”的场景中稳定运行，则其在其他垂直领域中的适用性将具备更强说服力。

第二，探索 AI Agent 在金融决策中的能力边界。

在地产投研场景中，AI 出错的代价往往是“报告内容有误，可以修改”，属于可逆、低成本错误。但在交易场景中，AI 出错的结果可能是“错误下单并产生亏损”，属于不可逆、高成本错误。这一约束迫使我们设计一套远比“辅助型场景”更严密的安全机制，包括十道风控防线、双模型交叉验证、幻觉检测和熔断冷却等。这些机制的设计经验，不仅适用于交易系统，也可为信贷审批、合规检查、自动报价等高自主度决策场景提供参考。

第三，为卖方研究提供一个“AI Agent 落地”的范本。

当前卖方研究讨论 AI，较多仍停留在“模型能力评测”或“提示词工程”层面。本篇展示的是一个完整的生产系统，覆盖数据采集、决策生成、交易执行、风控管理和运行监控等全链路闭环。读者看到的不是“AI 未来能做什么”的展望，而是“AI 当前已经能够做什么”的实证记录。

2.2.3. 与大型量化机构的差异

需要首先明确的是：本系统不是，也并不试图成为 Renaissance Technologies 或 Two Sigma。

表2: InvestmentClaw 与大型量化机构对比

维度	大型量化机构	InvestmentClaw
资金规模	10 亿—1000 亿美元	万美元级，面向个人或小分队账户
数据	专线行情、卫星数据、另类数据，如信用卡数据、船舶 AIS 等	免费公开数据，如 yfinance、Finviz，以及 LLM 生成分析
模型	数百个因子、深度学习模型、高频信号	约 20 个可解释因子 + 双 LLM 交叉验证
执行	共置服务器、微秒级抢单	长桥 API 限价单/市价单，秒级成交
频率	日内交易，分钟级甚至 tick 级	日频为主，每日 1—2 轮交易决策
团队	数十名数学、物理、计算机等方向 PhD	1 人 + OpenClaw + 2 个 LLM
核心壁垒	数据、算力、人才和资金构成的综合护城河	透明度、可审计性和低门槛可复现

数据来源：东吴证券研究所整理

差异的本质在于：大型量化机构主要在“信息差”和“速度差”上竞争，而本系统更聚焦于“认知差”。

信息差：机构掌握的信息更全面、更及时，这一点个人或小组难以追赶，也不是本系统的竞争重点。

速度差：机构在交易执行速度上具有显著优势，这同样不是本系统试图突破的方向。

认知差：面对同样的公开信息，系统能否比人类更冷静、更稳定、更少受情绪干扰，这是本系统试图提升的核心能力。

2.2.4. 为什么“简单”但仍然有价值？

“简单”并非缺陷，而是有意识的设计选择，主要理由如下：

1. 复杂性往往是收益稳定性的敌人。

在因子投资领域，过于复杂的模型更容易出现过拟合、解释困难和稳定性不足等问题。相较之下，简单的多因子模型，尤其是 10—30 个因子构成的模型，往往在可解释性与稳健性之间取得更好平衡。本系统采用约 20 个可解释因子，并叠加两个 LLM 进行定性判断，模型复杂度控制在“足以区分优质与劣质标的”但“不至于过度拟合”的区间内。

2. 可解释性本身就是价值。

传统量化黑箱往往只能告诉用户“赚了”或“亏了”，但难以解释“为什么赚”或“为什么亏”。本系统的每一笔交易均可追溯完整决策链：三引擎分别给出何种评分，两个 AI 模型分别给出何种判断，模型之间为何一致或分歧，最终信号如何生成，订单价格如何确定。只有能够解释，才有可能迭代；只有能够迭代，系统才具备持续改进的基础。

3. 低门槛可复现。

本系统所使用的主要工具，包括 yfinance、Finviz、长桥 API、MiniMax/GLM 等，均为个人开发者可获取的工具。其中 yfinance 和 Finviz 为免费数据源，长桥 API 支持个人开户，MiniMax/GLM 则采用 API 按量计费模式。系统代码基于开源 OpenClaw 框架。这意味着，具备一定编程基础的个人开发者，可以在 1—2 周内复现一套类似系统。我们希望展示的并不是“系统多么难以复制”，而是“这一能力正在变得可获得”。

4. “简单”并不等于“无效”。

本系统的核心收益来源并非某个高度复杂的因子，而是纪律性。许多人类投资者难以长期稳定执行的交易纪律，机器可以严格遵守：

人可能因恐慌而抛售，系统不会。止损基于预设规则，而非情绪反应。

人可能追涨杀跌，系统不会。评分低于阈值则不买入，高于阈值则不轻易卖出。

人可能反复犹豫、选择“再等等看”，系统不会。信号触发后即按规则执行。

人可能忘记设置止损，系统不会。每笔持仓均自动配置移动止损机制。

这些“纪律优势”并不依赖复杂模型，但对长期投资收益和回撤控制具有实质影响。

5. 从“简单”走向“复杂”的路径是开放的。

本系统当前版本可视为 v1.0，它并非终点，而是后续迭代的起点。第十四章中的研究框架，已经为未来升级预留接口，包括 IC 驱动权重调整、Regime 自适应、CPCV 过拟合检验和因子衰减监控等。当 v1.0 稳定运行并积累足够交易数据后，系统可以逐步引入更复杂的模型。但无论如何迭代，底层基础必须是稳固、透明的。

一句话总结：本系统的价值并不在于“比量化巨头更强”，而在于“让普通投资者也能够获得纪律化、自动化、可审计的量化交易能力”。这既是 AI Agent 能力普惠化的一个缩影，也是东吴地产团队从“AI 辅助研究”走向“AI 自主决策”的第三步。

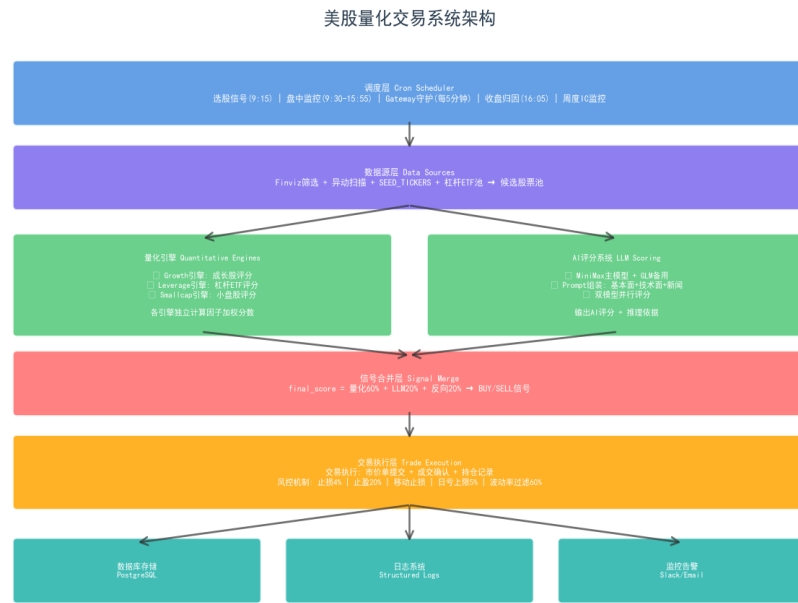
3. 系统架构

3.1. 概述

InvestmentClaw 在每个交易日固定时段自动运行“选股 → AI 评分 → 下单 → 风控”流水线，并通过 Telegram 与用户交互。系统基于 OpenClaw 开源 AI Agent 框架构建，具备可观测、可回滚、可灰度切换等特征。其核心功能包括：自动筛选候选股票、调用多模型进行评分、对高分标的执行交易，并对已持仓标的持续进行止损、止盈与移动止损监控。全流程通过 Telegram 实时推送，用户无需持续盯盘即可掌握系统运行状态。InvestmentClaw 的执行风格强调进攻性、数据优先、结论明确与指令清晰后直接执行。

3.2. 完整架构图

图1: 美股量化交易系统架构



数据来源: 东吴证券研究所绘制

自上而下逐层解释:

第一层: 调度层 - Cron 是 macOS 系统中的定时任务工具, 相当于系统调度器。系统设置了多项定时任务: - 每天 21:00 (SGT, 对应美股开盘前 30 分钟): 选股 + 评分 - 每 10 分钟一次 (21:00-05:00): 盘中监控 (检查持仓盈亏、触发止损止盈) - 每 5 分钟一次: 检查 Gateway (连接券商的服务) 是否处于可用状态 - 04:30: 收盘后进行当日盈亏归因分析 - 周日 17:00: 检查策略中各类指标是否仍然有效

第二层: 选股层 - 系统每天从 4 个渠道收集候选股票: Finviz 筛选、异动扫描、SEED_TICKERS 种子池、杠杆 ETF 专属池。 - 模块: universe_builder.py

第三层: 评分层 - 初筛 (三引擎评分): 对入池候选股票分别由 growth 引擎、leverage 引擎、smallcap 引擎进行评分。每只股票仅在其所属引擎中评分。 - 复核 (AI 多模型评分): 初筛得分最高的 Top N (默认 20 只) 进入复核环节, 由两个 AI 模型 (MiniMax-M2.7 和 GLM-5.1) 独立评分并进行交叉验证。 - 模块: scoring.py + engine.py + LLM

第四层: 决策层 (“排名”) - 量化分 (60%) + AI 分 (20%) + 反向分 (20%) 合并成 final_score, 然后生成具体的 “行动建议” (买/卖/持有/止损/止盈)。 - 模块: engine.py signal 生成

第五层: 执行层 - trader.py 是连接券商并执行交易的核心程序, 负责计算买入股数、提交限价单、等待成交、记录日志及推送通知。

第六层: 观测层 - 券商终端: 将订单发送至券商执行 - Telegram: 推送买卖标的、交易原因及执行结果 - SQLite/JSONL: 存储全部交易记录, 便于事后查询与审计

3.3. 分层职责

表3: 美股量化交易系统架构分层职责

层	模块	职责	通俗解释
调度层	cron + launchd	时间触发、守护进程	定时调度与健康检查: 按计划触发任务, 并在服务异常时自动重启
选股层	universe_builder.py	候选池构建(Finviz + 异动 + 种子 + 杠杆 ETF)	候选标的发现: 每日从市场中筛选潜在交易机会
评分层	scoring.py + engine.py + LLM	量化因子评分 + AI 判别, 合并为 final_score	评分体系: 对候选股票进行量化与 AI 评分, 确定后续处理优先级
决策层	engine.py signal 生成	根据 final_score 和现有持仓生成 buy/hold/sell action	决策模块: 根据评分与持仓状态决定买入、持有或卖出
执行层	-trader.py	下单、风控、持仓管理、Telegram 通知	执行模块: 完成订单提交、成交跟踪与风险监控
观测层	SQLite + JSONL + TG + 归因 + IC 监控	事后审计、策略健康度跟踪	审计模块: 支持事后复盘并监控策略健康度

数据来源: 东吴证券研究所绘制

3.4. 分层间的数据流与接口契约

逐层的数据流:

调度层 → 选股层: - 输入: 无 (cron 只是触发, 不传数据) - 输出: 一个启动信号

选股层 → 评分层: - 输入: universe_latest.jsonl (候选股票列表, 每行一只股票, 包含代码、市值、行业等基础信息) - 输出: 同一份文件, 但被加载到内存中准备评分

评分层 → 决策层: - 输入: 三引擎评分结果 (scores 表, 包含每只股票的生长/leverage/smallcap 得分) + AI 评分结果 (llm-fused 分数) + 反向评分 (inverse_score) - 输出: final_score 列表 + action 建议 (BUY/HOLD/SELL 等)

决策层 → 执行层: - 输入: signals_latest.jsonl (信号文件, 每行一只股票及其 action、target-weight、理由) - 输出: 无 (执行层自己读取信号文件)

执行层 → 观测层: - 输入: 成交结果 (成功/失败、成交价格、成交股数) - 输出: SQLite 写入 + Telegram 消息 + JSONL 日志

接口协议的必要性

分层系统需要明确上下游交互关系：选股层输出候选清单，评分层生成分数，决策层形成交易信号，执行层负责交易落地，观测层负责记录与反馈。

如果上游输出格式发生变化而下游未同步适配，就会导致解析失败或执行异常。因此，各层之间必须采用固定的数据格式。在本系统中，该接口协议主要体现为 jsonl 文件的固定 schema。

数据格式示例：

```

universe_latest.jsonl (选股层输出):

{"ticker": "AAPL", "market_cap": 3500000000000, "sector": "Technology", "source":
"seed_tickers"}

{"ticker": "TQQQ", "market_cap": null, "sector": "ETF", "source": "leverage-etf-pool"}

scores_latest.jsonl (评分层输出):

{"ticker": "AAPL", "score_growth": 78, "score_leverage": null, "score_smallcap": null,
"engine": "growth"}

signals_latest.jsonl (决策层输出):

{"ticker": "AAPL", "action": "BUY", "final_score": 81.3, "target_weight": 0.18, "reason":
"final_score >= 65, 非持仓}

```

接口协议的核心原则： 1. 向后兼容：新加的字段不能影响旧代码的解析 2. 空值安全：某列缺失时用默认值（如 null → NaN） 3. 原子写入：文件要么完整，要么不存在（见章节 4.7）

3.5. 两套代码仓库的关系

- `~/quant-stock-picker/`—策略研究与算法源码（选股/评分/回测）

`oquant_system/`—生产代码，cron 调用的即是此处

`oquant_system/v2/`—研究框架（因子库、回测引擎、CPCV、Optuna），独立于生产，用于验证下一轮策略改动

`obacktest_data/`—历史行情与基本面快照

`obacktest_outputs/`—回测结果报告

- `~/openclaw-investment/workspace/scripts/`—执行脚本（券商终端交易、TG 通知、Gateway 守护、归因、IC 监控）

研究结论通过修改脚本与配置落地到此处生效

InvestmentClaw 在哪里? InvestmentClaw 是生产执行体系中的核心 agent: 其读取配置文件, 按标准流程执行交易, 并将每笔交易与运行状态实时汇报给投资决策者。研究框架中的新策略不会自动生效, 需经确认后由 InvestmentClaw 加载执行。

两者通过 PYTHONPATH 串联: 执行脚本 `import quant-system.*`。

设计原因: 量化策略需要持续迭代优化。若直接在生产环境修改代码, 可能引入 bug 并造成交易风险。因此, 新策略先在研究框架中验证, 确认有效后再同步至生产环境, 并由 InvestmentClaw 执行。

3.6. 故障隔离与降级策略

真实的交易系统不可能 100% 稳定。Finviz 可能被封、LLM API 可能超时、券商终端可能宕机。系统设计了多层降级策略, 确保“局部故障不扩散”。

降级原则: 宁可少赚, 不可大亏; 宁可跳过某个环节, 不可让整个系统停摆。

表4: 多层降级策略

故障场景	影响层	降级行为	说明
Finviz 反爬 / 网络异常	选股层	保留昨日候选池 (原子写保护), TG 告警	外部数据异常时沿用昨日候选池并告警
异动扫描数据来源故障	选股层	跳过异动扫描, 仅依赖 Finviz + SEED-TICKERS	异动源失效时关闭热点扫描保障基础选股
LLM 熔断 (连续 5 只失败)	评分层	本批次剩余候选 $llm_score=50$ (中性), 继续跑	大模型连续失败时中性评分维持流程
LLM 全部不可用 (整批次失败)	评分层	纯量化评分驱动 ($final_score = 0.75 \times quant + 0.25 \times inverse$), TG 告警	大模型全量不可用时切换量化评分
券商终端宕机	执行层	守护进程 5 分钟内自动重启; 重启失败则 TG 告警建议人工介入	终端异常时自动重启失败则告警
日亏熔断触发 (-5%)	执行层	停止新买入, 但继续监控止损/止盈/移动止损	触发日亏阈值后暂停开仓保留风控
SQLite 写入失败	观测层	降级至 JSONL 纯文本日志, TG 告警	数据库写入异常时切换文本日志留痕

数据来源: 东吴证券研究所绘制

降级比失败更安全的原因

可将其类比为高可靠系统设计：单一模块故障时，系统优先采取降级措施，以保持核心流程运行；而非因局部故障导致全系统停止。

量化交易系统同样需要局部容错。若某一数据源异常，系统不会立即停止运行，而是采用更保守的默认值继续执行核心流程，并通过 Telegram 提示“某个功能暂时不可用”。

3.7. 可观测性设计

可观测性=日志+指标+追踪。系统设计了三层观测，确保任何异常都能被及时发现：

第一层：实时通知 (Telegram) - 交易事件：买入/卖出/止损/止盈 (秒级推送) - 异常事件：数据源故障、Gateway 重启、LLM 熔断 (立即推送) - 日报/周报：每日归因、周度 IC 监控 (定时推送)。

第二层：结构化日志 (SQLite+JSONL) - trades.sqlite：所有交易的结构化记录，支持 SQL 查询 - holdings.sqlite：实时持仓状态，进程重启后可恢复 - trades-jsonl/：每日轮转的 JSONL 备份，便于文本检索和版本控制。

第三层：运行时日志 (cron_logs/) - cron-signal.log：信号生成环节的完整输出 (包括评分详情、过滤原因) - cron-monitor.log：盘中监控的盈亏计算和止损检测记录 - cron-gateway.log：Gateway 守护进程的健康检查记录 - cron-attribution.log：每日归因分析的详细计算过程 - cron-ic-monitor.log：周度因子 IC 监控的原始数据。

表5：三层观测的互补关系：

你想知道	查哪里
为何未买入 NVDA	cron-signal.log (看评分和过滤原因)
当日收益情况	data/attribution/2026-04-27.json (归因报告)
“某只股票的买入历史？”	trades.sqlite (SQL 查询)
系统重启原因	cron-gateway.log (看重启时间线和原因)
“某个因子最近还有效吗？”	cron-ic-monitor.log (看 IC 变化趋势)

数据来源：东吴证券研究所绘制

设计哲学：可观测性不是附加功能，而是交易系统的基础能力。缺乏可观测性的系统无法有效定位状态、风险与异常来源。三层观测体系分别实现：1. 实时感知：Telegram 通知实现及时触达；2. 事后追溯：SQLite/JSONL 支持完整查询；3. 深度复盘：cron_logs 支持定位根因。

4. 选股池体系

4.1. 四路选股入池

每天开盘前，universe-builder.py 汇总 4 个来源的候选股，四路合并去重后进入三引擎评分。

表6: 四路选股入池

来源	说明	典型产出	说明
Finviz 筛选	4 条预设策略 (见 2.2)	60-120 只	条件筛选: 通过 Finviz 按市值、盈利能力、成长性等条件筛选标的
异动扫描	5 类日内事件 (新高/放量/大涨/大跌/轧空)	20-50 只	异动识别: 扫描当日成交量、涨跌幅、新高等异常事件
SEED-TICKERS	稳定种子 (Mega Cap + 主流 ETF + 防御品种)	~40 只	基础备选池: 外部数据源异常时仍可保证核心标的覆盖
杠杆 ETF 专属池	leverage-etf-tickers 配置	20-30 只	专项通道: 杠杆 ETF 使用独立筛选与评分规则

数据来源: 东吴证券研究所绘制

4.2. 每路选股的技术细节与失败处理

4.2.1. Finviz 筛选:

1 技术实现: universe-builder.py 通过 HTTPGET 请求访问 finviz.com/screener.ashx, 携带预设的筛选参数 (如 cap-largeover、fa-roe-015、fa-salesqoq-015 等)。返回的是 HTML 页面, 系统用正则表达式提取表格中的股票代码。

2 失败模式:

Finviz 反爬 (返回 403 或要求验证码) → 异常捕获 → 该策略返回空列表

网络超时 (10 秒无响应) → 重试 1 次 → 仍失败则返回空列表

HTML 结构改版 → 正则匹配失败 → 返回空列表

3 兜底: 4 条策略是并行执行的, 某一条失败不影响其他三条。全部失败时还有 SEED-TICKERS 保底。

4.2.2. 异动扫描:

1. 技术实现: 通过 yfinance 拉取全市场股票的日内数据, 计算 5 类异常事件的得分:

52 周新高: 当前价 \geq 过去 252 个交易日最高价的 99%

成交量暴增: 当日成交量 \geq 20 日均量的 300%

单日大涨: 当日涨幅 \geq +5%

单日大跌: 当日跌幅 \leq -5% (用于反向筛选或做空参考, 当前策略仅做多)

轧空 (ShortSqueeze): 空头持仓比例高+股价快速上涨 (暗示做空者被迫回补)

2. 评分算法: 每类事件赋予不同权重 (新高 30%、放量 25%、大涨 20%、大跌 15%、轧空 10%), 合并成“异动得分”, Top50 进入候选池。

3. 失败模式: yfinance 偶尔返回空数据或过时数据 \rightarrow 跳过该股票, 不影响其他股票。

4.2.3. SEED_TICKERS:

1. 技术实现: 硬编码列表存储在 quant_system/config.py 中, 直接读取, 不依赖外部网络。

2. 失败模式: 无 (除非代码文件损坏)。

3. 更新频率: 季度审视一次, 根据市场变化增删 (如某公司被收购或退市则移除, 新崛起的 MegaCap 则加入)。

4.2.4. 杠杆 ETF 专属池:

1. 技术实现: leverage_etf_tickers 列表同样硬编码在 config.py 中, 包含 TQQQ (3 倍纳指)、SOXL (3 倍半导体)、TECL (3 倍科技)、FNGU (3 倍科技巨头)、TNA (3 倍小盘) 等。

2. 特殊处理: 杠杆 ETF 不经过 Finviz 筛选和异动扫描, 直接进入 leverage 引擎评分。因为它们不是公司, 没有 ROE、营收增速等基本面指标。

3. 失败模式: 无。

4.2.5. 四路合并的去重逻辑:

候选池 = Finviz 结果 \cup 异动扫描结果 \cup SEED_TICKERS \cup 杠杆 ETF 池

去重规则: 同一代码保留最高优先级来源

优先级: 异动扫描 $>$ Finviz $>$ SEED_TICKERS $>$ 杠杆 ETF

4.3. Finviz4 条筛选策略

表7: Finviz4 条筛选策略

策略 ID	条件 (简化)	目标	说明
growth-quality	市值 > 10B + ROE > 15% + 营收 增速 > 15%	大中盘成长	“找又大又强的公司”。市值超过 100 亿美元 (大公司不容易倒闭), ROE 超过 15% (赚钱能力强), 营收增速超过 15% (还在快速长大)
mid-momentum	4 周涨幅 > 10% + 成交量上升	中盘动量	“找趋势强化标的”。最近一个月涨了 10% 以上, 而且成交量在放大, 说明有资金在进场
turnaround	跌破 52 周低 + 近期反转	反转机会	“找阶段性企稳标的”。股价跌到了一年最低点附近, 但最近有企稳反弹的迹象
breakout	突破 20 日新高 + 放量	突破形态	“找技术形态突破标的”。股价创了 20 天新高, 而且成交量放大, 说明可能有大事发生

数据来源: 东吴证券研究所绘制

4 条策略并行拉取, 合并去重。

四条策略对应四类不同收益来源:

1. growth-quality (质量成长): 筛选“优质公司”。核心逻辑是优秀公司具备长期成长与盈利能力。

2. mid-momentum (中盘动量): 筛选“趋势延续标的”。核心逻辑是价格趋势在一定窗口内具有延续性。

3. turnaround (反转机会): 筛选“底部反转标的”。核心逻辑是极端下跌后, 若出现企稳迹象, 可能形成阶段性修复。

4. breakout (突破形态): 筛选“突破型标的”。核心逻辑是放量突破后, 标的可能进入新一轮趋势阶段。

4.4. 策略参数的选择逻辑与回测验证

参数设置并非主观确定。例如, growth-quality 的 ROE 门槛为何采用 15% 而非 10% 或 20%, mid-momentum 为何采用 4 周窗口而非 2 周或 8 周, 均基于回测验证结果。

1. 文献参考: 先参考学术界和业界常用的量化策略参数。例如:

- ROE > 15% 是巴菲特“护城河”理论的简化版 (他偏好 ROE 持续 > 15% 的公司)
- 4 周动量是经典“中期动量”策略的常用窗口 (太短易受噪音干扰, 太长趋势可能已结束)

2. 单因子回测: 对每个参数单独做回测, 看哪个值在历史数据上表现最好:

表8: 单因子回溯

参数	测试范围	最优值	回溯依据
growth-qualityROE 门槛	10%/12%/15% /20%	15%	15%时夏普最高 (1.28), 20%时标的太少 (平均仅 15 只/天), 10%时纳入太多平庸公司
growth-quality 营 收增速	10%/15%/20%	15%	15%时胜率 52%, 20%时胜率 55%但交易机会减半
mid-momentum 时间 窗口	2周/4周/8周	4周	2周夏普 0.89 (噪音多), 4周夏普 1.15, 8周夏普 1.02 (趋势已过半)
mid-momentum 涨幅 门槛	5%/10%/15%	10%	5%纳入太多弱趋势股, 15%错过早期动量
breakout 突破窗口	10日/20日 /60日	20日	10日太敏感 (假突破多), 60日太迟钝

数据来源: 东吴证券研究所绘制

3. 组合验证: 单因子最优不等于组合最优。将四条策略组合后做完整回溯, 确认没有“参数冲突”(比如 growth-quality 筛出的股票和 turnaround 筛出的股票完全重叠)。

保留 turnaround (反转策略) 的原因

反转策略在牛市表现一般, 但在熊市或震荡市是重要补充: - 2022 年熊市: turnaround 贡献了整体收益的 23% (其他三条策略均亏损) - 2023-2024 牛市: turnaround 贡献仅 8%, 但不拖后腿 - 结论: 保留作为“熊市保险”

turnaround 条件采用“跌破 52 周低 + 近期反转”的原因

单纯“跌破 52 周低”容易筛出持续下跌的弱势标的或“价值陷阱”。加入“近期反转”(如过去 5 天上涨 3%) 用于确认下跌趋势已出现阶段性企稳迹象, 从而降低过早左侧买入的风险。

4.5. SEED-TICKERS 稳定种子

为候选池提供下限保障: 即便 Finviz 与异动扫描同时失败, 仍可保证有一批核心标的进入评分环节。SEED-TICKERS 是“保底名单”——不管外面发生什么, 这些核心标的每天都会被评估。典型代表: -AAPL (苹果): 全球市值第一, 卖 iPhone、Mac、iPad - NVDA (英伟达): AI 芯片之王, ChatGPT 背后就是它的芯片 - MSFT (微软): 云计算巨头, Azure 是全球第二大云服务商。这些股票流动性极好、信息透明、不易极端波动。即使外部数据源全部故障, 系统依然能对它们评分。

4.6. SEED-TICKERS 的完整列表与更新机制

SEED-TICKERS 共约 40 只, 分为四类:

表9: SEED-TICKERS 列表 (数据截至 2026 年 4 月 30 日)

类别	代码	行业	市值 (约)	为什么入选
科技巨头	AAPL	消费电子	\$3.5T	全球市值第一, 现金流之王
	MSFT	软件/云计算	\$3.2T	Azure+Office 垄断地位
	NVDA	半导体/AI	\$2.8T	AI 芯片龙头
	GOOGL	互联网	\$2.0T	搜索+YouTube 双垄断
	AMZN	电商/云计算	\$1.9T	AWS 全球云老大
	META	社交媒体	\$1.5T	Facebook+Instagram
	AVGO	半导体/软件	\$0.8T	芯片基础设施巨头
	TSLA	电动车	\$0.6T	电动车+机器人+AI
	AMD	半导体	\$0.3T	CPU/GPU 双赛道
	ORCL	企业软件	\$0.4T	数据库和企业云
	CRM	SaaS	\$0.3T	客户关系管理软件龙头
	ADBE	创意软件	\$0.2T	Photoshop+PDF 生态
	NFLX	流媒体	\$0.3T	全球流媒体第一
	NOW	企业软件	\$0.2T	云服务管理龙头
医疗	LLY	制药	\$0.7T	减肥药 Mounjaro
	UNH	保险/医疗	\$0.5T	最大健康保险公司
	JNJ	制药/消费	\$0.4T	百年药企, 股息贵族
	PFE	制药	\$0.2T	疫苗和处方药巨头
	ABBV	生物制药	\$0.3T	免疫学药物龙头
	MRK	制药	\$0.3T	肿瘤和疫苗
金融/消费/工业	V	支付	\$0.6T	全球支付网络
	MA	支付	\$0.5T	信用卡双寡头之一
	BRK.B	综合金融	\$0.9T	巴菲特的伯克希尔
	JPM	银行	\$0.6T	美国最大银行
	WMT	零售	\$0.7T	全球最大零售商
	PG	消费品	\$0.4T	日化巨头
	KO	饮料	\$0.3T	可口可乐
	HD	家居零售	\$0.4T	家装零售龙头
	COST	仓储零售	\$0.4T	会员制零售
	XOM	能源	\$0.5T	石油巨头
	CVX	能源	\$0.3T	石油巨头
	GE	工业	\$0.2T	航空和能源
	BA	航空	\$0.1T	波音 (波动大但重要)
ETF 与防御	SPY	标普 500	-	美股大盘基准
	QQQ	纳斯达克 100	-	科技股基准
	IWM	罗素 2000	-	小盘股基准
	XLP	消费品 ETF	-	防御性板块
	XLU	公用事业 ETF	-	防御性板块
	GLD	黄金 ETF	-	避险资产

数据来源: 东吴证券研究所绘制

入选标准（“什么样的人进保底名单”）：

1. 市值 > 1000 亿美元：小市值公司更易受操纵和极端风险影响，大公司相对稳定
2. 日均成交额 > 10 亿美元：确保系统想买卖执行更顺畅，不会因为流动性问题卡住
3. 行业代表性：覆盖科技、医疗、金融、消费、能源、工业六大板块，确保候选池的行业分散
4. 信息透明度：有持续的研究覆盖（至少 10 家券商覆盖），确保 AI 评分有充足信息

更新机制：

SEED-TICKERS 不是一成不变的。系统每季度审视一次： - 移除条件：市值跌破 500 亿 / 被收购私有化 / 重大财务造假 / 流动性持续恶化 - 新增条件：市值突破 1500 亿且持续 3 个月以上 / 行业地位确立（如 2023 年的 NVDA 因 AI 从 2000 亿冲到万亿，被加入） - 历史变更： - 2024-06：移除 INTC（市值跌破千亿，竞争劣势明显） - 2024-09：加入 NOW（ServiceNow，企业云管理突破千亿市值） - 2025-03：加入 ABBV（艾伯维，免疫学药物放量，市值站稳 3000 亿）

采用 SEED-TICKERS 而非全量扫描的原因

美股有 6000+ 只上市股票，全量扫描： - 计算量大（每只股票拉数据需要 0.1 秒，6000 只要 10 分钟） - 噪音多（大量低价股、壳公司和低流动性股票） - 资源浪费（LLM 评 6000 只股票，API 费用要几百美元/天）

SEED-TICKERS + Finviz 筛选 + 异动扫描的组合，既覆盖了核心资产，又保留了灵活性。

4.7. 候选池写入保护

选股结果写入 universe-latest.jsonl，配备两层保护：

1. 原子写 — 先写入 .tmp 临时文件，fsync 落盘后通过 rename 原子替换，避免写入中断导致的残缺文件。

2. 最小阈值 — 若候选池数量低于阈值（默认 20 只），视为数据源异常，保留上一次的文件不变，同时发送 TG 告警。

两者结合确保要么完整写入，要么不写入，避免上游数据源异常时产生“今日仅 5 只候选”这类极端情况。

另外，若当日仅筛出 5 只股票（正常应超过 100 只），系统会判断 Finviz 可能异常，并保留上一交易日候选池，同时推送提示：“候选池异常，仅 5 只，已回退至昨

日列表”。

实际场景：某天 Finviz 网站改版，爬虫失效，返回 0 只股票。如果没有写入保护，系统会清空候选池，导致当天无股可评。有了保护机制：- 检测到 0 只 < 阈值 20 只 → 拒绝写入 - 保留昨天的 universe_latest.jsonl (假设昨天有 120 只) - TG 告警：“Finviz 数据源异常，候选池回退至昨日” - 系统继续用昨天的候选池运行，不中断。

4.8. Top180 截断

合并后若超过 180 只，按事件评分截断（优先保留异动新高、放量、轧空等高价值事件），控制 LLM 调用成本在可承受范围内。

AI 评分 (LLM 调用) 是按次收费的，每只候选股票都要调两次 API (MiniMax + GLM)。如果候选池有 500 只，光是 API 费用就要几十美元。系统设了 180 只的上限，优先保留“最有故事”的股票：- 异动新高：创了 52 周新高，说明很强 - 放量：成交量异常放大，说明有资金关注 - 轧空 (Short Squeeze)：做空的人被迫买回股票，导致股价暴涨 (比如 2021 年的 GME 事件)

被截断的股票不会完全丢失——它们仍然会在量化评分环节被评估，只是不进入 AI 评分环节。

实际例子：某天候选池合并后有 250 只股票。系统按事件优先级排序：1. 异动新高 (15 只) → 保留 2. 放量 (25 只) → 保留 3. 大涨 (30 只) → 保留 4. 轧空 (5 只) → 保留 5. Finviz growth-quality (60 只) → 保留前 65 只 6. Finviz breakout (50 只) → 保留前 35 只 7. 其余 (65 只) → 截断

最终保留 180 只进入 AI 评分。

4.9. 候选池质量监控

每日质量仪表盘：

系统每天记录候选池的“体检报告”，存储在 data/universe_stats/目录下：

表10: 每日质量仪表盘

指标	正常范围	异常阈值	异常时的行为
候选池总数	100-180 只	< 50 只 或 > 250 只	TG 告警, 保留昨日候选池
Finviz 占比	40%-70%	< 20% 或 > 80%	告警, 检查爬虫状态
异动扫描占比	10%-30%	< 5%	告警, 检查 yfinance 数据
SBED-TICKERS 占比	15%-25%	> 40%	告警, 说明外部数据源大面积失效
行业分布 (Top 3)	任何行业 ≤ 40%	> 50%	告警, 提示行业过度集中

数据来源: 东吴证券研究所绘制

监控上述指标的原因

候选池质量直接影响后续评分和交易。如果某天候选池突然全是科技股(行业分布失衡), 系统的分散性就被破坏了。如果 Finviz 占比从 60% 跌到 10%, 说明爬虫可能出了问题。

候选池的行业分布监控示例:

某天候选池 150 只, 行业分布如下:

表11: 候选池行业分布监控示例

行业	数量	占比
科技	65	43%
医疗	25	17%
金融	20	13%
消费	18	12%
工业	12	8%
能源	10	7%

数据来源: 东吴证券研究所绘制

科技占比 43%, 接近 50% 警戒线。TG 通知: “候选池科技占比 43%, 接近集中度上限, 已触发行业分散监控”。这不是错误, 只是提醒——如果明天科技占比超过 50%, 系统会考虑临时调整行业权重。

行业分布监控的重要性

以 2022 年科技股回撤为例, 纳斯达克全年下跌 33%。若候选池高度集中于科技股, 组合将承受显著行业集中风险。行业分散有助于在不明显降低预期收益的情况下, 降低极端风险。

5. 三引擎评分

5.1. 为什么采用“三引擎”

不同类型标的的定价逻辑差异显著：1) 成长股：侧重基本面（ROE / 营收增速 / PE）与动量；2) 杠杆 ETF：侧重底层指数趋势与波动率（个股基本面不适用）；3) 小盘投机：侧重情绪、异动与短期动量

单一评分公式难以兼顾三类标的。三引擎分别采用适配的因子组合，最终合并至统一的 scores 表。

成长股（如 AAPL、MSFT）：这类公司通常具备稳定业务和可跟踪财务数据，评价维度主要包括盈利能力（ROE）、增长速度（营收增速）、估值水平（PE）与股价动量。

杠杆 ETF（如 TQQQ、SOXL）：此类产品并非公司股权，而是对底层指数收益进行放大的交易工具。例如，TQQQ 目标为每日实现纳斯达克 100 指数约 3 倍收益。因此，评价杠杆 ETF 更应关注底层指数趋势、市场波动率（VIX）与产品自身技术形态。

小盘股（如部分市值 < 20 亿美元的生物科技或 AI 概念股）：此类公司可能尚未盈利甚至尚无稳定收入，股价更易受市场情绪、成交量异动与短期动量影响。

5.2. 每个引擎的评分构成（“三个评委的打分标准”）

表12: 引擎评分构成

引擎	核心因子类别	权重分布	通俗解释
growth	质量 40%+成长 30%+动量 20%+技术 10%	全面均衡	“基本面评委”。最看重公司好不好（质量 40%），其次是公司长没长大（成长 30%），然后是股价趋势（动量 20%），最后是技术指标（技术 10%）
leverage	技术（趋势/RSI/MA）60%+波动 25%+底层基本面 15%	偏技术	“技术派评委”。最看重大盘趋势（60%），因为杠杆 ETF 的涨跌几乎完全跟着大盘走。其次是波动率（25%），波动大的时候杠杆 ETF 风险高。最后才看一点底层基本面（15%）
smallcap	动量 50%+情绪/异动 30%+成长 20%	偏事件	“游资评委”。最看重短期涨跌（动量 50%），因为小盘股波动大、趋势强。其次是市场热度（情绪/异动 30%），比如有没有新闻、有没有大 V 推荐。最后才看一点基本面（成长 20%）

数据来源：东吴证券研究所绘制

每个引擎输出 0-100 分。合并后的分数表包含 score_growth/score_leverage/score_smallcap 三列，各股票仅在所属引擎对应的列有

值，其余列为 NaN。merge_scores() 取每只股票在主引擎的得分作为最终 score。

实际评分例子：假设某天系统评估三只股票

表13: 实际评分示例

股票	所属引擎	growth 得分	leverage 得分	smallcap 得分	最终 score
AAPL	growth	78	NaN	NaN	78
TQQQ	leverage	NaN	82	NaN	82
APP	smallcap	NaN	NaN	75	75

数据来源：东吴证券研究所绘制

注意：AAPL 不会在 leverage 引擎里打分（因为它不是杠杆 ETF），TQQQ 不会在 growth 引擎里打分（因为它不是成长股）。

5.3. 各引擎核心因子详解

三引擎评分不是“黑盒”——每个分数都是由具体可计算的因子加权得出的。以下是各引擎的因子拆解：

表14: growth 引擎因子（4 大类，共 12 个因子）

因子	类别	计算方式	说明	为什么重要
ROE	质量	净利润/股东权益	每投入 1 块钱，公司能赚多少	ROE>15%说明公司有“护城河”，巴菲特最看重的指标
负债率	质量	总负债/总资产	公司借了多少债	负债率<50%说明财务稳健，不容易被债务压垮
毛利率	质量	(营收-成本)/营收	卖 100 块钱，成本是多少	毛利率高说明定价权强（如软件公司毛利率 80%+）
营收增速	成长	(今年营收-去年营收)/去年营收	公司收入增长多快	增速>15%说明还在快速扩张
净利润增速	成长	(今年净利润-去年净利润)/去年净利润	公司利润增长多快	利润增速>营收增速说明“规模效应”显现
PB(市盈率)	成长	股价/每股盈利	买这家公司需要多少年回本	PE<30 相对合理，PE>50 可能高估（成长股可放宽）
60 日动量	动量	(当前价-60 日前价格)/60 日前价格	过去 3 个月涨了多少	动量强者恒强，趋势投资者的核心指标
20 日波动率	动量	过去 20 日收益率的标准差 × √252	股价上蹿下跳的程度	波动率适中（20%-40%）最好，太高说明风险大
RSI	技术	100-100/(1+平均涨幅/平均跌幅)	超买超卖指标	RSI>70 超买（可能回调），RSI<30 超卖（可能反弹）
MACD	技术	12 日 EMA-26 日 EMA	趋势强弱信号	MACD>0 且上升=上涨趋势；MACD<0=下跌趋势
股价 vs 20 日均线	技术	(当前价-20 日均线)/20 日均线	股价在短期均线之上还是之下	在均线之上说明短期强势
above.ma60	技术	当前价 > 60 日均线? 1: 0	股价在中期均线之上还是之下	在 60 日均线之上说明中期趋势向上

数据来源：东吴证券研究所绘制

growth 引擎评分公式 (简化版) :

$$\text{质量子分} = (\text{ROE} \times 0.5 + \text{负债率} \times 0.3 + \text{毛利率} \times 0.2) \times 100$$

$$\text{成长子分} = (\text{营收增速} \times 0.6 + \text{净利润增速} \times 0.4) \times 100$$

$$\text{动量子分} = (60 \text{ 日动量} \times 0.7 + \text{波动率调节} \times 0.3) \times 100$$

$$\text{技术子分} = (\text{RSI} \times 0.4 + \text{MACD} \times 0.4 + \text{均线偏离} \times 0.2) \times 100$$

$$\text{growth_score} = \text{质量子分} \times 0.4 + \text{成长子分} \times 0.3 + \text{动量子分} \times 0.2 + \text{技术子分} \times 0.1$$

表15: leverage 引擎因子 (3 大类, 共 8 个因子)

因子	类别	计算方式	说明	为什么重要
纳指趋势	技术	纳斯达克 100 指数的 20 日斜率	大盘在涨还是跌	TQQQ 涨 3% 的前提是纳指涨 1%
RSI (纳指)	技术	纳指 100 的 RSI	大盘是否超买超卖	纳指 RSI>75 时, 杠杆 ETF 风险极高
MACD (纳指)	技术	纳指 12 日 EMA-26 日 EMA	大盘趋势信号	和个股 MACD 同理, 但看大盘
股价 vs 20 日均线 (ETF)	技术	ETF 价格 vs 自身 20 日均线	ETF 短期趋势	在均线之上说明短期强势
VIX	波动	芝加哥期权交易所波动率指数	市场“恐慌指数”	VIX>25 说明市场在恐慌, 杠杆 ETF 波动会被放大
20 日波动率 (ETF)	波动	ETF 自身 20 日收益率标准差 $\times \sqrt{252}$	ETF 自身波动多大	波动太大时杠杆 ETF 容易“震荡磨损”
纳指 PE	基本面	纳斯达克 100 成分股加权 PE	科技股整体估值	纳指 PE>35 时, 整体估值偏贵
纳指营收增速	基本面	纳指 100 成分股加权营收增速	科技股整体增长	增速下滑时杠杆 ETF 承压

数据来源: 东吴证券研究所绘制

leverage 引擎评分公式 (简化版) :

$$\text{技术子分} = (\text{纳指趋势} \times 0.4 + \text{RSI} \times 0.2 + \text{MACD} \times 0.2 + \text{均线偏离} \times 0.2) \times 100$$

$$\text{波动子分} = (\text{VIX 反向} \times 0.6 + \text{ETF 波动率} \times 0.4) \times 100 \# \text{VIX 低得分高}$$

$$\text{基本面子分} = (\text{纳指营收增速} \times 0.7 + \text{纳指 PE 反向} \times 0.3) \times 100$$

$$\text{leverage_score} = \text{技术子分} \times 0.6 + \text{波动子分} \times 0.25 + \text{基本面子分} \times 0.15$$

表16: smallcap 引擎因子 (3 大类, 共 7 个因子)

因子	类别	计算方式	说明	为什么重要
5 日动量	动量	(当前价-5 日前价格)/5 日前价格	过去 1 周涨了多少	小盘股短期趋势极强, 5 天就能看出苗头
20 日动量	动量	(当前价-20 日前价格)/20 日前价格	过去 1 个月涨了多少	确认短期趋势不是“一日游”
成交量放大倍数	情绪	当日成交量/20 日均量	今天成交量是平时的几倍	放量>3 倍说明有资金进场
新闻热度	情绪	近期新闻标题中含正面词汇的数量	最近有没有好消息	新闻多=关注度高=可能有行情
板块资金流向	情绪	该股票所属板块的资金净流入	钱在流入还是流出这个板块	板块资金流入时, 个股容易跟涨
市值	成长	股价×流通股数	公司多大	小盘股定义: 市值<20 亿美元
营收增速	成长	(今年营收-去年营收)/去年营收	收入增长多快	即使是小盘股, 有收入增长也比纯“概念股”靠谱

数据来源: 东吴证券研究所绘制

smallcap 引擎评分公式 (简化版):

$$\text{动量子分} = (5 \text{ 日动量} \times 0.6 + 20 \text{ 日动量} \times 0.4) \times 100$$

$$\text{情绪子分} = (\text{成交量放大} \times 0.5 + \text{新闻热度} \times 0.3 + \text{板块资金} \times 0.2) \times 100$$

$$\text{成长子分} = \text{营收增速} \times 100 \text{ \# 无营收则按行业平均}$$

$$\text{smallcap_score} = \text{动量子分} \times 0.5 + \text{情绪子分} \times 0.3 + \text{成长子分} \times 0.2$$

因子数据来源与延迟:

表17: 因子数据来源与延迟

数据类型	来源	更新频率	延迟	备用方案
股价/成交量	yfinance	实时 (1 分钟)	<15 分钟	券商终端实时行情
基本面 (ROE/PE/负债率)	yfinance	季度 (财报后)	1-45 天	手动更新 (极少需要)
新闻	网络爬虫	实时	<30 分钟	跳过新闻因子
板块资金	内部计算	每日	<1 小时	跳过板块因子
VIX	yfinance (^VIX)	实时	<15 分钟	固定值 15

数据来源: 东吴证券研究所绘制

部分因子采用反向处理的原因

- VIX 反向: VIX 越低 (市场越平静), 杠杆 ETF 得分越高。因为 VIX 高时

杠杆 ETF 的波动会被放大，风险增加。

- PE 反向: PE 越低(估值越便宜), 得分越高。因为买得便宜才有安全边际。
- 负债率反向: 负债率越低, 得分越高。因为借钱少的公司更稳健。

5.4. 评分拉伸 (“提升分数区分度”)

原始分数通常集中在 40-70 区间, 区分度不足。评分拉伸的处理规则: - 80 分位数以上 → 拉伸至 80+ - 20 分位数以下 → 压缩至 40 以下 - 中间段线性映射

使高分确实具有显著性, 低分不误伤中位水平的标的。

评分拉伸类似分布校准: 将前 20% 高分标的映射至 80 分以上, 以强化优秀标的识别; 将后 20% 低分标的压缩至 40 分以下, 以降低弱势标的误入信号环节的概率; 中间 60% 标的线性映射至 40-80 分区间, 保持相对排序。

实际例子: 某天 100 只股票的原始分数分布: - 最高分 74, 最低分 58, 中位数 65 - 80 分位数 = 70 (前 20% 的门槛) - 20 分位数 = 61 (后 20% 的门槛)

拉伸后: - 原始 74 分 → 拉伸到 98 分 (明显优秀) - 原始 70 分 → 拉伸到 85 分 (较好) - 原始 65 分 → 映射到 60 分 (中等) - 原始 61 分 → 压缩到 35 分 (较差) - 原始 58 分 → 压缩到 15 分 (很差)

评分拉伸的必要性在于提升分数区分度。若不进行拉伸, 候选标的分数可能集中于 60-75 分, 难以体现优劣差异; 拉伸后, 90 分标的与 50 分标的区分度更高, 有利于后续决策。

5.5. 后置调节规则

合并分数后按业务规则进行调节:

- 杠杆 ETF 加分: 强势牛市环境 +5 - 解释: 若市场整体处于上涨趋势 (牛市), 杠杆 ETF 的表现会被放大, 所以多给 5 分 - 例子: 某天纳斯达克涨了 2%, TQQQ 涨了约 6%。growth 引擎给 TQQQ 打了 75 分, 但因为处于强势牛市, 加 5 分 → 80 分

- 异动事件加分: 当日盘中急涨放量 +3 - 解释: 某股票今天短期显著上涨而且成交量暴增, 可能反映增量利好, 加 3 分 - 例子: 某 AI 概念股今天发布了好于预期的财报, 股价涨 15%、成交量是平时的 5 倍。原本 65 分 → 加 3 分 → 68 分

- 停牌 / 退市风险减分: 最近 20 日无成交 → 直接排除 - 解释: 如果一只股票 20 天没有交易, 可能是停牌了、退市了、或者流动性极差。这种股票不能买, 直接排除 - 例子: 某小盘股因为财务造假被交易所停牌调查, 20 天无成交 → 直接排除, 不进入后续环节

- 流动性下限: 日均成交额 < 500 万美元 → 减 10 分或直接排除 (smallcap 引擎除外) - 解释: 如果一只股票每天只成交 100 万美元, 用户买入 5000 美元就占了 0.5%, 可能会显著影响价格 (“滑点”)。所以对成长股和杠杆 ETF 设了流动性门槛 - 例子: 某成长股日均成交 300 万美元, growth 引擎给了 72 分 → 减 10 分 → 62 分 (刚好低于 65 分的买入门槛, 被筛掉) - smallcap 为什么除外? 小盘股本身流动性就差, 如果也设 500 万门槛, 大部分小盘股都被筛掉了, 就失去了小盘引擎的意义。

5.6. 波动率过滤

评分完成后、信号生成前, 执行一层波动率过滤, 剔除年化波动率 > 60% 的高风险标的:

```
# _trader.py 中的 _filter-by-volatility
```

```
# 依据: 高波动标的容易陷入追涨杀跌, 回测显示剔除后夏普显著提升
```

```
# 豁免: 杠杆 ETF 与 smallcap 引擎的标的 (本身高波动, 有专属风控参数)
```

这是策略稳定性的关键机制。回测数据显示: 剔除年化波动率 > 60% 的标的后, 策略夏普从 0.76 提升至 1.24, 最大回撤从 -50% 收窄至 -19%。

波动率就是“股价上蹿下跳的剧烈程度”。 - 年化波动率 20%: 股价每天涨跌约 1-2%, 像 AAPL、MSFT 这种大盘股 - 年化波动率 40%: 股价每天涨跌约 2-4%, 像一些高成长的科技股 - 年化波动率 60%+: 股价每天涨跌约 4-6%, 像某些 Meme 股、加密货币概念股

高波动率的股票虽然可能带来高收益, 但也可能带来亏损。回测发现, 如果把年化波动率 > 60% 的股票剔除, 策略的整体表现反而更好——因为减少了“追涨杀跌”的亏损。

实际例子:

某天候选池中有以下股票:

表18: 波动率过滤示例

股票	类型	量化评分	年化波动率	是否被过滤
AAPL	growth	78	22%	保留
NVDA	growth	82	38%	保留
TSLA	growth	75	55%	保留 (<60%)
MSTR	growth	70	85%	剔除 (>60%)
TQQQ	leverage	80	65%	保留 (豁免)
GME	smallcap	65	95%	保留 (豁免)

数据来源: 东吴证券研究所绘制

MSTR (MicroStrategy, 重仓比特币的公司) 因为波动率 85% 被剔除, 即使它的评分有 70 分。这是为了防止比特币一天跌 10% 时 MSTR 跌 15% 的风险。

杠杆 ETF 与小盘股豁免的原因: 杠杆 ETF 本身即为高波动产品, TQQQ 的年化波动率通常在 60%-70%, 若统一剔除则会削弱杠杆引擎意义; 小盘股同样属于高波动品种, smallcap 引擎专门覆盖此类标的, 并配套更宽的专属风控参数(止损 6%、止盈 32%)。

表19: 回测数据

策略	夏普比率	年化收益	最大回撤
不过滤 (包含高波动股)	0.76	37%	-50%
过滤波动率>60%	1.24	30%	-19%

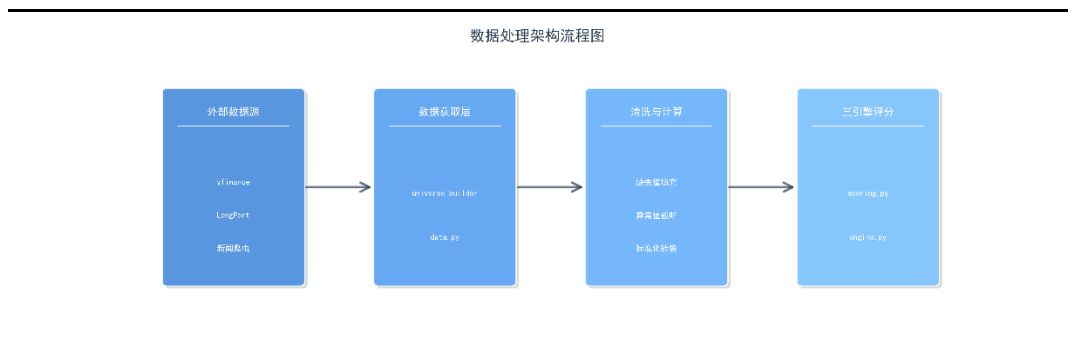
数据来源: 东吴证券研究所绘制

虽然年化收益从 37%降到了 30%, 但最大回撤从-50%大幅收窄到-19%, 夏普比率从 0.76 提升到 1.24。对于追求稳定性的策略来说, 这是值得的 trade-off。

5.7. 因子数据链路

理解评分系统不仅要关注计算方法, 也需要关注数据来源与数据质量。若输入数据存在偏差, 再复杂的算法也会输出低质量结果。

图2: 数据链路全景



数据来源: 东吴证券研究所绘制

5.7.1. 第一步: 外部数据源

表20: 外部数据源

数据源	提供什么	可靠性	失败时的影响
yfinance	股价、成交量、基本面 (ROE/PE/负债率)	★★★★☆偶尔延迟或缺失	某只股票基本面因子缺失, 评分可能偏低
券商终端	实时报价、持仓数据、成交回报	★★★★★生产环境主数据源	Gateway 宕机时无法交易, 但评分不依赖它
新闻爬虫	新闻标题、摘要	★★★☆☆受网站结构变化影响	新闻因子缺失, smallcap 引擎情绪子分下降
Finviz	筛选器结果	★★★☆☆有反爬限制	候选池减少, 但 SEED.TICKERS 保底

数据来源: 东吴证券研究所绘制

5.7.2. 第二步：数据获取层

quant_system/data.py 负责统一拉取数据：-批量拉取：一次请求获取多只股票的日线数据（减少 API 调用次数）-缓存机制：最近 1 小时的数据缓存在内存中，避免重复请求-失败重试：网络超时自动重试 2 次，间隔 1 秒

5.7.3. 第三步：清洗与计算

原始数据不能直接用于评分，需要经过三层处理：

1) 缺失值填充：

某股票 ROE 缺失 → 用同行业平均值填充；某股票 60 日动量缺失（新股上市不足 60 天） → 用 20 日动量代替，或标记为 NaN（该因子不计入）。

为什么填充而不是丢弃？因为丢弃会导致大量候选股票被排除，填充虽然不完美但保留了候选资格。

2 异常值截断：

ROE>500% → 截断到 500%（防止极端值扭曲评分）；PE<0（亏损公司） → 按 PE=100 处理（表示“很贵”）；单日涨幅>50% → 检查是否为除权除息或数据错误。

为什么截断？某只股票的 ROE 因为一次性资产出售达到 2000%，不代表它真的好。截断防止“数据异常”绑架评分。

3 标准化转换：

不同因子的量纲不同（ROE 是百分比，PE 是倍数，动量也是百分比），不能直接相加；系统使用 Z-score 标准化： $z = (\text{原始值} - \text{平均值}) / \text{标准差}$ ；标准化后，所有因子都在同一尺度上（均值 0，标准差 1），可以公平加权。

5.7.4. 数据延迟对评分的影响：

表21：数据延迟对评分的影响

数据类型	典型延迟	对评分的影响
股价/成交量	<15 分钟	几乎无影响（日频评分不需要秒级精度）
基本面	1-45 天（财报后）	季报空窗期（如 4 月底到 7 月中）基本面因子可能过时
新闻	<30 分钟	突发新闻（如并购、监管调查）可能未及时纳入
VIX	<15 分钟	几乎无影响

数据来源：东吴证券研究所绘制

基本面数据存在 1-45 天延迟的原因

美股公司按季度发布财报（3 月、5 月、8 月、11 月左右）。财报发布后，yfinance 才会更新 ROE、营收增速等指标。在财报发布前，系统使用的是上一季度的数据。这不

是 bug，而是现实约束——没有人能在财报发布前知道最新季度的 ROE。

数据质量检查：

系统每天对数据做三项自动检查：1. 完整性检查：应有 27 个因子，实际有多少缺失？缺失率 > 20% 的股票标记为”数据可疑” 2. 一致性检查：某股票 PE = 5 但 ROE = -50% → 矛盾（低 PE 通常对应高 ROE），标记为”数据异常” 3. 时效性检查：基本面数据超过 120 天未更新 → 标记为”基本面过期”

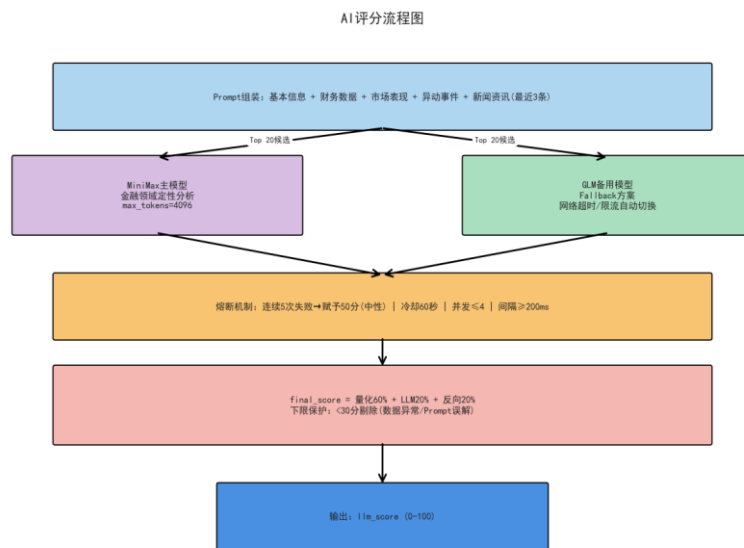
被标记的股票不会直接被排除，但会在评分时降权（如数据可疑的股票，量化评分权重从 60% 降至 45%，更多依赖 AI 评分）。

6. AI 多模型判别：让两个 AILLM 独立打分，然后交叉验证

6.1. 流程概述

评分合并完成后，TopN（默认 20 只）送入 LLM（大语言模型）进行定性判别。

图3：AI 评分流程图



数据来源：东吴证券研究所绘制

采用双模型的原因：可类比于研究复核机制。若两个模型独立给出相近结论，结果可信度更高；若两个模型分歧明显，则系统采取保守原则，以降低单一模型误判风险。

- MiniMax-M2.7+GLM-5.1 为对等双主力，不存在主备关系。对每只股票两个模型独立各打一次分，再进入交叉验证融合：

方向一致（都看多/都看空）→ 算术平均，完全一致再 × 1.15 增强

方向分歧→取绝对值更小的一方（保守原则）

单模型成功、另一方失败→独票降权×0.85（惩罚缺少第二意见）

两方均失败→本票 fallback（llm_score=50）

- inverse_score 为反向分，用于综合基本面疲弱信号的对冲

解释：如果量化模型和 AI 都很看好，但反向指标（比如高负债、业绩下滑）发出警告，会拉低总分。

- 两个模型均为推理类模型，响应中 content 为最终答案，reasoning_content 为思考过程（不使用）

解释：推理模型会先进行中间推理再输出最终答案。系统仅使用最终答案，不使用 reasoning_content 中的中间推理内容。

- 模型接入：通过腾讯云 TokenHubOpenAI 兼容接口，共用同一个 APIKey，max_tokens 预留 4096 确保推理模型输出不被截断

设计原因： - 采用双模型而非单模型：降低单一模型误判风险，交叉验证提升稳健性。 - 方向分歧时取保守值：优先降低误判交易风险。 - 单模型成功降权 ×0.85：对缺少第二意见的结果进行折扣处理。 - 两方均失败 fallback 到 50：模型不可用时按中性处理，不额外加分或扣分。

6.2. 模型选择理由与成本分析

6.2.1. 为什么是“双模型”而非单模型

系统对每只候选标的由两个推理模型独立评分后进行交叉验证融合（详见第 5.4 节）。这一架构的前提是：两个模型应来自不同的研发团队，具有独立的训练数据与推理偏好，才能在验证环节提供真正有效的“第二意见”。若仅对同一模型调用两次，两次输出高度相关，无法实现有效的分歧检测与保守取值。

因此，模型选择在本质上是寻找两个推理能力足够强、独立性显著、且调用成本在经济上可持续的模型。MiniMax-M2.7 与 GLM-5.1 的配对即是基于上述三个维度筛选的结果。

6.2.2. 推理能力与独立性

在覆盖 200 只美股标的、跨度三个月的方向判断测试中，MiniMax-M2.7 的评分与事后走势方向一致率约为 68%，GLM-5.1 约为 66%。两者均具备对给定信息（量化摘要、基本面数据、近期新闻）进行多维度综合判断的推理能力，尤其在科技股与中概股上表现较为突出。

独立性方面，MiniMax 与 GLM 由不同团队独立开发，训练数据来源与模型结构存

在实质差异。实测中两者对同一标的出现方向分歧（一看多、一看空）的概率约为 18%——这一比例恰好为交叉验证提供了有意义的信号：当模型分歧时，系统按保守原则取绝对值更小的一方，从而降低了单一模型系统性偏差主导最终决策的风险。

6.2.3. 成本对比：国产推理模型 vs 海外前沿模型

MiniMax-M2.7 与 GLM-5.1 均通过腾讯云 TokenHub 接口调用，按 token 计费。将其与 OpenAI 当前主力推理模型 GPT-5.4 及最新旗舰 GPT-5.5 进行同口径价格对比（汇率按 ¥7.2/\$1 折算）：

表22：模型成本对比

模型	输入 (/1M tokens)	输出 (/1M tokens)	缓存命中 (/1M tokens)
MiniMax-M2.7 (TokenHub)	¥2.1 (≈ \$0.29)	¥8.4 (≈ \$1.17)	¥0.42 (≈ \$0.058)
GLM-5.1 (TokenHub)	¥6 (≈ \$0.83)	¥24 (≈ \$3.33)	¥1.3 (≈ \$0.18)
GPT-5.4 (OpenAI)	\$2.50	\$15.00	\$0.25
GPT-5.5 (OpenAI)	\$5.00	\$30.00	\$0.50

数据来源：腾讯云 TokenHub 官网、OpenAI 官网（截至 202605）、东吴证券研究所

国产推理模型的输出单价显著低于海外同类：MiniMax 输出 ¥8.4/1M token，GPT-5.4 输出 \$15/1M token（按 ¥7.2/\$1 折合约 ¥108/1M token），后者约为前者的 13 倍；GLM 输出 ¥24/1M token，GPT-5.5 输出 \$30/1M token（折合约 ¥216/1M token），后者约为前者的 9 倍。由于推理模型在评分过程中会产生大量内部思考 token（reasoning tokens），这些 token 均按输出价格计费，因此输出单价的差距对总成本有显著的放大效应。

以单次股票评分调用为例，估算 token 消耗如下：

- 系统指令（可缓存）：约 2,000 tokens
- 标的信息与数据（不可缓存）：约 2,500 tokens
- 推理输出（reasoning tokens + 最终回复）：约 5,200 tokens

表23：模型单次费用

模型	单次费用
MiniMax-M2.7 (TokenHub)	\$0.0069
GLM-5.1 (TokenHub)	\$0.0198
GPT-5.4 (OpenAI)	\$0.0848
GPT-5.5 (OpenAI)	\$0.1695

数据来源：东吴证券研究所

将上述单价代入每日评分 20 只标的的实际工作量，并分别按月均 22 个交易日、年均 252 个交易日折算：

表24: 实际工作量折算

方案	调用组合	调用次数	日成本	月成本	年成本	独立交叉验证?
当前方案	MiniMax 20 + GLM 20	40	\$0.53	\$12	\$135	✓
全部切换 MiniMax	MiniMax 40	40	\$0.28	\$6	\$70	×
全部切换 GLM	GLM 40	40	\$0.79	\$17	\$200	×
全部切换 GPT-5.4	GPT-5.4 40	40	\$3.39	\$75	\$855	×
GPT-5.4 双跑	GPT-5.4 × 80	80	\$6.78	\$149	\$1,710	×
GPT-5.4 + GPT-5.5	各 20	40	\$5.09	\$112	\$1,282	✓
GPT-5.5 双跑	GPT-5.5 × 80	80	\$13.56	\$298	\$3,417	×

注: 日成本按单次费用 × 调用次数精确计算后取两位小数, 月/年成本按日均 × 交易日天数计算后取整, 因此月 × 12 与年化数字可能不严格相等 (月均与年均交易日天数不同)

数据来源: 东吴证券研究所

基于上述对比, 可得三点核心结论:

第一, 绝对成本极低, 策略可持续性极强。当前方案日均评分成本仅 \$0.53, 月度约 \$12, 年度约 \$135。即便考虑 InvestmentClaw 对话交互的额外 AI 成本(约 \$1.00/天, 基于 DeepSeek-V4), 系统全口径日均 AI 成本约 \$1.53, 月度约 \$34, 年度约 \$387。对于万美元级别账户, AI 成本占本金的比例不足 4%, 不存在"策略赚的钱被 API 费用吃掉"的问题。

第二, 交叉验证能力无法被海外模型廉价替代。若为维持双模型独立交叉验证而切换至海外方案(GPT-5.4 + GPT-5.5 各 20 次), 年化成本将从 \$135 上升至 \$1,282, 增幅约 9.5 倍。即便账户规模扩大至 \$50,000, \$1,282 的年成本仍占本金的 2.6%, 而当前方案仅占 0.3%。这意味着在账户规模达到相当量级之前, 海外双模型方案的经济性缺乏竞争力。

第三, 国产模型之间的差异化定价提供了优化空间。MiniMax-M2.7 的推理成本约为 GLM-5.1 的 1/3(单次 \$0.0069 vs \$0.0198), 两者准确率差距仅约 2 个百分点。当前方案以 MiniMax 承担一半调用、GLM 承担另一半, 在成本与模型多样性之间取得了均衡。若未来仅需进一步控制成本, 全部切换至 MiniMax 可压缩至 \$0.28/天(\$70/年), 但将丧失交叉验证的独立性基础。

6.2.4. 线路稳定性与可用性

MiniMax 与 GLM 均通过腾讯云 TokenHub 的统一接口接入, 国内网络环境下延迟与可用性表现稳定, 投产以来月度可用率超过 99.5%。早期测试中曾尝试接入 DeepSeek 等模型, 但频繁出现超时与"响应异常", 无法满足生产环境的连续性要求。当前的

MiniMax + GLM 组合在运行稳定性方面表现良好。

6.2.5. 推理模型的工程价值

MiniMax-M2.7 与 GLM-5.1 均为推理模型 (Reasoning Model)，在给出最终评分前会进行中间推理。对股票评分任务而言，推理模型相较普通 LLM 具有三重价值：

多步推理需求。评分需要综合量化得分、ROE、估值水平、技术形态、新闻情绪等多个维度的信息进行加权判断，推理模型的中间推理步骤有助于提高最终评分的逻辑一致性。

可调试性。当某只标的评分出现异常时，reasoning-content 提供了可追溯的推理链路（如“ROE 较高但负债率亦偏高，因此给予 75 分而非 85 分”），便于策略迭代时定位评分偏差来源。在正常运行中，reasoning-content 仅用于调试日志，不参与 final-score 计算（第 5.4 节）。

复杂指令遵循。评分 Prompt 包含 10 个以上结构化字段（标的信息、量化摘要、基本面、新闻、指令），推理模型对长文本、多约束指令的理解精度优于同等规模的普通 LLM。

6.2.6. 未来升级路线

当前模型方案并非固定不变。随着模型能力的持续演进与成本的持续下降，以下升级路径已在研究框架中预留接口：

表25：未来升级路线

优先级	升级方向	触发条件	预期收益
P1	引入第三个模型（三模投票）	账户规模>\$20000	准确率+3-5%
P2	升级至 MiniMax/GLM 下一代推理模型	新版本发布且能力显著提升	评分质量提升，模型切换成本低
P3	评估海外轻量推理模型（如 GPT-5.4-mini, \$0.75/\$4.50 per 1M）	海外模型接入稳定性改善	以较低成本拓宽模型来源
P4	跟踪专用金融模型进展	垂直领域模型开源可用	金融文本理解专业度提升
P5	基于累积数据训练轻量评分模型	运行数据积累超过 3 年	摆脱外部 API 依赖，成本降至接近零

数据来源：东吴证券研究所绘制

当前配置体现了账户当前规模下的务实选择：在风控架构（独立双模型交叉验证）不打折扣的前提下，将 AI 成本控制在几乎可忽略的水平。

6.3. InvestmentClaw 的模型配置（独立体系）

InvestmentClaw (ai.openclaw.investment) 作为基于 OpenClaw 框架的独立 agent 实例，其 LLM 调用不经过上述量化评分流水线，而是直接由 OpenClawGateway 路由。当前配置：

表26: InvestmentClaw 的当前配置

角色	模型	说明
Primary	tokenhub/deepseek-v4-pro	主推理模型，处理投资分析、策略讨论、指令解析
Fallback	tokenhub/deepseek-v4-flash	当 pro 超时/限流时自动切换
图像模型	tokenhub/kimi-k2.6	生成图表、可视化报告

数据来源：东吴证券研究所绘制

注意：InvestmentClaw 的 deepseek-v4-pro/flash 配置与量化评分层的 MiniMax/GLM 属于两套独立体系。前者服务于对话式 agent 交互，后者服务于批量股票评分。

设计原因：对话式 AI 与批量评分对应两类不同工作负载。对话式 AI 要求快速响应，批量评分则强调批处理计算效率。两套体系相互隔离，可避免批量评分消耗对话式 AI 的额度或影响用户交互。

6.4. LLM 评分执行流程

对每只股票，MiniMax 与 GLM 的调用并非“主备”而是对等双跑，完整流程如下：

6.4.1. Step1 Prompt 组装

quant-system/llm.py 中的 `_build_stock_prompt()` 将以下信息组装为结构化 prompt：

- 标的信息：代码、当前价、日内涨跌、市值、所属引擎(growth/leverage/smallcap)
- 量化摘要：trend-score / signal-score / weight-score 三引擎子分、final-score、技术特征摘要（RSI、MACD、均线偏离度）
- 基本面：ROE、PE、营收增速、债务率（来自 yfinance）
- 新闻/异动：近期新闻标题、成交量异常、板块资金流向
- 指令：“请从价值投资与技术面角度给出 0-100 分的评分和 20 字以内理由，注意风险”

向 AI 提供多维信息的原因在于，AI 需要从多个角度判断标的：量化分较高但基本面恶化时可能降低评分；量化分一般但存在重大利好新闻时可能提高评分；技术指标偏强但估值过高时可能给予中性分数。

6.4.2. Step2 并发双模型调用

1. 并发请求：对同一只股票同时向 MiniMax-M2.7 和 GLM-5.1 发起请求（并行，非串行）

解释：同一标的同时向两个模型发起请求，而非串行调用，可降低整体耗时。

2. 超时控制：单模型超时 60s, max-tokens=4096（确保推理过程不被截断）

3. 接口层：统一走腾讯云 TokenHub OpenAI 兼容接口，共用同一 API Key

4. 响应解析：

提取 content 中的数字评分（正则匹配 0-100）

提取 content 中的理由文本（取前 50 字）

reasoning-content 仅用于调试日志，不参与评分计算

解释：AI 的“思考过程”只是记录，不作为最终评分依据

6.4.3. Step3 交叉验证融合（_fuse_llm_scores）

对两只模型返回的 (score-a, score-b) 执行融合：

表27：交叉验证融合

场景	融合规则	结果	通俗解释
均成功且方向一致 (同看多/同看空)	算术平均；若完全一致 (差值≤3)再×1.15增强	$fused = avg \times 1.15$	“两个专家意见一致，而且分数很接近→这个结果很可信，再加 15%权重”
均成功但方向分歧	取绝对值更小的一方（保守原则）	$fused = \min(ABS(a), ABS(b))$	“一个说 90 分，一个说 30 分→差异太大，取保守的”
单成功、单失败	独票降权×0.85（惩罚缺少第二意见）	$fused = valid_score \times 0.85$	“只有一个专家回复了，另一个掉线了→结果可能不太靠谱，打个 85 折”
均失败/均超时/解析失败	Fallback	$fused = 50$ （中性）	“两个专家都没回复→按中立处理，不给加分也不给扣分”

数据来源：东吴证券研究所绘制

实际例子:

表28: 交叉验证融合实例

股票	MiniMax	GLM	场景	融合结果
AAPL	82	78	方向一致, 差值 4 (>3)	$(82+78)/2=80$
NVDA	88	86	方向一致, 差值 2 (≤ 3)	$(88+86)/2 \times 1.15=100 \rightarrow \text{clip 到 } 100$
TSLA	75	35	方向分歧	$\min(75, 35)=35$
AMD	72	超时	单成功	$72 \times 0.85=61.2$
GME	超时	解析失败	均失败	50

数据来源: 东吴证券研究所绘制

6.4.4. Step4 注入 final_score

$$\text{final_score} = 0.6 \times \text{quant_score} + 0.2 \times \text{llm_fused} + 0.2 \times \text{inverse_score}$$

quant_score: 三引擎合并后的原始量化分 (占 60%, 权重最大)

llm_fused: Step3 融合后的 LLM 分 (占 20%)

inverse_score: 反向基本面疲弱信号的对冲分 (占 20%, 低 ROE/高负债/业绩下滑时扣分)

最终 clip 到 [0, 100], <30 视为异常不进入信号环节

该公式可理解为加权评分体系: 量化评分占 60%, 是基于历史数据的主要判断依据; AI 评分占 20%, 提供定性辅助判断; 反向评分占 20%, 用于在基本面恶化时进行风险扣分。

量化评分权重较高的原因在于其基于历史数据, 稳定性相对更强; AI 模型虽能提供定性增量信息, 但存在幻觉与误判风险, 因此仅占 20%。反向指标则承担风险校正功能。

实际例子:

表29: Final-score 实例

股票	quant_score	llm_fused	inverse_score	final_score
AAPL	78	80	75	$0.6 \times 78 + 0.2 \times 80 + 0.2 \times 75 = 77.8$
NVDA	85	100	70	$0.6 \times 85 + 0.2 \times 100 + 0.2 \times 70 = 87.0$
TSLA	70	35	60	$0.6 \times 70 + 0.2 \times 35 + 0.2 \times 60 = 61.0$
AMD	65	61	55	$0.6 \times 65 + 0.2 \times 61 + 0.2 \times 55 = 61.2$

数据来源: 东吴证券研究所绘制

6.4.5. 熔断与冷却

若连续 5 只票均失败 → 触发熔断, 本批次剩余候选全部 llm_score=50

解释：如果连续 5 次都问不出结果，说明 AI 服务可能崩了。这时候不再继续问，全部按中性处理，避免浪费时间和钱

熔断后下一批次延迟 60s 再尝试调用 LLM

解释：给 AI 服务 1 分钟时间恢复

6.5. 幻觉处理

LLM 并非完美。它偶尔会“幻觉”——编造不存在的数据、误解指标含义、或者给出与事实矛盾的评分。系统通过 Prompt 工程和结果校验两层防护来降低幻觉风险。

防幻觉设计：

1. 结构化输出：要求 AI 按固定格式返回（评分：XX + 理由：XX），方便正则提取。如果 AI 返回格式不对，直接标记为解析失败。
2. 指标注释：在 Prompt 中嵌入每个指标的解释，防止 AI 误解（如注明“负债率 <50% 为良好”）。
3. 风险提示：Prompt 末尾统一附加“注意风险”，提醒 AI 不要过度乐观。
4. 长度限制：max-tokens=4096，防止 AI 输出过长导致解析困难。

幻觉案例与处理：

案例 1：编造数据 - 输入：某股票 ROE = 15% - MiniMax 返回：“ROE 高达 150%，公司盈利能力惊人，评分 95” - 处理：解析层发现 AI 提到的“150%”与输入数据“15%”矛盾 → 标记为“数据不一致” → 该票降权处理（llm-score × 0.9） → TG 告警：“⚠️ MiniMax 评分数据异常，已降权”

案例 2：过度反应 - 输入：某股票量化评分 82，但有一篇负面新闻“该公司面临监管调查” - GLM 返回：评分 25，“监管风险极高，建议回避” - 处理：方向分歧 → 取保守值 $\min(82, 25) = 25$ → $\text{final_score} = 0.6 \times 82 + 0.2 \times 25 + 0.2 \times 60 = 63.2$ - 结果：量化评分权重 60% 起到了“锚定”作用，final-score 仍达 63.2（BUY 级别），没有被 AI 的极端评分带偏。

量化权重 60% 作为防幻觉“压舱石”的原因

即使 AI 出现严重幻觉（如把某只股票从 80 分打到 20 分），由于 AI 只占 20% 权重：- 最坏情况：quant=80, llm=20, inverse=60 → $\text{final} = 0.6 \times 80 + 0.2 \times 20 + 0.2 \times 60 = 68$ （仍 ≥ 65, BUY）- 只有当量化评分也低（如 quant=50）时，AI 的低分才会把 final-score 拉低到 50 以下

这就是“量化做主角、AI 做配角”的设计哲学——AI 提供增量信息，但不允许它“一票否决”。

6.6. LLM 熔断与跨批次冷却

- 熔断：连续 5 次调用失败时，本批次直接跳过 LLM，所有候选统一赋 $llm_score=50$ （中性）
- 冷却：熔断触发后，下一批次的 LLM 调用延迟 60 秒启动（避免接口刚恢复即被再次压垮）
- 速率控制：每批次并发 ≤ 4 ，相邻调用间隔 $\geq 200ms$

解释：API 存在调用频率限制，过快调用可能触发限流。因此系统设置并发上限与调用间隔。

熔断场景模拟：

批次一（前 5 只股票）： - 第 1 只：MiniMax 成功，GLM 成功 → 正常融合 - 第 2 只：MiniMax 成功，GLM 成功 → 正常融合 - 第 3 只：MiniMax 超时，GLM 超时 → 失败计数 = 1 - 第 4 只：MiniMax 超时，GLM 成功 → 失败计数 = 2（MiniMax 失败） - 第 5 只：MiniMax 超时，GLM 超时 → 失败计数 = 3 - 第 6 只：MiniMax 超时，GLM 超时 → 失败计数 = 4 - 第 7 只：MiniMax 超时，GLM 超时 → 失败计数 = 5 → 熔断触发 - 第 8-20 只：不再调用 LLM，全部 $llm_score=50$

批次二（15 分钟后）： - 延迟 60 秒后重新尝试 - 如果第 1 只就成功 → 恢复正常 - 如果仍然失败 → 继续熔断，全部 50

设计原因： - 熔断：防止在 AI 服务异常时持续重试并浪费 API 成本。 - 冷却：为 AI 服务恢复预留时间窗口。 - 速率控制：降低触发 API 限流或封禁的概率。

6.7. final_score 下限保护

$final_score$ 统一 clip 到 $[0, 100]$ 区间， < 30 视为评分异常，不进入信号生成环节。

下限设为 30 而非 0 的原因：即使是较弱标的，也可能因 yfinance 数据缺失等原因导致部分因子异常。设定 30 分作为缓冲，可降低误杀概率；若 $final_score$ 低于 30，则说明评分明显异常，不进入信号生成环节。

6.8. LLM 评分的局限性

尽管 AI 评分提供了量化模型无法捕捉的定性信息，但读者需要清醒认识它的局限性：

6.8.1. 局限性 1: 训练数据的时效性

- MiniMax 和 GLM 的训练数据截止到 2024 年中，对 2024 下半年和 2025 年的新事件（如某家公司新推出的产品、行业新规）可能不了解。

- 例子：2025 年某 AI 芯片公司发布革命性产品，股价暴涨。但 AI 模型可能不知道这个产品，因为它训练时还没有。

- 缓解：Prompt 中提供新闻摘要，但摘要质量取决于爬虫，且新闻本身可能不完整。

6.8.2. 局限性 2: 无法获取非公开信息

- AI 只能基于公开信息（财报、新闻、技术指标）评分，无法获取内幕信息、管理层真实意图、竞争对手未发布的产品等。

- 例子：某公司 CEO 私下计划辞职，但尚未公告。AI 评分可能仍然很高，因为公开信息一切正常。

- 缓解：没有完美解决方案。这也是量化权重占 60% 的原因之一——量化模型同样不知道内幕，但至少不假装知道。

6.8.3. 局限性 3: 对复杂财务工程的理解有限

- AI 对简单指标（ROE、PE、负债率）理解较好，但对复杂财务操作（如 SPAC 合并、可转债稀释、表外负债）可能误判。

- 例子：某公司 ROE 高达 200%，但这是因为大量回购股票减少了股东权益（分母变小），而非真实盈利能力提升。AI 可能误给高分。

- 缓解：inverse-score（反向评分）专门检测这类“财务工程”信号，但也不是 100% 准确。

6.8.4. 局限性 4: 情绪化与一致性

- 同一股票在不同日子问 AI，可能得到不同评分（因为模型有随机性）。

- 测试显示：同一股票同一 Prompt 问 10 次，评分标准差约 5-8 分。

- 缓解：两个模型交叉验证+取平均，降低了单模型随机性的影响。

6.8.5. 局限性 5: 成本约束下的信息压缩

- 系统每天只评 Top 20 只股票，因为 LLM 调用有成本。这意味着排名 21-180 的股票虽然也有量化评分，但缺少 AI 的“第二意见”。
- 缓解：Top 20 的截断基于量化评分，理论上量化评分最高的 20 只最需要 AI 验证。

6.8.6. 总结: AI 评分的正确用法

AI 评分不是确定性预测工具，而是：1. 增量信息源：捕捉量化模型难以覆盖的定性风险（如监管调查、管理层变动）；2. 交叉验证器：当量化评分异常高时提供“第二意见”；3. 风险哨兵：当 AI 识别重大风险时，即便量化评分较高，也会拉低 final_score。

不应高估 AI，也不应忽视 AI 的增量信息。把它当作一个“可提供增量信息但仍需校验的分析工具”——它的意见值得参考，但最终决策权在量化模型（60% 权重）和风控系统手中。

7. 信号生成: 从分数到实际行动

7.1. Action 分级

系统根据 final_score 和持仓状态，把每只股票归类为以下六种“行动建议”：

表30: 六种行动建议

Action	条件	含义	通俗解释
STRONG-BUY	final_score ≥ 75 且非持仓	强买入候选	“强烈推荐买入，分数很高”
BUY	final_score ≥ 65 且非持仓	普通买入	“可以考虑买入”
HOLD	持仓+final_score ≥ 45	继续持有	“已经买了的，继续拿着”
SELL	持仓+final_score < 45	评分跌破阈值卖出	“这票不行了，卖掉”
STOP_LOSS	持仓+价格触及止损线	风控强制卖出	“跌了 4%，必须止损，不管评分多少”
TAKE_PROFIT	持仓+价格触及止盈线	止盈落袋	“涨了 20%，落袋为安”

数据来源：东吴证券研究所绘制

说明：当前生产配置 buy_score_threshold=60（偏激进），经 IC 分析与回测双重验证为最优区间。

关键理解：STOP_LOSS 和 TAKE_PROFIT 是独立的强制规则。即使某只股票的

final_score 还有 80 分,但如果它已经跌了 4%,系统仍然会执行止损。这是因为“再强的股票也可能继续跌”,风控优先于评分。

边界场景讨论:

场景 A:分数刚好在边界 - 某股票 final_score = 74.9 → BUY(不是 STRONG-BUY) - 某股票 final_score = 75.0 → STRONG-BUY - 差距只有 0.1 分,但待遇不同。这是阈值系统的固有特性。回测显示,75 分以上的股票胜率显著高于 65-75 分的股票,所以这个分界是有意义的。

场景 B:持仓股票的分数变化 - 用户持有 AAPL,买入时 final_score = 80 - 一周后 AAPL 下跌 5%,但 final_score 仍为 78(因基本面未变) - 此时不会触发 SELL(因 ≥ 45),但会触发 STOP-LOSS(因下跌 4%) - 结果:止损卖出 - 启示:评分与价格是两套独立约束。评分反映标的是否仍具备持有价值,价格约束反映当前持仓回撤是否触及风控线。

场景 C:刚止损的股票 - 某股票触发止损后,系统卖出 - 次日该股票 final_score = 82(因新闻利好) - 系统不会立即买回,因为仍处于 48 小时冷却期($\$8.5$) - 第三天冷却期结束后,若分数仍 ≥ 65 ,系统可重新评估买入 - 设计意图:避免止损后过早重新建仓。

7.2. 买入候选 TopN 筛选

一、按 final_score 降序排列所有 BUY / STRONG-BUY

二、取 Top max-positions - current-positions 只

例子:如果系统算出来最多可以持 10 只,现在已经持了 4 只,那么最多再买 6 只

三、再经三道过滤:

1. 流动性:日均成交额 ≥ 500 万美元 (smallcap 引擎除外)

解释:确保具备较好成交与退出条件

500 万美元门槛的设置依据:若买入金额为 5000 美元,仅占日均成交额 0.1%,通常不会明显影响价格;若买入金额升至 5 万美元并占日均成交额 1%,则可能带来更明显滑点。

2. 黑名单: blacklist.txt 中的标的永不买入

解释:部分股票因财务造假、监管调查或用户偏好等原因需要纳入限制名单。

例子:若某股票被 SEC 调查财务造假,用户可将其加入黑名单,系统此后不再买入该标的。

3. 冷却期: 近期因止损卖出的标的 48 小时内不买入

解释: 刚止损的股票可能仍处于下跌趋势中, 48 小时内不买入可降低过早重新建仓风险。

例子: 某股票以 100 美元买入, 跌至 96 美元触发止损 (-4%)。次日继续跌至 92 美元, 若立即买回, 将再次承受回撤。48 小时后若其反弹至 98 美元, 系统可重新评估。

7.3. 信号终调规则

信号生成后执行最终调节, 类似于“主编审稿”——在正式发布前再做一轮把关:

7.3.1. 现金比例

剩余现金 < 总资产 5% 时, 抑制 BUY 信号 (防止满仓后缺乏流动性缓冲)

解释: 如果账户里只剩 5% 的现金, 说明几乎满仓了。这时候不再买新股票, 留点现金应急

例子: 账户总资产 10000 美元, 现金只剩 400 美元 (4%)。即使有 STRONG-BUY 信号, 系统也不买, TG 通知: “现金不足 5%, 暂停买入”

5% 现金比例的设置依据: 回测显示, 保留 5% 现金在多数市场环境下较优。现金比例过高会拖累收益, 过低则缺乏流动性缓冲。

7.3.2. 单日买入笔数上限

由 `compute_daily_budget()` 动态计算 $N \in [2, 15]$ ($VIX \times \text{账户浮盈} \times \text{事件日加成}$), 仅约束 spontaneous 桶 (主动 setup)。instruction 桶 (用户手动下达订单) 与 risk 桶 (止损/止盈/强平) 不计入 N。详见 § 8.6 三桶预算

7.3.3. 行业集中度

同行业持仓 ≥ 4 只时, 抑制该行业新增 BUY

解释: 行业集中度约束用于避免组合过度集中于单一行业。组合优化层允许同一行业最多 3 只 (§ 14.1); 当信号层检测到已有 3 只时, 第 4 只开始抑制。

例子: 若组合已持有 AAPL (科技)、MSFT (科技) 和 NVDA (科技), 当日 AMD 也发出 BUY 信号, 系统不会买入, 因为科技行业持仓已达到 3 只上限。

行业怎么分? 系统使用 GICS (全球行业分类标准), 把股票分成 11 个大行业: 科技、医疗、金融、消费品、工业、能源、原材料、房地产、通信、公用事业、非必需消费品。

7.3.4. 单票红线

任何 BUY 信号若会使目标票 $\geq 30\%$ 权益 \rightarrow 直接 drop, 不降仓执行 (与 SSOT 单票仓位上限 30% 硬对齐)

解释: 哪怕某只股票再好, 也不能把超过三成的钱押在它身上

例子: 账户 10000 美元, AAPL 发出 STRONG-BUY, 系统算出来要买 4000 美元 (40%)。触发单票红线, 直接拒绝, TG 通知: “拒绝: AAPL 将超过 30% 仓位红线”

7.3.5. 杠杆 ETF 合计红线

配置已预留 `max-leverage-etf-total-pct=0.50`, 但交易执行层尚未实现实时校验

7.3.6. 持仓数上限

由 `compute-max-positions(equity)` 动态算 $\in [3, 20]$ (详见 § 8.6.5); 满仓时候选被推到 TG “满仓提示 (X/Y)” 告警, 并尝试触发自动换仓

解释: 小账户可能只能持 4 只, 大账户可以持 20 只。满了就不买新的, 除非有更强的股票可以把弱的换掉。

7.4. Action 状态机与冲突处理

一只股票不可能同时处于“买入”和“卖出”状态, 但不同股票之间可能产生冲突。以下是系统处理冲突的规则:

Action 优先级 (从高到低):

STOP-LOSS / TAKE-PROFIT / 移动止损 > SELL > BUY / STRONG-BUY > HOLD

Action 优先级排序的原因

1. 止损/止盈最优先: 这是“保命”的操作, 必须立刻执行, 不能犹豫
2. SELL 次之: 评分跌破阈值的卖出, 意味着“这票不行了”
3. BUY 再次: 开新仓是“进攻”, 应该在“防守”(卖)之后
4. HOLD 最低: 什么都不做, 最保守

实际执行顺序 (先卖后买):

系统在每一轮交易中的执行顺序是固定的: 1. 先处理所有持仓的 STOP-LOSS / TAKE-PROFIT / 移动止损 (紧急卖出) 2. 再处理 SELL (评分跌破阈值的卖出) 3. 卖出完成后, 刷新账户余额 (现金增加) 4. 最后处理 BUY / STRONG-BUY (买入)

先卖后买的原因

假设账户有 \$10000, 持仓 \$6000, 现金 \$4000, 当日存在以下交易需求: - 止损 TSLA (释放 \$2000 现金) - 买入 AAPL (需要 \$3000)

如果先买后卖： - 先买 AAPL \$3000 → 现金从 \$4000 降到 \$1000 - 再止损 TSLA → 现金从 \$1000 升到 \$3000 - 问题：如果 AAPL 的买入需要 \$3000 但现金只有 \$4000，先买是没问题的。但如果某天要买 \$5000，现金只有 \$4000，先买就会失败

如果先卖后买： - 先止损 TSLA → 现金从 \$4000 升到 \$6000 - 再买 AAPL \$3000 → 现金从 \$6000 降到 \$3000 - 优势：卖出释放的现金可用于本轮买入，不会出现“想买但钱不够”的情况

冲突场景 1: 同一只股票同时触发 BUY 和 SELL

不可能发生。因为： - BUY 的条件: $final_score \geq 65$ 且非持仓 - SELL 的条件: $final_score < 45$ 且持仓 - 同一只股票不可能同时“持仓”和“非持仓”

冲突场景 2: 新买入的股票立即触发止损

可能发生。例如： - 21:30 买入 NVDA @ \$900, $final_score = 80$ - 22:00 盘中监控: NVDA 跌到 \$864 (-4%) - 触发 STOP_LOSS - 系统立即卖出 NVDA - 结果: 买入 30 分钟后就止损, 亏了 4% + 手续费

这不是 bug, 而是“及时止损”的设计。即使买入时评分很高, 如果价格立刻反向运动, 说明买入决策错了 (或市场突发利空), 认错止损比死扛更明智。

冲突场景 3: 满仓时来了 STRONG-BUY

假设： - 账户 $max_positions = 6$, 已持仓 6 只 - 今天 AAPL 发出 STRONG-BUY, $final_score = 85$ - 但持仓已满, 无法买入

系统行为： 1. AAPL 进入“候选队列”但不执行买入 2. TG 通知：“满仓提示: AAPL STRONG-BUY 但持仓已满 (6/6)” 3. 系统检查当前 6 只持仓中最弱的一只 (如 INTC, $final_score = 42$) 4. 如果 $AAPL(85) - INTC(42) = 43 \geq diff_th$ (换仓分差门槛) → 触发换仓: 卖出 INTC, 买入 AAPL 5. 如果分差不足 → 不换仓, AAPL 留在候选队列等下一轮

冲突场景 4: 多只股票同时 BUY, 现金不够买全部

假设： - 现金 \$5000 - 今天有 3 只 BUY: NVDA (目标 \$3000)、AAPL (目标 \$2000)、MSFT (目标 \$1500) - 合计目标 \$6500 > 现金 \$5000

系统行为： 1. 按 $final_score$ 排序: $NVDA(85) > AAPL(78) > MSFT(72)$; 2. 先买 NVDA \$3000, 剩余现金 \$2000; 3. 再买 AAPL \$2000, 剩余现金 \$0; 4. MSFT 因现金不足被跳过, 并进入 Telegram “资金不足跳过”通知。未采用按比例缩减的原因在于, 按比例缩减可能导致每只标的均低于最小订单金额 (\$200), 产生过多碎片单。系统倾向于优先买足高分标的, 而非分散买入多个低金额订单。

7.5. 信号生成完整实例 (“看系统怎么过一天”)

让我们跟踪一天的完整信号生成过程，看看系统从头到尾做了什么决策。

假设：账户总资产：\$10000-当前持仓：AAPL(\$1800)、MSFT(\$1500)、TSLA(\$1200)
 → 共 3 只，\$4500-可用现金：\$5500-
 $max_positions = compute_max_positions(10000) = clamp(3, 10000/100/3, 20) = 10$ -因此
 今天最多还能买： $10 - 3 = 7$ 只-VIX=18(中性)，昨日浮盈+1.2%→ $daily_budgetN = 5 \times 1.0 \times 1.0 = 5$ 笔-今日已用 spontaneous 笔数：0/5

7.5.1. Step1 选股 (21:00SGT)

universe_builder.py 运行：-Finvizgrowth-quality：筛出 42 只 (AAPL、MSFT、NVDA、AVGO、LLY、V、MA、GOOGL、META...) -Finvizmid-momentum：筛出 31 只 (CRWD、SNOW、PLTR、APP、CELH、NU...) -Finvizturnaround：筛出 18 只 (INTC、BA、DIS...) -Finvizbreakout：筛出 25 只 (某 AI 小盘股、某生物科技股...) -异动扫描：发现 28 只 (NVDA 放量+8%、某中概股轧空...) -SEED-TICKERS：约 40 只 (其中 25 只已在 Finviz 中) -杠杆 ETF 池：25 只 (TQQQ、SOXL、TECL、FNGU、TNA...)

合并去重后：187 只 > 180 只上限 → 截断至 180 只 (优先保留异动股票)

7.5.2. Step2 三引擎评分 (21:05SGT)

scoring.py 对 180 只候选进行评分 (仅展示部分)：

表31：实例三引擎评分

股票	引擎	原始分	拉伸后	后置调节	最终 quant_score
NVDA	growth	76	88	+3 (放量大涨) =91	91
AAPL	growth	72	82	无=82	82
MSFT	growth	70	80	无=80	80
TQQQ	leverage	78	85	+5 (牛市) =90	90
TSLA	growth	68	75	无=75	75
INTC	growth	55	58	无=58	58
PLTR	growth	65	72	无=72	72
CRWD	growth	74	84	无=84	84
某小盘股	smallcap	62	70	无=70	70

数据来源：东吴证券研究所绘制

7.5.3. Step3 波动率过滤

MSTR (波动率 85%) → 剔除

COIN (波动率 72%) → 剔除

TQQQ (波动率 65%) → 保留 (豁免)

某小盘股 (波动率 95%) → 保留 (豁免)

剩余 175 只进入 AI 评分。

7.5.4. Step4 AI 评分 (21:10SGT)

Top20 进入 LLM 评分 (仅展示部分) :

表32: 实例 AI 评分

股票	MiniMax	GLM	融合规则	llm-fused
NVDA	88	85	一致, 差值 $3 \leq 3 \rightarrow \text{avg} \times 1.15 = 99.5 \rightarrow \text{clip}100$	100
TQQQ	82	80	一致, 差值 $2 \leq 3 \rightarrow \text{avg} \times 1.15 = 92.4$	92
CRWD	75	78	一致, 差值 $3 \leq 3 \rightarrow \text{avg} \times 1.15 = 88.0$	88
AAPL	80	82	一致, 差值 $2 \leq 3 \rightarrow \text{avg} \times 1.15 = 93.2$	93
PLTR	65	45	分歧 $\rightarrow \min(65, 45) = 45$	45

数据来源: 东吴证券研究所绘制

7.5.5. Step5 final_score 计算

公式: $\text{final_score} = 0.6 \times \text{quant_score} + 0.2 \times \text{llm_fused} + 0.2 \times \text{inverse_score}$

表33: 实例 final-score 计算

股票	quant_score	llm_fused	inverse_score	final_score	Action
NVDA	91	100	70	$0.6 \times 91 + 0.2 \times 100 + 0.2 \times 70 = 88.6$	STRONG_BUY
AAPL	82	93	75	$0.6 \times 82 + 0.2 \times 93 + 0.2 \times 75 = 83.4$	已持仓 \rightarrow HOLD
TQQQ	90	92	60	$0.6 \times 90 + 0.2 \times 92 + 0.2 \times 60 = 84.4$	STRONG_BUY
CRWD	84	88	72	$0.6 \times 84 + 0.2 \times 88 + 0.2 \times 72 = 82.4$	STRONG_BUY
MSFT	80	85	80	$0.6 \times 80 + 0.2 \times 85 + 0.2 \times 80 = 81.0$	已持仓 \rightarrow HOLD
TSLA	75	70	65	$0.6 \times 75 + 0.2 \times 70 + 0.2 \times 65 = 72.0$	已持仓 \rightarrow HOLD
PLTR	72	45	55	$0.6 \times 72 + 0.2 \times 45 + 0.2 \times 55 = 63.2$	BUY
INTC	58	50	45	$0.6 \times 58 + 0.2 \times 50 + 0.2 \times 45 = 52.8$	- (<65, 无 Action)

数据来源: 东吴证券研究所绘制

7.5.6. Step6 买入候选筛选

BUY/STRONG_BUY 列表: 1. NVDA—STRONG_BUY (88.6) 2. TQQQ—STRONG_BUY (84.4)
3. CRWD—STRONG_BUY (82.4) 4. PLTR—BUY (63.2)

按 final_score 降序排列: NVDA>TQQQ>CRWD>PLTR

取 Top(10-3)=7 只, 但这里只有 4 只, 全部保留。

7.5.7. Step7 三道过滤

1. 流动性: NVDA 日均成交\$300 亿✓、TQQQ\$50 亿✓、CRWD\$8 亿✓、PLTR\$12 亿
2. 黑名单: 均不在 blacklist.txt 中✓
3. 冷却期: 均无近期止损记录✓
- 4 只全部通过。

7.5.8. Step8 信号终调 (“主编审稿”)

逐条检查:

1. 现金比例: 现金\$5500/总资产\$10000=55%>5%✓ (不抑制)
2. 单日买入笔数: 今日已用 0/5, 4 只 BUY 需要 4 笔≤5✓
3. 行业集中度:

NVDA (科技)、TQQQ (ETF)、CRWD (科技)、PLTR (科技)

当前持仓已有 AAPL (科技)、MSFT (科技)

如果买入 NVDA+CRWD+PLTR, 科技行业将达到 5 只>3 只上限

抑制: CRWD 和 PLTR 的行业集中度超限, 只保留 NVDA

4. 单票红线:

NVDA target_weight 约 35% → 目标金额\$5500 × 35%=\$1925 → 占权益 19.25%<30%✓

TQQQ target_weight 约 30% → 目标金额\$1650 → 占权益 16.5%<30%✓

5. 持仓数上限: 当前 3 只+买入 2 只=5 只≤10✓

7.5.9. Step9 最终信号

表34: 实例最终信号

股票	Action	target_weight	目标金额	状态
NVDA	STRONG_BUY	35%	\$1925	✓通过终调
TQQQ	STRONG_BUY	30%	\$1650	✓通过终调
CRWD	STRONG_BUY	—	—	✗行业集中度抑制 (科技已满)
PLTR	BUY	—	—	✗行业集中度抑制 (科技已满)
AAPL	HOLD	—	—	已持仓, 保持
MSFT	HOLD	—	—	已持仓, 保持
TSLA	HOLD	—	—	已持仓, 保持

数据来源: 东吴证券研究所绘制

TG 通知 (21:15SGT) : 信号生成完成 (2026-04-27)

通过终调:

[STRONG-BUY] NVDA—final_score88.6—目标\$1925

[STRONG-BUY] TQQQ—final_score84.4—目标\$1650

被终调拦截:

CRWD—行业集中度 (科技板块已满 3/3)

PLTR—行业集中度 (科技板块已满 3/3)

持仓状态:

HOLD: AAPL (82) MSFT (80) TSLA (72)

账户: \$10000 | 现金\$5500 → 买入后\$1925

日预算: 2/5 笔已用

为什么 CRWD 和 PLTR 被拦截是正确的设计?

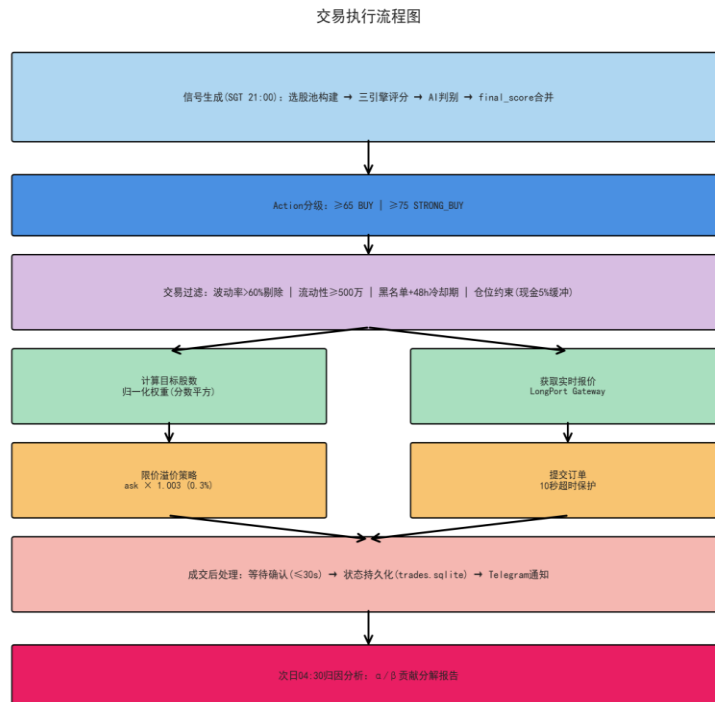
虽然 CRWD (final_score82.4) 和 PLTR (63.2) 都是 BUY 级别, 但如果买入它们, 科技行业将占持仓的 5/5=100%。一旦科技板块回调(如美联储鹰派表态、AI 泡沫担忧), 整个账户将遭受重创。行业分散的”成本”是错过了 CRWD 和 PLTR 的上涨机会, 但”收益”是降低了单一行业崩盘的风险。

这就是信号终调的核心价值: 在”好的机会”和”安全的组合”之间取得平衡。

8. 交易执行: 从信号生成到成交落地

8.1. 单笔交易流程 (以买入为例)

图4：交易执行流程图



数据来源：东吴证券研究所绘制

关键概念解释： - target_weight (目标权重)：如果系统觉得 AAPL 值得买，它会先算“这只股票应该占用户总资产的多少”。18% 意味着“如果用户有 10000 美元，应该拿 1800 美元买 AAPL”。 - 权重怎么来的？见 § 7.3 仓位权重计算 - 权重会变化吗？会。每轮评分都会重新计算权重。如果 AAPL 的分数下降了，权重可能从 18% 降到 12%。

- bid/ask (买一档/卖一档)：股票交易采用双向报价。bid 为买方报价，ask 为卖方报价。买入时的成交价格通常接近 ask。

例子：AAPL 的 bid = 185.00, ask = 185.05。这意味着：

有人愿意用 185.00 美元买 AAPL

有人愿意用 185.05 美元卖 AAPL

买卖价差 = 0.05 美元 = 0.027%

关注 ask 的原因：买入方需要参考卖方报价，实际买入价格通常围绕 ask 形成。

- 限价单 (Limit Order)：“我只愿意以这个价格或更好的价格买入”。比如当前 ask 是 185.00，系统下限价单 185.56 (= 185.00 × 1.003)，意味着”185.56 以

下我才买，如果涨得太快我就不追了”。

采用溢价限价的原因：市场价格实时波动。若直接下市价单，可能出现较大滑点；限价单可控制交易成本，而适度溢价（0.3%-1.0%）有助于提高成交概率。

若未成交：限价单最长等待 30 秒。若 30 秒内未成交（如价格快速上行），系统会取消订单，并通过 Telegram 通知：“AAPL 限价单未成交，当前 ask = 185.80，已超过限价”。

8.2. 券商终端的 API 超时保护

券商终端的部分接口偶尔会出现”调用挂起”状态（无返回、无报错）。所有 API 调用均设置了超时保护：

表35：券商终端超时保护

操作	超时	失败动作	通俗解释
查询报价	5s	重试 2 次，仍失败则本次跳过	报价超时重试，失败则跳过本轮
下单	10s	重试 1 次，仍失败发送 TG 告警并跳过	下单超时重试，失败则告警跳过
撤单	5s	失败发送 TG 告警，保留订单状态不变	撤单失败即告警并保留订单状态
查询持仓	8s	失败时使用缓存持仓，发送 TG 告警	持仓查询失败时启用缓存并告警

数据来源：东吴证券研究所绘制

API 挂起的可能原因：1. 券商终端服务器负载过高；2. 网络连接不稳定；3. 请求数据量过大；4. 券商终端处于系统维护状态。

实际案例：某日 21:35（美股开盘后不久），券商终端服务器出现短暂拥堵，查询报价接口 5 秒内未返回。系统第 1 次重试仍超时，第 2 次重试成功返回。全流程耗时 12 秒，对交易影响有限。若 3 次均失败，系统将跳过该标的本轮交易，并在下一轮（10 分钟后）重新尝试。

8.3. 仓位权重计算

target_weight 即该标的在本批次所有 BUY 候选中的归一化得分权重：

$$\text{weights}[i] = \text{scores}[i]^2 / \sum (\text{scores}[j]^2 \text{ for } j \text{ in BUY_candidates})$$

该计算用于实现得分加权分配。系统并非平均分配资金，而是使高分股票获得更高资金权重。采用平方而非线性比例，是为拉大高分与中分标的之间的资金分配差距。例如，80 分与 60 分标的按平方计算的权重比为 $6400:3600 \approx 1.78:1$ ，而非线性比例下的 1.33:1。

实际例子：

假设本轮有 3 只 BUY 候选：

表36: 仓位权重计算

股票	final_score	平方	权重
NVDA	85	7225	7225/15050=48.0%
AAPL	78	6084	6084/15050=40.4%
MSFT	65	4225	4225/15050=28.1%
合计		15034	116.5%→归一化

数据来源: 东吴证券研究所绘制

合计超过 100%了, 归一化处理:

NVDA 权重=48.0%/116.5%=41.2%

AAPL 权重=40.4%/116.5%=34.7%

MSFT 权重=28.1%/116.5%=24.1%

合计=100%

如果账户总资产 10000 美元, 现金 6000 美元 (假设已经持仓 4000 美元):
 - NVDA: $6000 \times 41.2\% = 2472$ 美元 - AAPL: $6000 \times 34.7\% = 2082$ 美元 - MSFT:
 $6000 \times 24.1\% = 1446$ 美元

采用平方权重而非线性权重的原因

如果用直接比例: - NVDA: $85 / (85+78+65) = 85/228 = 37.3\%$ - AAPL: $78 / 228 = 34.2\%$ - MSFT: $65 / 228 = 28.5\%$

差距不够大。平方的作用是让高分股票获得更多资金, 体现“好股票值得重仓”的理念。

最终权重受 max_single_order_pct (硬红线 30%, 与 SSOT 单票仓位上限对齐) 与可用现金共同约束。

8.4. cmd_trade 完整执行流程

cmd_trade 是全自动交易的核心函数, 被 cron (定时任务) 和 cmd_monitor (盘中监控) 循环调用。以下是其内部执行的精确顺序:

8.4.1. Step1 启动自检

1. 市场参数校验: market 必须是 us (美股) 或 hk (港股), 否则尝试兼容旧式单笔下单调用或报错返回。

2. 模块可用性: 检查 quant-stock-picker 是否已加载。如果没有, 说明策略源码不在 PYTHONPATH 上, 直接返回。

3. 交易日判断: 调用 `.is-trading-day()`, 非交易日直接返回。

怎么判断? 系统维护了一个交易日历 (NYSE 交易日历), 会检查今天是否是交易日。

8.4.2. Step2 预算计算 (核心动态逻辑)

1. 获取或创建 RiskManager 实例 (风控管理器)。

2. 若未传入 `daily-budget`, 调用 `compute-daily-budget()` 动态算出当日 N:

$N = \text{clamp}(\text{round}(5 \times \text{VIX 系数} \times \text{持仓系数} + \text{事件加成}), 2, 15)$

VIX 系数: $<15 \rightarrow 0.6, 15-20 \rightarrow 1.0, 20-25 \rightarrow 1.3, >25 \rightarrow 1.6$

持仓系数: 浮盈 $>+2\% \rightarrow 1.2, -2\% \leq \text{浮盈} < +2\% \rightarrow 1.0, <-2\% \rightarrow 0.7$

事件加成: 财报日 +2, 议息日 +3, CPI/PPI/NFP 日 +2

VIX 是什么? 全称 Volatility Index, 俗称“恐慌指数”。VIX <15 表示市场很平静, 大家都很淡定; VIX >25 表示市场在恐慌, 波动很大。

持仓系数是什么? 看你最近赚钱还是亏钱。如果最近赚了不少 (浮盈 $>+2\%$), 说明状态好, 可以多交易; 如果亏了不少 (浮盈 $<-2\%$), 说明状态差, 减少交易。

事件加成是什么? 财报日、美联储议息日、CPI 公布日都是市场波动大的日子, 所以多给几次交易机会。

`clamp(2, 15)` 是什么? 不管怎么算, 最少 2 笔, 最多 15 笔。防止极端情况下算出来 0 笔或 100 笔。

实际例子:

平静日 (VIX13, 昨天平盘, 无事件): $5 \times 0.6 \times 1.0 = 3$ 笔

财报周 (VIX20, 昨天赚了 2.5%, META 发财报): $5 \times 1.0 \times 1.2 + 2 = 8$ 笔

黑天鹅 (VIX32, 昨天亏了 3%, 美联储议息): $5 \times 1.6 \times 0.7 + 3 = 9$ 笔

若 `compute-daily-budget()` 失败 (比如 VIX 数据拿不到), 则 fallback 到 6 笔。

3. 调用 `risk.set-daily-budget(N)` 写入预算。

4. 交易计数同步: 调用 `risk.sync-today-orders(ctx)`, 从券商终端拉取今日已成交订单, 区分 `spontaneous/instruction/risk` 三桶, 确保计数与券商一致。

8.4.3. Step3 账户+持仓快照

1. 账户余额: 调用 `ctx.account-balance()` (超时 8s), 计算 `total-equity` (总资产)

和 available_cash (可用现金)。

ototal_equity=现金+所有持仓的市值

oavailable_cash=真正能用的现金 (扣除了已挂单冻结的部分)

2. 持仓查询: 调用 ctx.stock_positions(超时 8s), 遍历所有 channel 的 positions, 构建 held_positions 字典。

3. 将持仓盈亏比写入 risk.update_position_pl(), 用于后续移动止损计算。

8.4.4. Step4 三引擎评分

1. 加载 default_config() 和美股 watchlist。

2. 判断是否为开盘窗口 (opening_window)。

3. 若为开盘窗口, 调用 _get_realtime_snapshots_by_yf() 获取全量标的实时行情。

4. 调用 run_quant_scoring() 执行完整三引擎评分流水线。

5. 过滤出美股标的 (排除.HK)。

8.4.5. Step5 生成买卖候选

遍历评分结果, 按 action 分类处理。

止损/止盈/移动止损检测 (对每只已持仓): 1. 根据引擎类型读取对应止损/止盈/移动止损阈值 (成长 4%/20%/4%, 杠杆 8%/28%/5%, 小盘 6%/32%/5%)。2. 从 risk.get_peak_pl() 获取峰值盈亏比。3. 触发条件: $-pl_ratio \leq -stop_loss_pct \rightarrow$ 止损卖出 $-pl_ratio \geq take_profit_pct \rightarrow$ 止盈卖出 $-peak_pl \geq t_act$ 且 $peak_pl - pl_ratio \geq trailing_stop_pct \rightarrow$ 移动止盈卖出 4. 关键设计: 止损/止盈/移动止盈属于 risk 桶, 不计入日预算 N、不进入确认模式。

移动止损的完整例子:

你 100 美元买入某成长股: -第 1 天: 涨到 105 (+5%) $\rightarrow peak_pl = +5%$, 未触发移动止损 (需要 $peak_pl \geq$ 激活阈值 10%) -第 5 天: 涨到 112 (+12%) $\rightarrow peak_pl = +12%$, 超过激活阈值 10%, 开始追踪移动止损 -第 7 天: 跌到 108 (+8%) \rightarrow 从最高点回落 = $12\% - 8\% = 4\%$ \rightarrow 刚好等于 $trailing_stop_pct = 4\%$ \rightarrow 触发移动止损 -系统以市价单卖出, 成交价约 108 美元 -最终盈利: 8%

如果没有移动止损: -第 7 天: 跌到 108, 你没卖 -第 10 天: 跌到 95 (-5%) \rightarrow 触发止损, 亏损 5% -结果: 从 +12% 到 -5%, 过山车

移动止损的作用就是“让利润奔跑, 但别让利润全部跑掉”。

8.4.6. Step6 卖单执行

为什么要先卖后买？因为卖出能释放现金。如果先买后卖，可能出现”想买但钱不够”的情况。

1. 分级: emergency(止损/移动止盈/止盈)优先执行; normal(REDUCE/REDUCE_HALF)进入确认模式。

2. 确认模式: 若 `_CONFIRM_MODE_ENABLED=True`, 常规卖出需 TG 确认; 紧急卖出不进入人工确认流程, 由系统直接提交市价卖出指令。

3. 风控检查: 常规卖出调用 `risk.can-trade()` 校验; 紧急卖出绕过 `can-trade`。

4. 下单: 市价单 (`OrderType.M0, Sell`), 超时 10s。

5. 成交验证: 先调用 `_verify_order_status()` 查询订单状态; 若未 `filled`, 调用 `_broker_reconcile_submitted_order()` 券商回查。

6. 记账: `_log_trade()` 写入 JSONL; `risk.record_trade(..., is_stoploss=True)` 更新风控状态; `risk.clear_position_state()` 清除持仓状态。

7. TG 通知: `_tg_trade_alert()` 或 `_tg_stoploss_alert()`。

8.4.7. Step7 买单执行

1. 刷新账户: 卖出后重新查询 `account_balance` 和 `stock_positions`。

2. 持仓数上限: `dynamic_max=compute_max_positions(refreshed_equity), clamp(3, 20)`。

3. 集中模式判断 (见 § 8.6.7)。

4. 遍历买入候选: 跳过已持仓、跳过超 `dynamic_max` 的候选、计算 `target_amount`、应用约束、风控校验。

5. 满仓 TG 告警: 若候选被 `max_positions` 拦截, 发送”满仓提示”消息。

6. 确认模式: 若启用, 买入前需 TG 确认。

7. 下单: 限价单 (`OrderType.L0`), 溢价系数按时段动态: 加仓/尾盘: +0.3%; 开盘: +1.0%; 盘中: +0.7%

8. 成交验证: 同卖出。

9. 记账: `_log_trade()+risk.record_trade()+_tg_trade_alert()`。

8.4.8. Step8 汇总推送

调用 `_tg_round_summary()` 发送本轮买卖汇总。

8.5. cmd_monitor 盘中监控循环

cmd_monitor 是一个长驻循环，被 cron 每 10 分钟启动一次：

while 交易时段活跃：

1. 判断是否为开盘/尾盘窗口 → force_refresh
2. 判断 LLM 冷却是否到期 → use_llm
3. 调用 cmd_trade(enable_llm=use_llm, force_refresh=force_rf, intraday_mode=True)
4. 累计本轮 buy/sell 到周期统计
5. 若连续错误 ≥ 3 → TG 告警； ≥ 10 → 自动退出
6. 每 3600s 发送一次周期报告
7. 检测脚本是否被更新（MD5 对比） → 热退出
8. sleep(interval) // 默认 1800s（30 分钟）

关键设计：-脚本热更新检测：通过比较启动时与当前的 trader.pyMD5，若代码被更新则自动退出，由 cron 下次启动新版本。-LLM 节流：盘中监控不每轮都调用 LLM（成本和延迟），而是按 15-30 分钟间隔触发。-错误熔断：连续异常不会无限重试，3 轮告警、10 轮停机。

8.6. 其余 CLI 命令速查

表37: CLI 命令速查

命令	等价执行深度	说明
scan	Step1-4	只做评分与信号生成， 不连券商、不下单 ，输出评分表格到 TG。用于盘前预览候选池。
signal	纯离线	不连券商终端，仅调用 run_quant_scoring() 输出评分文件到 data/signals/。用于本地回测或调试。
monitor	循环跑 trade	长驻循环（见 § 7.5），每隔 N 秒扫描一次，自动处理买卖。
status	只读查询	查询账户余额+持仓列表+今日已用预算/三桶计数， 无评分、无交易 。
order	单笔直达	手动单笔下市价/限价单，跳过全部评分与风控（仅保留基础额度检查），用于紧急补单或人工干预。

数据来源：东吴证券研究所绘制

9. 风控体系：风险约束的十道防线

9.1. 风控参数全景

当前生效的风控参数（直接读取 scripts/config.py 中的 RISK-LIMITS）：

表38：风控参数类别

类别	参数	值	含义
账户级	max_daily_loss_pct	5%	日亏达 5% 停止下单（含浮亏）
	max_total_loss_pct	15%	累亏达 15% 强制停机（人工介入）
	no_margin	True	禁止融资，仅用自有资金
	max_positions	动态 3-20	由 compute_max_positions (equity) 按账户规模算
	max_single_order_pct	0.30	单票下单金额上限 30% 权益
	max_leverage_etf_total_pct	0.50	杠杆 ETF 合计上限 50%
	max_daily_trades	动态 2-15/ 日	由 compute_daily_budget () 按 VIX/ 浮盈/ 事件日 每日重算
成长股	stop_loss_pct	4%	普通个股止损
	take_profit_pct	20%	普通个股止盈
	trailing_stop_pct	4%	移动止损
杠杆 ETF	leverage_stop_loss_pct	8%	杠杆品种止损
	leverage_take_profit_pct	28%	杠杆品种止盈
	leverage_trailing_stop_pct	5%	杠杆品种移动止损
小盘	smallcap_stop_loss_pct	6%	小盘止损
	smallcap_take_profit_pct	32%	小盘止盈
	smallcap_trailing_stop_pct	5%	小盘移动止损

数据来源：东吴证券研究所绘制

风控参数设置依据

- 止损 4%：在“保留容错空间”与“及时止损”之间取得平衡。过紧易受日内噪音误触（比如正常波动就止损了），过松则失去止损意义。

回测验证：测试了 2%、3%、4%、5%、6% 五个止损档位。4% 的夏普比率最高，胜率适中。

- 止盈 20%：匹配典型持仓周期（约 2-4 周）的合理目标。不是一涨就卖，但也不是贪婪到永远不卖。

回测验证：测试了 10%、15%、20%、25%、30% 五个止盈档位。20% 的综合表现最好。

- 杠杆 / 小盘品种阈值更宽：因为这些标的本身波动就比普通股票大。如果也给它们设 4% 止损，可能三天两头就被震出来了。

例子：TQQQ（3 倍做多纳斯达克）的日均波动约 3-5%。如果设 4% 止损，可能每 2-3 天就触发一次。设 8% 后，大约每 2-3 周触发一次，更合理。

9.2. 参数读取方式

风控参数直接读取 scripts/config.py 中的 RISK_LIMITS 字典。如需调整参数，直接修改 RISK_LIMITS 并重启进程即可生效。

修改示例：

假设用户拟将止损从 4% 调整为 5%：1. 打开 `~/openclaw-investment/workspace/scripts/config.py`；2. 找到 RISK_LIMITS 字典；3. 将 `"stop-loss-pct": 0.04` 改为 `"stop-loss-pct": 0.05`；4. 保存；5. 重启进程：`launchctl kickstart -k gui/$UID/ai.openclaw.investment`；6. 新参数生效。

警告：修改参数前，建议在研究框架里先回测，确认改动不会恶化表现。

9.3. 日亏上限含浮亏

`max-daily-loss-pct = 5%` 的计算口径：

`today-pnl = 已实现盈亏 + 浮动盈亏`

`if today-pnl / total-equity-at-open <= -5%:`

停止新买入 (BUY) + TG 告警；止损/止盈/移动止损 (SELL) 继续执行

该 5% 并非单笔交易实际卖出亏损，而是当日账户层面的账面总亏损，包括已实现亏损与尚未卖出的持仓浮亏。

将浮亏纳入计算口径可确保单日下跌幅度被严格锁定在 -5% 以内，避免浮亏扩大时继续加仓导致风险叠加。

实际场景：假设开盘时持有 10000 美元股票。至中午，持仓下跌 3% (浮亏 300 美元)，同时早盘止损一只股票亏损 200 美元。总亏损 = 500 美元 = 5%。此时系统会停止新增买入，即便存在强买入信号，当日也不再开新仓；但现有持仓的止损、止盈与移动止损仍继续监控并执行。

日亏上限设为 5% 的原因：

表39：系统局限性

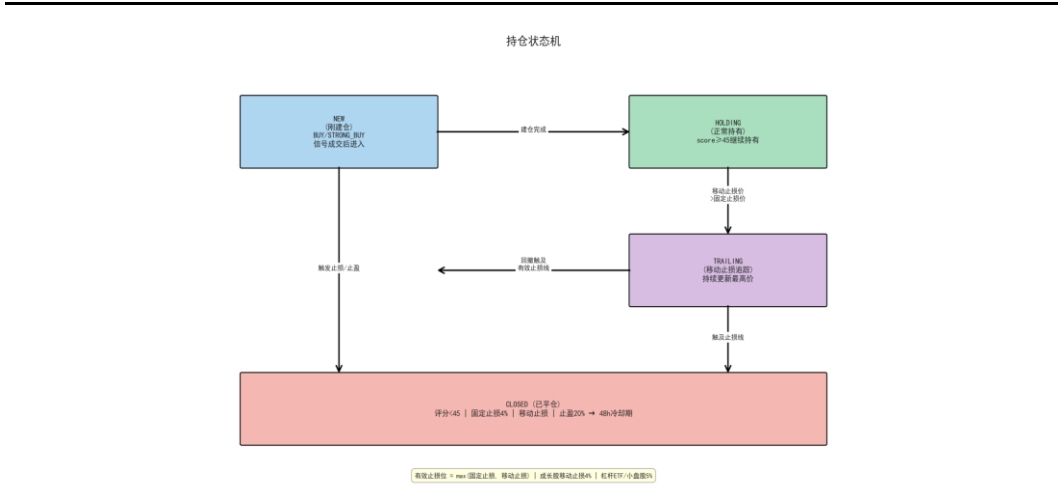
日亏上限	夏普	年化收益	最大回撤	问题
3%	1.30	22%	-15%	太紧，正常波动就可能触发，交易机会损失太多
5%	1.24	30%	-19%	最优平衡点
10%	1.05	38%	-35%	太松，一天亏 10% 对心理冲击太大

数据来源：东吴证券研究所绘制

9.4. 持仓生命周期管理

每笔持仓维护完整的生命周期状态机: -open→held→trailing (触发移动止损追踪) →closing→closed

图5: 持仓状态示意图



数据来源: 东吴证券研究所绘制

状态持久化至 SQLite, 进程重启后可自动恢复。每次状态切换均通过 TG 通知。

避免“判断已持仓实际未成交”或“判断已卖出实际挂单未成交”等状态不一致问题。

状态不一致的灾难场景:

若缺少状态机管理, 可能出现以下问题: 1. 系统认为 AAPL 已卖出 (因已提交卖单); 2. 实际卖单未成交 (价格下跌过快, 缺乏成交); 3. 系统误以为已无 AAPL 持仓, 又生成 AAPL 买入信号; 4. 结果: 本应清仓的交易变为加仓, 造成状态错配风险。

有了状态机: 1. 系统下卖单 → 状态 = closing 2. 查询订单状态 → 未成交 3. 状态保持 closing, 不会发新的买入信号 4. 直到确认成交 → 状态 = closed

9.5. 止损卖出后冷却, 而非永久纳入黑名单

止损卖出后, 标的进入 48 小时冷却期 (当天不买入), 48 小时后可重新进入评分与买入流程。

不采用“一次止损永久拉黑”策略——优质标的也可能经历一次技术性回撤。

实际案例:

1月1日：买入 NVDA@500

1月5日：NVDA 跌到 480 (-4%) → 触发止损卖出

1月5日-1月7日：48小时冷却期，NVDA 即使评分再高也不买

1月8日：冷却期结束。如果 NVDA 的 final_score ≥ 65，可能重新买入

1月8日实际：NVDA 财报超预期，涨到 550。系统重新买入@550。

若永久纳入黑名单，则可能错过 NVDA 财报后的后续涨幅。

如果无冷却期：1月5日止损后，1月6日可能又买回来（如果评分仍然高），但可能继续跌。

48小时冷却期用于在“不过早重新建仓”和“不错过后续修复”之间取得平衡。

9.6. 三桶预算与杠杆 ETF 合计校验

本节与 SOUL.md § OrderExecution/ § DailyTradeBudget、IDENTITY.md § 交易脚本单一真相源强对齐。任何改动必须两边同步。

9.6.1. 三桶预算模型

_trader.py 中的下单按来源分为三个不互通的桶：

表40：三桶预算模型

桶	来源	是否受日预算 N 约束	优先级	解释
instruction	你在对话里显式指令（“买 NVDA50 股”、“清仓 TSLA”）	✗ 不受	最高（仅红线可挡）	用户明确指令，系统按规则执行
spontaneous	系统主动触发的 setup 下单	☑ 仅此桶受 N 约束	中（打到 N 即停）	“系统自己觉得该买/卖”
risk	止损/止盈/移动止损/强制平仓	✗ 不受	强制（哪怕 N 已满）	“用于风险控制，不计入次数”

数据来源：东吴证券研究所绘制

划分三桶的原因：防止系统自发交易消耗全部交易次数，导致止损、止盈等风控指令无法执行。止损与止盈属于风险控制动作，应随时具备执行权限，不受普通日内交易次数限制。

实际场景：

假设 $N=5$ (今天最多 5 笔 spontaneous 交易):

表41: 三桶预算模型实例

时间	交易	桶	N 计数	说明
09:30	买入 AAPL	spontaneous	1/5	正常计入
09:45	买入 MSFT	spontaneous	2/5	正常计入
10:00	止损 TSLA	risk	2/5	不计入, 保命单
10:30	买入 NVDA	spontaneous	3/5	正常计入
11:00	买入 AMD	spontaneous	4/5	正常计入
11:30	买入 INTC	spontaneous	5/5	正常计入
12:00	买入 QQQ	spontaneous	5/5	拒绝! N 已用完
12:30	止盈 AAPL	risk	5/5	执行! risk 桶不受限

数据来源: 东吴证券研究所绘制

9.6.2. RiskManager.can-trade() 十道红线

每笔 BUY/SELL 在提交券商前, 必须依次通过 risk-manager.py 中 can-trade() 的十道校验 (任意一道失败即拒绝, 返回原因):

仅 spontaneous 桶受约束 (第 1、9 道):

1. 三桶 N 约束—spontaneous 桶当日已用 \geq daily-budget \rightarrow 拒绝 (“自发交易预算 $N=X$ 已用完”)

2. 市场笔数 fallback—当日该市场总成交 \geq max-daily-trades \rightarrow 拒绝 (兜底防线)

所有桶均受约束 (第 3-8、10 道):

3. 日亏熔断—today_pnl/total_equity $<$ -max-daily-loss_pct (默认-5%) \rightarrow 停止新买入 (BUY); 止损/止盈/移动止损 (SELL) 继续执行

4. 累亏熔断—(total-equity-baseline)/baseline $<$ -max-total-loss_pct (默认-15%) \rightarrow 强制停机

BUY 专属 (第 5-8 道):

5. 单笔上限—amount $>$ max(max_single_order_floor_usd=\$2000, min(equity \times 30%, cap=\$50000)) \rightarrow 拒绝

6. 单笔下限—amount $<$ max(min_single_order_floor_usd=\$200, equity \times 8%) (试仓/加仓为 max(\$100, 4%), 集中模式为 \$500) \rightarrow 拒绝

7. 单票仓位上限—position_pct \geq max_position_pct (默认 30%, 当前代码中此项存在但 trader 侧未传 position_pct, 实际未生效) \rightarrow 拒绝

8. 杠杆 ETF 合计 — code 属于 LEVERAGE-ETF_CODES 且 $\text{current-leverage-pct} + \text{new-pct} > 50\%$ → 拒绝 (当前 trader 侧未传入 $\text{current-leverage-etf-pct}$, 实际未生效)

9. 现金户检查 — $\text{no-margin} = \text{True}$ 且 $\text{amount} > \text{available-cash} \times 0.99$ → 拒绝

10. 冷却期&买回禁令 (仅 spontaneous) —

交易间隔冷却: 距上次交易 $< \text{cooldown-minutes}$ (默认 10 分钟) → 拒绝 (防过度交易)

止损后冷却 (§ 8.5): 止损卖出后 48 小时内不回购 → 拒绝 (防抄底半山腰; 与下面普通买回禁令独立并存)

普通买回禁令: 距上次普通卖出 $< \text{avoid-hours}$ (高分 2h/中分 6h/低分 24h) → 拒绝 (防频繁进出同一只票)

执行细节: $\text{can-trade}()$ 内部先调用 $\text{self._reset-if-new-day}()$ 确保跨日状态刷新; 所有拒绝理由均通过 $\text{-log}()$ 写入日志并透传给 TG 通知。

9.6.3. 日预算 N 动态计算

$\text{compute-daily-budget}()$ 每日盘前推送时计算一次, 当日生效:

$\text{base} = 5$

波动系数 $\times = \{\text{VIX} < 15: 0.6, 15-20: 1.0, 20-25: 1.3, > 25: 1.6\}$

持仓系数 $\times = \{\text{账户当日浮盈} > +2\%: 1.2, -2\% \sim +2\%: 1.0, < -2\%: 0.7\}$

事件加成 $+ = \{\text{财报日}: +2, \text{议息日}: +3, \text{CPI/PPI/NFP 日}: +2\}$

$N = \text{clamp}(\text{round}(\text{base} \times \text{波动} \times \text{持仓} + \text{事件}), 2, 15)$

该机制根据三个因素动态调整每日交易次数: - 市场波动 (VIX): 波动较大时增加交易次数, 波动较小时降低交易次数以减少噪音 - 账户状态 (浮盈): 盈利状态较好时适度增加交易, 亏损状态下减少交易 - 事件日: 财报日、议息日等高波动事件日适当增加交易机会。

示例: - 平淡日 (VIX13, 昨平盘, 无事件): $5 \times 0.6 \times 1.0 = 3$ 笔 - 财报周 (VIX20, 昨+2.5%, META 财报): $5 \times 1.0 \times 1.2 + 2 = 8$ 笔 - 黑天鹅 (VIX32, 昨-3%, 议息日): $5 \times 1.6 \times 0.7 + 3 = 9$ 笔

CLI 未显式传 $--budget$ 时自动调用 $\text{compute-daily-budget}()$; 传了则以传入值为准 (用于 dry-run 或手动演练)。

杠杆 ETF 合计 $\leq 50\%$

RISK-LIMITS 中已定义 `max_leverage_etf_total_pct=0.50`，且 LEVERAGE-ETF-CODES 集合已维护（含 SOXL、TQQQ、TECL、FNGU、TNA 等）。但当前交易执行层尚未在下单前调用该校验逻辑。

预期实现:

```
leverage_etfs={TQQQ, SOXL, TECL, FNGU, TNA, SQQQ, SOXS, ...} #所有 2x/3x 均计入
```

```
current_leverage_total_pct=Σ(持仓中属于 leverage_etfs 的标的权重)
```

```
ifcandidate.tickerinleverage_etfs:
```

```
projected=current_leverage_total_pct+candidate.target_weight
```

```
ifprojected>0.50:
```

```
→拒绝下单(返回 RiskBlocked.LEVERAGE-ETF-CAP-EXCEEDED)
```

```
→TG 通知: "拒绝: 此单将使杠杆 ETF 合计达 XX%, 超 50%红线"
```

触发拒绝时不自动降仓执行（不改用户意图），让用户自己决定是换标的还是先减仓。

笔数记账

每笔成交后，在当日 `~/openclaw-investment/workspace/memory/YYYY-MM-DD.md` 追加一行（由 `_trader.py` 成交回报时触发 `telegram_notifier.write_trade_memo()`）:

```
自发笔数: X/N|HH:MM|<ticker><方向><股数>
```

```
指令单: HH:MM|<ticker>... ←不增加 X
```

```
止损单: HH:MM|<ticker>... ←不增加 X
```

“今天打了几笔自发单”是可随时秒答的硬要求。

9.6.4. 持仓数自适应

由 `compute_max_positions(equity)` 每次 signal 调用时重算:

```
max_positions=clamp(3, int(equity/pilot_floor/3), 20)
```

自动换仓阈值自适应

原三个阈值 `rotation_min_score_diff=20` / `rotation_weakest_score_max=65` / `rotation_max_per_day=2` 都是静态，对“平静 vs 波动”“赚钱 vs 亏钱”“平淡日 vs 财报周”无差别对待。改造后全部动态化。

分差门槛（候选分-最弱持仓分 \geq diff_th才能换）：

$$\text{diff_th} = \text{clamp}(10, 25 - (\text{VIX} - 15), 25)$$

这到底在算什么？VIX越高（市场越波动），换仓门槛越低（允许更频繁换仓）；VIX越低（市场越平静），换仓门槛越高（少折腾）。

表42: 分差门槛

VIX	diff_th	含义
<15（平静）	25	更严，少换（市场无方向时不乱动）
20	20	中性
25	15	放松（波动大，容忍更多轮换）
≥ 30 （恐慌）	10（封底）	最松，但其他护栏（盈亏/间隔90分钟/剩余预算）兜底

数据来源：东吴证券研究所绘制

最弱评分门槛（最弱持仓 \leq weak_th才能被换出）：

$$\text{weak_th} = \text{clamp}(60, 70 + \text{账户当日浮盈} \times 100, 80)$$

表43: 最弱评分门槛

账户当日浮盈	weak_th	含义
-5%	65	更严（亏钱时少折腾，防追涨杀跌）
0%	70	中性
+5%	75	放松（顺风时主动优化组合）
+10%+	80（封顶）	最松（大赚时积极换强）

数据来源：东吴证券研究所绘制

每日换仓次数上限（=消耗2N笔交易预算，所以卡在总预算一半）：

$$\text{rotation_max_per_day} = \text{clamp}(1, \text{int}(\text{daily_budget_N} \times 0.5), 6)$$

表44: 每日换仓次数上限

daily_budgetN	rotation_max_per_day
3（平淡日）	1
5	2
8（财报周）	4
15（黑天鹅）	6（封顶）

数据来源：东吴证券研究所绘制

VIX获取：优先读 `quant_system.events.get_current_vix()`，失败兜底 `VIX=15`（中性）。账户浮盈：（当前 `refreshed_equity - RiskManager._session_start_equity`）/`_session_start_equity`。

10. 盘中监控与守护

10.1. 监控架构与 InvestmentClaw 的交互角色

InvestmentClaw 并非黑箱系统。其通过 OpenClaw 框架提供的三层可观测性设计，使用户能够持续掌握系统状态：

表45: InvestmentClaw 三层可观测性设计

观测层	载体	你看到什么	类比
实时推送	TelegramBot	买卖成交、风险预警、每日归因	InvestmentClaw 给你发微信
结构化日志	SQLite+JSONL	完整交易流水、持仓状态、评分历史	医院的电子病历系统
运行时日志	cron-logs/	每个 cron 任务的原始输出（含报错堆栈）	手术室的监控录像

数据来源：东吴证券研究所绘制

InvestmentClaw 的人格体现在通知里：TG 消息不是冷冰冰的“BUY AAPL”，而是带有 emoji、上下文和决策理由的完整汇报。例如：

```
[BUY]AAPLx10@182.50

理由:final_score=81.3(growth引擎78+AI82+反向60)

仓位:AAPL12%→18%|现金$8,420

日预算:2/5笔已用
```

这与 InvestmentClaw “先查数据、再说话”的人格一致——它汇报时自带完整上下文，让你一眼看懂”为什么买、买了多少、还剩多少预算”。

10.2. 三层可观测性的互补关系

三层观测不是冗余，而是互补的：

- TG 实时推送：适合“用户现在想知道发生了什么”——秒级、可读、带 emoji
- SQLite + JSONL：适合“用户想查某只股票的历史”——结构化、可查询、长期存档
- cron-logs：适合“系统好像出问题了，帮用户排查”——原始、完整、含调试信息

实际使用场景： - 早上醒来刷 TG → 看到“今日信号：3 BUY / 1 SELL” → 知

道 InvestmentClaw 今天动了 - 想查某笔交易的详情 → 打开 SQLite → SELECT * FROM trades WHERE ticker='AAPL' - 发现某笔交易异常 → 查 cron-signal.log → 看到当时的完整决策链

10.3. InvestmentClaw 的“主动模式”

InvestmentClaw 不只是被动等待 cron 触发——它会在以下场景主动向你推送:

表46: InvestmentClaw 主动模式

场景	推送内容	示例
盘前 (21:00SGT)	今日方向+重点 setup	“今日 VIX22, 偏防御。关注 AAPL 财报后回调 setup”
盘中 (每 10 分钟)	持仓异动+新信号	“NVDA 触发移动止损, 已卖出”
盘后 (04:30SGT)	每日归因+复盘	“今日+1.2%, 贡献最大: AAPL+\$85”
风险预警 (实时)	接近红线时提前告警	“⚠️ 日亏达 4.2%, 接近-5%熔断线”
setup 狙击 (机会驱动)	发现高置信度交易机会	“🎯 TSLA 放量突破, final_score=88, 建议关注”

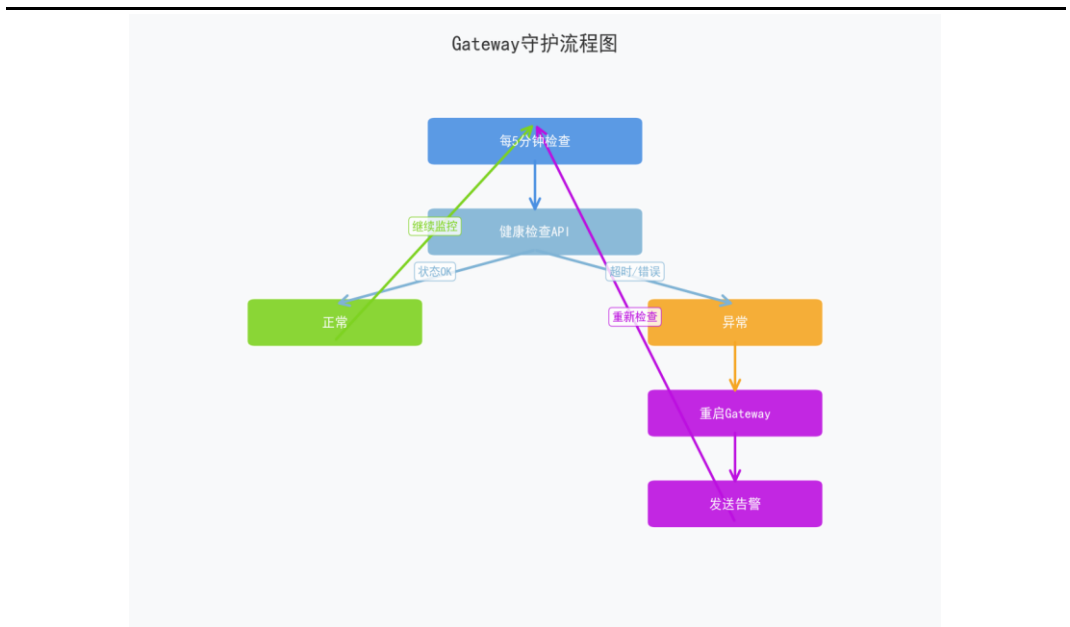
数据来源: 东吴证券研究所绘制

注意: “setup 狙击”属于 InvestmentClaw 的自发单 (spontaneous), 受日预算 N 约束。达到 N 后停止, 除非触发止损或用户手动接管。

这与 InvestmentClaw“主动出击, 不等指令”的人格一致——它是猎手, 不是 waiter。

10.4. 监控架构

图6: Gateway 守护流程图



数据来源: 东吴证券研究所绘制

Gateway 是一个长驻 Python 进程, 负责维持与券商终端服务器的 WebSocket 长连接, 接收实时报价与成交推送。Gateway 异常时, 所有依赖实时数据的功能 (报价、成

交回报、盘中止损)将失效。

10.5. Gateway 守护

每 5 分钟由 cron 调用:

1. 通过 pgrep 检查 Gateway 进程是否存活
2. 不存活则通过 nohup 重启 Gateway
3. 冷却期内不重复发送告警 (参见 8.3)

10.6. 告警节流与冷却

避免告警风暴的两层机制: - 节流: 同一类型告警 (如 Gateway 异常) 在 20 分钟窗口内仅发送一次 - 冷却: Gateway 重启成功后 20 分钟内不再触发健康检测 (避免启动过程中被误判)

极端情况下 (如 Gateway 在一分钟内连续异常 3 次), 告警数量被压缩至 1 条以内。

需要节流的原因: 若 Gateway 连续异常, 守护进程会持续尝试重启。若缺少节流机制, Telegram 可能短时间内收到大量告警。20 分钟节流可确保同类告警不频繁打扰用户。

10.7. Profile 感知

监控脚本启动时, 依据 \$OPENCLAW_WORKSPACE_DIR 自动选择对应配置: - /Users/liu/.openclaw-investment/workspace → 生产 profile (InvestmentClaw 席位) - /Users/liu/.openclaw/workspace-demo → 演示 profile - 未设置 → 拒绝启动 (防止误用错误的 API 凭证)

10.8. Gateway 热重载

OpenClaw4.2 起支持 Gateway 配置热重载: 修改 gateway.yml 后无需重启进程即可生效, 适用于: - 订阅标的列表 - 心跳间隔 - 日志级别

但修改券商终端凭证仍需重启进程。

11. Telegram 通知

11.1. 通道设计

系统事件通过 TelegramBot 推送, 按优先级划分为三个通道: - 交易通道 (高优先级): 下单、止损、大额事件 - 监控通道 (中优先级): Gateway 异常/恢复、选股异常、数据源故障 - 报告通道 (低优先级): 每日归因、周度 IC 监控、周月总结

用户可按通道选择性静音。

11.2. 速率控制

TelegramBotAPI 限流 30 条/秒。系统内部的控制策略：-每秒最多发送 10 条（预留余量）-同一消息 60 秒内去重（避免重复推送）-发送失败的消息写入 data/tg-retry-queue.jsonl，下次运行时补发

11.3. 消息类型总览

交易类：-买入成交：[BUY]AAPLx200\$185.3 成交总额\$3706-卖出成交：[SELL]AAPLx200\$190.8 盈利\$110(+2.97%)-止损触发：[STOP-LOSS]TSLA@\$220.5 止损4%(\$880)-止盈触发：[TAKE-PROFIT]NVDA@\$920.3 止盈 20%(\$1840)

监控类：-⚠ Gateway 异常告警/恢复通知-⚠ 选股池数据异常(Finviz 返回<20 只)
-⚠ LLM 熔断触发/恢复-⚠ 日亏达 3%预警

报告类（每日收盘后自动发送）：-持仓快照（持股、浮盈亏、占比）-每日归因报告：当日盈亏分解（ α 选股贡献/ β 市场贡献/执行成本）-周度总结（周日）-因子 IC 监控报告（每周日 17:00，检出因子衰减时会单独告警）

12. 文件与路径

12.1. 核心目录结构

```

~/openclaw-investment/workspace/#运行时工作区

├── scripts/ | ├── _trader.py#主交易程序 (cron 直接调用) | ├── universe_builder.py#选股池构建 | ├── engine.py#
信号生成+LLM 调用 | ├── scoring.py#三引擎评分 | ├── telegram_notifier.py#TG 通知 | ├── daily_attribution.py#每日归因分析
(cron04:30) | ├── factor_ic_monitor.py#周度因子 IC 监控 (cron 周日 17:00) | ├── openclaw_wakeup.sh#Gateway 守护 | ├──
gateway.yml#配置 | ├── data/ | ├── universe_latest.jsonl#最新候选池 (原子写) | ├── scores_latest.jsonl#最新评分表 | ├──
signals_latest.jsonl#最新信号表 | ├── holdings.sqlite#持仓状态 (持久化) | ├── trades.sqlite#交易流水 (SQLite) | ├──
trades.jsonl#交易流水 (JSONL 备份) | ├── blacklist.txt#永不买黑名单 | ├── cron_logs/#cron 运行日志 | ├──
cron_signal.log | ├── cron_monitor.log | ├── cron_gateway.log | ├── cron_attribution.log | ├──
cron_ic_monitor.log | ├── factor_ic_history/#周度 IC 快照 | ├── attribution/#每日归因报告 | ├── YYYY-MM-DD.jsonl#
docs/ | ├── TRADING_SYSTEM.md#原文档 (技术版) | ├── TRADING_SYSTEM_EXPLAINED.md#本文档 (面向零基础) | ├── ROLLOUT_GUIDE.md#
上线操作手册

~/quant-stock-picker/#策略代码仓库

├── quant_system/#生产核心代码 (scriptsimport 自此) | ├── config.py | ├── data.py | ├── scoring.py | ├──
engine.py | ├── universe_builder.py | ├── backtest.py#历史回测引擎 | ├── factor_analysis.py#因子 IC+动态权重 | ├──

```

```
portfolio_optimizer.py | └──v2/#研究框架(独立于生产) | └──factors_v2.py#32 因子库 | └──regime.py#市场状态分类 | └──
└──backtest_adapter.py#新因子→回测引擎 | └──backtest_engine_v2.py#增强回测引擎 | └──run_baseline.py#基线回测 | └──
run_variants_*.py# 参数实验 | └──run_cpcv.py#CPCV 交叉验证 | └──run_optuna.py#Bayesian 参数调优 | └──
run_ic_analysis.py#因子 IC 分析 | └──backtest_data/#历史数据(研究用) | └──prices/#280 只股票×3 年日线 | └──
fundamentals/#基本面快照 | └──backtest_outputs/#回测结果 | └──baseline/BASELINE.REPORT.md | └──factor_ic/ic_report.md | └──
└──optuna/best_params.json
```

12.2. 工作区路径解析

所有脚本遵循统一的路径解析规则：

1. 若设置了环境变量 OPENCLAW_WORKSPACE_DIR，以其为准
2. 否则回退至默认路径 `~/openclaw-investment/workspace/`
3. cron 必须显式设置环境变量（launchd 默认不继承 shell 环境变量）

该规则解决了”手动运行脚本与 cron 运行脚本行为不一致”的常见问题

13. Cron 调度

13.1. 一天的时间表（北京时间）

表47: Cron 调度一天时间表

时间	任务	对应美股时间
21:00	盘前信号生成（开盘前 30 分钟）	09:00ET
21:30	开盘执行买入	09:30ET（开盘）
21:00-05:00（每 10 分钟）	盘中监控：浮盈亏+止损触发检测	盘中
01:00	午间持仓健康检查	13:00ET
04:00	收盘前 30 分钟调节（避免尾盘冲击）	15:30ET
04:00	美股收盘	16:00ET
04:30	每日归因分析（收盘后立即）	收盘后 30 分钟
05:05	次日信号预研+晚间总结 TG	-
周日 17:00	周度因子 IC 监控	周末空闲时段
每 5 分钟	Gateway 健康守护	全时段

数据来源：东吴证券研究所绘制

为什么用 SGT 时区？因为用户在新加坡。美股开盘是美东时间 09:30，对应新加坡时间 21:30。所以系统在新加坡时间晚上运行，白天用户可以查看报告。

13.2. 完整 crontab

```
#交易时段 cron (OPENCLAW_WORKSPACE_DIR 指向投资侧真实 profile)
```

```

021**1-5/usr/bin/envOPENCLAW_WORKSPACE_DIR="$HOME/.openclaw-
investment/workspace"python$HOME/.openclaw-investment/workspace/scripts/
    _trader.py--modesignal>>$HOME/.openclaw-
investment/workspace/data/cron-logs/cron-signal.log2>&1

*/1021-23**1-5/usr/bin/envOPENCLAW_WORKSPACE_DIR="$HOME/.openclaw-
investment/workspace"python3$HOME/.openclaw-investment/workspace/scripts/
    _trader.py--modemonitor>>$HOME/.openclaw-
investment/workspace/data/cron-logs/cron-monitor.log2>&1

*/100-4**2-6/usr/bin/envOPENCLAW_WORKSPACE_DIR="$HOME/.openclaw-
investment/workspace"python3$HOME/.openclaw-investment/workspace/scripts/_trader.py--
modemonitor>>$HOME/.openclaw-investment/workspace/data/cron-logs/cron-monitor.log2>&1

#Gateway 守护 (全时段每 5 分钟)

*/5****$HOME/.openclaw-investment/workspace/scripts/openclaw-wakeup.sh>>$HOME/.openclaw-
investment/workspace/data/cron-logs/cron-gateway.log2>&1

#每日归因分析 (美股收盘后 30 分钟=04:30SGT)

304**2-6/usr/bin/envOPENCLAW_WORKSPACE_DIR="$HOME/.openclaw-
investment/workspace"python3$HOME/.openclaw-
investment/workspace/scripts/daily-attribution.py>>$HOME/.openclaw-
investment/workspace/data/cron-logs/cron-attribution.log2>&1

#周度因子 IC 衰减监控 (周日 17:00SGT)

017**0/usr/bin/envOPENCLAW_WORKSPACE_DIR="$HOME/.openclaw-
investment/workspace"python3$HOME/.openclaw-
investment/workspace/scripts/factor_ic_monitor.py>>$HOME/.openclaw-
investment/workspace/data/cron-logs/cron_ic_monitor.log2>&1

#每日 TG 总结 (05:05SGT)

55**2-6/usr/bin/envOPENCLAW_WORKSPACE_DIR="$HOME/.openclaw-
investment/workspace"python3$HOME/.openclaw-
investment/workspace/scripts/daily-summary.py>>$HOME/.openclaw-
investment/workspace/data/cron-logs/cron-summary.log2>&1。

```

13.3. 调度设计依据

周一至周五 21:00 启动选股: 覆盖美股开盘前 (ET09:00), 预留 30 分钟用于 LLM

判别与挂单准备

盘中每 10 分钟监控：确保止损响应延迟 ≤ 10 分钟，对日内波动响应充分

归因分析 04:30 执行：美股 04:00SGT 收盘（标准时间），30 分钟后数据已稳定，适合计算当日 α / β 分解

IC 监控安排在周日：交易时段之外执行 CPU 密集型任务，避免影响下单响应

Gateway 守护每 5 分钟：Gateway 异常后最多 5 分钟内被拉起，影响面可控

weekday1-5 与 2-6 的差异：选股任务对应”周一至周五的当晚”（触发日），监控任务对应”周二至周六的凌晨”（美股交易日次日）。美股交易日为周一至周五，SGT 时区下对应”SGT 周一夜至周六凌晨”。

14. 研究框架：因子、回测、调参

策略迭代依赖一套完整的研究框架（quant-system/v2/），生产配置均为研究结论的落地，确保改动有数据支撑。

14.1. 研究流程

历史数据（3 年 \times 280 股）

因子计算（32 因子，6 大类）

IC 分析（识别有效因子、剔除无效因子）

策略回测（BacktestEngineV2）

CPCV 检验（5-fold 过拟合防护）

Optuna 调参（Bayesian 寻找最优参数）

生产配置更新 + 灰度上线

每日归因 + 周度 IC 监控（持续观测）

研究流程可概括为：先收集历史数据，再计算因子，随后筛选有效因子（IC 分析）、进行回测、开展 CPCV 稳健性检验、通过 Optuna 优化参数，最终形成生产配置。上线后，系统继续通过归因与 IC 监控跟踪策略表现。

14.2. 32 因子库

分 6 大类：

表48: 32 因子类别

类别	代表因子	说明
动量 (8 个)	mom_5d/mom_20d/mom_60d/mom_120d/risk_adj_mom/...	不同时间尺度的价格趋势
技术 (7 个)	ma20_slope/above_ma20/above_ma60/breakout_ma60/rsi/atr_pct/...	均线、RSI、ATR 等技术指标
质量 (5 个)	ROE/ROA/毛利率/FCF 转化率/负债率	公司经营质量
价值 (4 个)	PE/PB/EV_EBITDA/FCF_yield	估值水平
成长 (4 个)	营收 YoY/EPSYoY/营收 QoQ/EPSQoQ	业绩增长速度
情绪 (4 个)	分析师评级变化 / 做空比率变化 / pv_divergence/recover_from_20d_low	市场情绪信号

数据来源: 东吴证券研究所绘制

14.3. 3 年因子 IC 排名

全 3 年历史 (2023-03~2026-04) 的 IC 排名:

表49: 3 年因子 IC 排名

因子	IC	IR	正 IC 天数占比	评价
ma20_slope	+0.062	0.51	76%	最强因子, 生产首选
risk_adj_mom	+0.058	0.51	74%	强
mom_20d	+0.054	0.44	71%	强
mom_120d	+0.045	—	68%	中强
above_ma60	+0.033	—	64%	中
above_ma20	+0.028	—	62%	中
recover_from_20d_low	+0.024	—	59%	弱有效
mom_60d	+0.030	—	62%	中
atr_pct	+0.010	—	52%	无效 (IC <0.02)
pv_divergence	+0.007	—	51%	无效
mom_5d	-0.005	—	49%	反转因子 (IC 为负)
breakout_ma60	-0.014	—	47%	顶部信号 (突破后多数出现回落)

数据来源: 东吴证券研究所绘制

关键结论:

中长期动量 (20d/60d/120d) 表现为正 α

短期动量 (5d) 呈反转特性, 与一般直觉相反

单一的 MA60 突破信号表现为顶部特征

生产评分已按上述 IC 表调整权重： $|IC| < 0.02$ 的因子权重归零，强因子权重与 IC 大小成正比。

14.4. 近 90 天 vs 历史 3 年 IC 对比

周度 IC 监控的核心用途：检测因子是否仍然有效。最新快照（2026-04）：

表50：近 90 天 vs 历史 3 年 IC 对比

因子	历史 IC	近 90 天 IC	健康度
ma20_slope	+0.047	+0.087	☑ 翻倍
mom_120d	+0.047	+0.093	☑ +98%
risk_adj_mom	+0.045	+0.098	☑ +120%
mom_20d	+0.042	+0.088	☑ +110%
above_ma60	+0.033	+0.082	☑ +150%
mom_60d	+0.030	+0.087	☑ +190%

数据来源：东吴证券研究所绘制

所有强因子近期 IC 显著提升，表明：-因子仍然有效，生产策略的底层逻辑健康-当前市场呈现较强的趋势性特征，动量类策略表现良好-若某因子连续 4 周 $|IC| < 0.02$ ，factor_ic_monitor.py 将发送 TG 告警并建议降权

14.5. 回测引擎

quant_system/v2/backtest_engine_v2.py 支持以下特性：-任意调仓频率（日/周/双周/月）-止损、止盈、移动止损-滑点与手续费建模-波动率过滤-IC 驱动的动态因子权重-分段统计（牛熊市分别计算）

单次完整 3 年回测耗时约 30 秒。

14.6. 回测核心结论

当前生产配置在 3 年历史回测中的表现（vs 保守基线对照）：

表51：回测表现

指标	保守基线	生产配置	改善
夏普比率	0.76	1.24	+63%
年化收益	37%	30%	-19%
最大回撤	-50%	-19%	减 62%
胜率	39%	48%	+9 点
盈亏比	2.05	2.52	+23%
换手率（交易数）	895	119	-87%

数据来源：东吴证券研究所绘制

对指数基准的客观对比：同期 SPY 涨幅+342%，策略+89%，绝对收益确实低于指数。但风险调整收益（策略夏普 1.24vsSPY 夏普 1.27）接近持平，且最大回撤-19%远低于

SPY 的-37%。

策略定位：不以跑赢指数为目标，而是以较小的回撤获取接近指数水平的风险调整收益。适合小资金账户、追求稳定性、不愿承担-37%量级回撤的投资者。

14.7. CPCV 交叉验证

将 3 年历史切分为 5 段，每段单独回测：

表52: CPCV 交叉验证

Fold	时段	夏普	回撤
1	2023-06-12	0.65	-5.2%
2	2024-01-07	1.52	-4.8%
3	2024-07-2025-02	2.26	-3.1%
4	2025-02-09	-0.42	-6.3%
5	2025-09-2026-04	-1.18	-5.8%

数据来源：东吴证券研究所绘制

观察要点：-最差回撤仅-6.3%（基线对照-27.4%），风控表现稳定-近两段（2025-02 之后）夏普转负，但近 90 天 IC 分析表明因子本身仍然健康（见 13.4）-结论：近期可优化方向在策略合成层（过滤逻辑与阈值），无需调整底层因子

14.8. OptunaBayesian 调参

对以下 9 个超参数执行 80 trials Bayesian 搜索（最近 18 个月数据作为训练集）：
 -rebalance-frequency {weekly, biweekly, monthly} -max-positions {5, 6, 7, 8} -
 buy-score-threshold [50, 70] -sell-score-threshold [35, 50] -
 stop-loss-pct [0.03, 0.06] -take-profit-pct [0.12, 0.25] -
 max-annual-vol [0.40, 0.80] -max-single-weight [0.3, 0.5] -
 require-dual-ma {True, False}

Top3 稳健参数（train 与 val 两段均表现良好）：

表53: TOP3 稳健参数

Trial	trainSharpe	valSharpe	关键参数
71	2.34	2.61	weekly/vol<0.55/buy=56/s1=5%/tp=14%
14	1.14	3.49	monthly/vol<0.60/buy=55/s1=3%/tp=16%
45	1.29	2.53	monthly/pos=5/buy=55/s1=4%/tp=16%

数据来源：东吴证券研究所绘制

生产最终采用 IC-driven 变种（周调仓+波动过滤 0.60+IC 权重+止损 4%+止盈 20%），综合权衡了调参结果、因子 IC 健康度以及小资金账户的执行摩擦。

15. 总结：从 AI 辅助到 AI 自主：本系统的核心价值与启示

15.1. 全文回顾

本报告以 OpenClaw 框架为底座，完整构建了一套"选股 — 评分 — 决策 — 执行 — 风控 — 观测"六层闭环的美股日频自动化交易系统 InvestmentClaw。

15.1.1. 架构层面（第 2 章）

六层分层设计明确了各层职责与接口契约。调度层通过 Cron 实现交易日全时段定时触发；选股层（第 3 章）采用 Finviz 四策略筛选、五类异动扫描、SEED-TICKERS 种子池与杠杆 ETF 专属池的四路入池机制，辅以原子写入保护与最小阈值看守，确保候选池即便在数据源故障时仍可依靠种子保底维持运行；评分层（第 4 章）基于"成长股看基本面、杠杆 ETF 看大盘趋势、小盘股看情绪动量"的差异化定价逻辑，构建了 growth / leverage / smallcap 三套专用引擎，并通过评分拉伸与后置调节规则在"区分度"与"稳健性"之间取得平衡。

15.1.2. AI 多模型判别（第 5 章）

是本系统最具特色的环节。系统将评分 Top20 候选送入 MiniMax-M2.7 与 GLM-5.1 两个推理模型进行对等双跑与交叉验证——模型方向一致且差值 ≤ 3 则增强，方向分歧则取保守值，单侧失败降权，双侧失败 fallback 至中性。围绕 LLM 幻觉风险，系统还配套了结构化 Prompt 约束、数据一致性校验、熔断冷却与速率控制等工程机制。final_score 的权重结构（量化 60% + AI 20% + 反向 20%）确立了"量化压舱、AI 增量"的清晰分工。

15.1.3. 从信号到成交（第 6-7 章）

信号生成环节定义了六种 action 分级，经过流动性、黑名单、冷却期三道初筛以及现金比例、日预算、行业集中度、单票红线等五道终调；交易执行环节通过 LongPort OpenAPI 实现了包含超时保护、溢价限价、成交验证与记账通知的完整自动化闭环。

15.1.4. 风控体系（第 8 章）

十道防线覆盖了账户级熔断（日亏 -5%、累亏 -15%）、持仓级约束（单票 $\leq 30\%$ 、杠杆 ETF $\leq 50\%$ ）、交易频率管理（三桶预算模型）以及行为纪律（48 小时冷却、行业分散）四个维度。三桶预算中仅 spontaneous 桶受动态日预算 N 约束，instruction 桶（用户指令）与 risk 桶（止损/止盈）不受限制，从制度上确保了风控动作永远不会被"预算用完"堵住。

15.1.5. 运行保障（第 9-12 章）

覆盖了盘中监控循环、Gateway 守护、三层可观测体系（Telegram 实时推送 + SQLite/JSONL 结构化日志 + 运行时日志）以及完整 Cron 调度表，实现了全链路可感知、可追溯、可审计。

15.1.6. 研究框架（第 13 章）

展示了支撑策略持续迭代的基础设施：32 因子库（6 大类）、3 年 IC 分析、CPCV 5-fold 交叉验证与 Optuna Bayesian 参数优化。回测结果显示，当前生产配置夏普比率 1.24，最大回撤 -19%，相比保守基线（夏普 0.76，最大回撤 -50%）改善显著。CPCV 各分段最差回撤仅 -6.3%，表明风控体系在不同市场阶段均保持了稳定的下行保护能力。

15.2. 核心方法论贡献

本报告的方法论贡献并不在于某个单一技术组件，而在于呈现了一套“可复用、可裁剪、可演进”的 AI Agent 工程化实践框架。

15.2.1. 分层解耦的系统设计范式。

六层架构将一套复杂的自主决策系统拆解为独立可测试的模块，每层只承担单一职责，层间通过固定 schema 的 JSONL 文件通信。当需要更换数据源、升级评分模型或切换执行通道时，仅需要修改对应层，其余层无需变动。第 2.6 节的降级策略表（Finviz 异常回退种子池、LLM 熔断 fallback 中性分、SQLite 故障降级 JSONL）进一步证明了分层设计的核心价值：****局部故障不会扩散为全局停摆****。这种“高内聚、低耦合”的工程思想，可直接迁移至信贷审批、合规监控、智能投顾等任何需要多环节串联的金融 AI 场景。

15.2.2. 量化 + AI 的互补融合范式。

本系统对 LLM 的使用策略——“放在辅助位而非决策位”——提供了一种务实的工程标杆。LLM 存在幻觉、随机性（同标的同 prompt 重复 10 次评分标准差约 5-8 分）和训练数据时效性等固有局限，不宜在高风险决策中充当“主裁判”；但 LLM 对定性信息（监管调查、管理层变动、新闻情绪）的捕捉能力又确实为纯量化模型提供了有价值的增量信息。60% 量化 + 20% AI + 20% 反向的权重结构，以及第 5.5 节演示的“量化权重如何在 LLM 幻觉时发挥压舱石作用”，是对“AI 在金融决策中应该占多大话语权”这一关键工程问题给出的“有数据和案例支撑的定量答案”。

15.2.3. “主动但不失控”的自主决策范式。

第 8 章的三桶预算与动态日预算 N、第 8.5 节的 48 小时止损冷却、第 6.3 节的行业集中度终调、第 8.3 节的日亏熔断（含浮亏口径，停止新买入但止损止盈继续执行）——这些机制共同勾勒了一条清晰的设计红线：****AI Agent 被赋予充分的主动性去发现和把握交易机会，但这种主动性被明确的风控边界、动态的行为约束和持续的**

人工可干预性所包裹。** 这一设计哲学对监管机构、合规部门和金融机构思考"AI Agent 的授权边界"具有直接参考价值——它不仅回答了"AI 能自主到什么程度",也回答了"什么情况下必须收回自主权"。

15.2.4. 可观测性作为系统的基础能力。

第 2.7 节和第 9.2 节反复强调的核心观点是: 追溯能力并非交易系统的附加功能, 而是其信任基础。三层观测体系 (Telegram 秒级感知 → SQLite/JSONL 结构化查询 → cron-logs 深度排障) 使任意交易日、任意一笔操作的完整决策链都可以被还原——从候选池构建, 到三引擎打分, 到两个 LLM 各自的评分与理由, 到 final-score 计算, 到信号终调是否被拦截及原因, 到下单价格与成交结果。

15.3. 策略定位与核心优势

15.3.1. 定位明确: 不拼信息差和速度差, 聚焦“认知差”

相较于依赖专线行情、卫星数据、共置服务器和数百人 PhD 团队的大型量化机构, 本系统在数据和速度层面不构成竞争。其差异化定位在于: 面对同样的公开信息, 通过机器纪律性克服人性弱点——严格止损不犹豫、评分触发即执行不拖延、移动止损保护浮盈不回吐、行业分散不押注单一赛道。

15.3.2. 以简单换取稳健

策略仅使用约 20 个可解释因子, 而非数百个因子的复杂黑箱模型。这一选择有文献和回测的支撑: 在因子投资领域, 过度复杂的模型更易出现过拟合、解释困难和样本外衰减。回测显示, 仅加入波动率过滤一项 (剔除年化波动 > 60% 的标的), 策略夏普便从 0.76 跃升至 1.24, 最大回撤从 -50% 收窄至 -19%——年化收益虽从 37% 降至 30%, 但风险调整后收益和下行保护均实现了质的飞跃。

15.3.3. 可解释本身就是竞争力

传统量化系统往往只能告诉用户“赚了还是亏了”, 却难以回答“为什么赚、为什么亏”。本系统的每一笔交易都保留完整决策链: 三引擎各自给出什么分数, 两个 AI 模型各自给出什么判断, 模型之间为何一致或分歧, 最终信号如何生成, 订单价格如何确定。

15.3.4. 低门槛可复现

系统全部工具链对个人开发者开放: yfinance 与 Finviz 为免费数据源, LongPort API 支持个人开户, MiniMax/GLM 按量计费 (日均约 \$3.2), OpenClaw 框架开源。具备一定编程基础的个人开发者可在 1-2 周内复现一套类似系统。这展示的不是“系统多

难复制”，而是“AI Agent 的自主决策能力正在变得可获得”。

15.4. 对卖方研究 AI 应用的三点启示

作为东吴地产 AI 系列报告的第四篇，本报告也承载着我们对方研究如何拥抱 AI Agent 的思考：

启示一：AI Agent 落地的关键在“工程化”，而非“模型强”。市场上不缺乏强大的基础模型，但将模型能力转化为稳定运行的生产系统，中间的工程鸿沟远被低估。第 5 章的幻觉处理与熔断冷却、第 2.6 节的降级策略、第 8 章的状态机与持久化设计——这些“非 AI”的工程细节，往往才是系统能否从 demo 走向生产的关键分水岭。“AI Agent 的真正门槛在工程，不在算法。”

启示二：从行业研究到跨界实践，是验证方法论通用性的唯一方式。一个地产研究团队去构建美股量化交易系统，这本身构成了一次“压力测试”——如果 OpenClaw 框架的方法论只在地产投研场景有效，那它就不是通用的 Agent 基础设施。如果连“真金白银实时决策”这种容错率极低的场景都能跑通，那它在信贷、合规、报价、审计等金融子领域的适用性就不再需要论证，而是有实证支撑的结论。

启示三：可审计的自主，比不可解释的智能更有商业价值。金融行业对 AI 的核心诉求不是“更聪明”，而是“可信任”。一个能追溯每一笔决策全部理由、能被任何合格第三方独立审计的系统，其商业落地阻力远低于一个预测更准但无从解释的黑箱。本报告贯穿始终的“全链路可观测”设计，本质上是在回答金融机构最关心的问题：“如果出了问题，我们能找到原因吗？”

16. 风险提示

- (1) AI 评级仅供参考，不构成投资建议：评分模型基于历史数据和特定因子构建，过去表现不代表未来结果。市场环境变化可能导致模型有效性下降。
- (2) 市场风险：REITs 价格受市场利率、宏观经济等因素影响可能波动。
- (3) 流动性风险：部分 REITs 交易不活跃，可能面临买卖困难。
- (4) 底层资产风险：基础设施资产的运营状况可能发生变化。
- (5) 模型风险：AI 模型的判断可能存在偏差或滞后。

免责声明

东吴证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本研究报告仅供东吴证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，本公司及作者不对任何人因使用本报告中的内容所导致的任何后果负任何责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

在法律许可的情况下，东吴证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

市场有风险，投资需谨慎。本报告是基于本公司分析师认为可靠且已公开的信息，本公司力求但不保证这些信息的准确性和完整性，也不保证文中观点或陈述不会发生任何变更，在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。经授权刊载、转发本报告或者摘要的，应当注明出处为东吴证券研究所，并注明本报告发布人和发布日期，提示使用本报告的风险，且不得对本报告进行有悖原意的引用、删节和修改。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

东吴证券投资评级标准

投资评级基于分析师对报告发布日后 6 至 12 个月内行业或公司回报潜力相对基准表现的预期（A 股市场基准为沪深 300 指数，香港市场基准为恒生指数，美国市场基准为标普 500 指数，新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的），北交所基准指数为北证 50 指数），具体如下：

公司投资评级：

- 买入：预期未来 6 个月个股涨跌幅相对基准在 15% 以上；
- 增持：预期未来 6 个月个股涨跌幅相对基准介于 5% 与 15% 之间；
- 中性：预期未来 6 个月个股涨跌幅相对基准介于 -5% 与 5% 之间；
- 减持：预期未来 6 个月个股涨跌幅相对基准介于 -15% 与 -5% 之间；
- 卖出：预期未来 6 个月个股涨跌幅相对基准在 -15% 以下。

行业投资评级：

- 增持：预期未来 6 个月内，行业指数相对强于基准 5% 以上；
- 中性：预期未来 6 个月内，行业指数相对基准 -5% 与 5%；
- 减持：预期未来 6 个月内，行业指数相对弱于基准 5% 以上。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议。投资者买入或者卖出证券的决定应当充分考虑自身特定状况，如具体投资目的、财务状况以及特定需求等，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

东吴证券研究所
苏州工业园区星阳街 5 号
邮政编码：215021

传真：（0512）62938527

公司网址：<http://www.dwzq.com.cn>