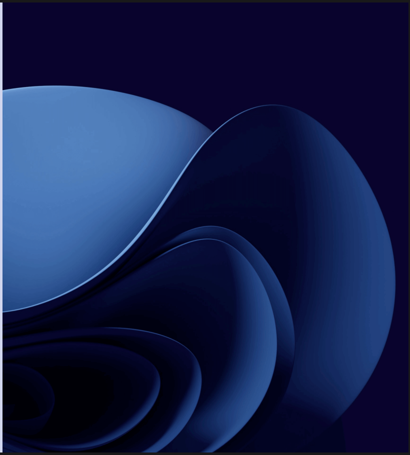
20260119

大模型
在量化交易策略中的应用



放在前面的问题——投资者的困惑

说到股票交易，会想到的问题有哪些呢?

A股5000多只股票，怎么筛选?

这只股票30块，是贵还是便宜?

每天那么多新闻，哪些值得关注?

什么时候买入？买多少?

买入后该关注什么？什么时候该卖出?

选择股票 → 估值判断 → 消息监控 → 买入时机/买多少 → 持有管理 → 平仓条件

基础概念（量化交易/大模型）

内容

量化系统架构（6个核心模块）

大模型在量化系统中的应用

问题闭环（解决投资者的实际问题）

关于大模型应用的局限性的探讨

01

基本概念——建立整体认知框架

什么是量化交易

传统投资方式	量化投资方式
靠"感觉"和"经验"决策	用数据和规则决策
情绪化：追涨杀跌	系统化：有明确的买卖标准
信息处理能力有限	可处理海量数据
难以坚持纪律	严格执行既定策略

核心理念：把模糊的投资直觉 → 转化为可执行、可检验的规则

量化交易的核心优势

纪律性

严格执行规则，不受情绪影响
不会因为恐惧割在最低点，也不会因为贪婪坐过山车

处理能力

同时分析上千只股票
人脑最多关注10-20只股票，计算机可以全覆盖

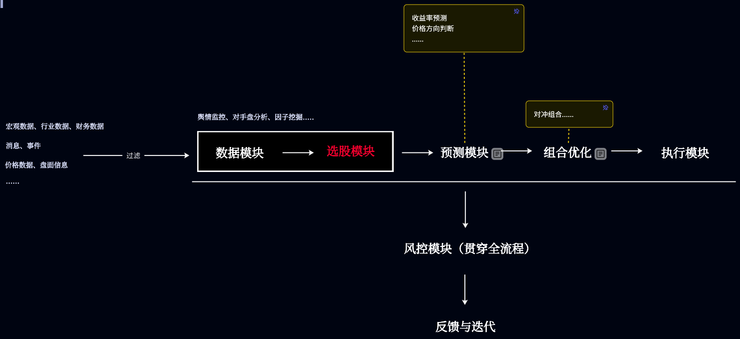
可回测

用历史数据验证策略是否有效
上线前就知道策略过去表现如何

速度快

毫秒级响应市场变化
新闻出来后自动分析并决策

量化策略完整架构



宏观数据、行业数据、财务数据
消息、事件
价格数据、盘面信息
……

过滤

舆情监控、对手盘分析、因子挖掘.....

数据模块 → 选股模块 → 预测模块 → 组合优化 → 执行模块

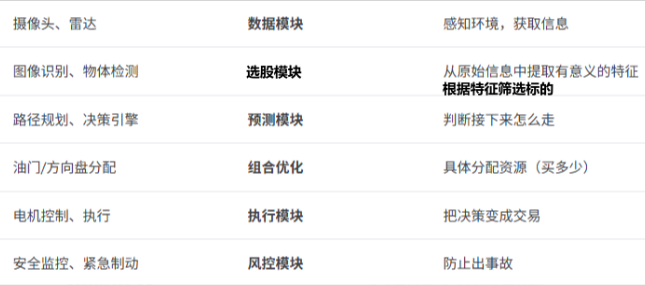
收益率预测
价格方向判断
……

对冲组合......

风控模块（贯穿全流程）

反馈与迭代

自动驾驶	量化策略	功能



摄像头、雷达	**数据模块**	感知环境，获取信息
图像识别、物体检测	**选股模块**	从原始信息中提取有意义的特征 **根据特征筛选标的**
路径规划、决策引擎	**预测模块**	判断接下来怎么走
油门/方向盘分配	**组合优化**	具体分配资源（买多少）
电机控制、执行	**执行模块**	把决策变成交易
安全监控、紧急制动	**风控模块**	防止出事故

模块与基本问题的基本对应

数据模块 信息从哪里来?	**选股模块** 如何筛选股票?	**预测模块** 未来会涨还是跌?	**组合模块** 买多少?	**执行模块** 怎么下单?	**风控模块** 如何防止大亏?
· 行情数据 · 财务数据 · 新闻资讯 · 另类数据	· 设计筛选条件 · 计算因子值 · 评估有效性	· 收益预测 · 方向判断 · 排序预测	· 权重分配 · 风险控制 · 约束满足	· 最优执行 · 降低成本 · 市场冲击	· 风险监控 · 止损止盈 · 异常检测

完整流程示例:
数据 收集茅台的价格、财务数据、最新新闻
因子 计算茅台的估值因子(PE=25)、动量因子(近期涨10%)
预测 基于因子预测：未来一周有65%概率上涨
组合 考虑风险后决定：买入5%仓位
执行 分3次下单买入，避免冲击价格
风控 全程监控，设置止损线
关键理解：这些模块紧密耦合，任何一个环节的短板都会拖累整体表现

大模型和自然语言处理的基本任务

为什么叫"大"模型?
- 参数多：GPT-4有上万亿参数（人脑神经突触约100万亿）
- 数据多：用互联网上的海量文本训练
- 能力强：能对话、写文章、写代码、做分析

文本分类

情感分析

文本摘要/信息抽取

意图识别

槽位填充

机器翻译

……

NLP 基本任务	在金融交易领域的具体应用场景	价值与目标
文本分类	1. 新闻/公告分类：自动将海量新闻、研报、公司公告分类为“财报发布”、“并购重组”、“政策变动”、“风险警示”等，便于交易员快速关注特定领域。 2. 主题识别：识别市场当前讨论的热点主题（如“人工智能”、“碳中和”、“供应链紧张”），用于主题投资和趋势跟踪。 3. 信号分类：将文本信息分类为“买入信号”、“卖出信号”或“中性信号”，为量化策略提供输入。	提高信息处理效率，实现自动化事件监控与筛选。
情感分析	1. 市场情绪分析：分析新闻标题、社交媒体、财经论坛的舆情，判断对某个股票、行业或整个市场的情绪是积极、消极还是中性，构建“情绪指数”。 2. 财报电话会议分析：分析公司高管在电话会议中的语调、用词情感，判断管理层信心，作为基本面分析的补充。 3. 事件影响评估：在突发事件（如财报超预期、产品事故）后，快速评估市场主流情绪的走向。	量化市场“贪婪与恐惧”，提供反直觉的交易信号或趋势确认，是量化 alpha 和风险管理的核心因子。
文本摘要/信息抽取	1. 自动研报摘要：将几十页的券商研报自动浓缩为核心观点、目标价、评级变化，节省阅读时间。 2. 关键信息抽取：从公司公告、政策文件中自动抽取关键数据，如“净利润同比增长X%”、“新签订单金额Y亿元”、“利率上调Z个基点”，并直接写入数据库或触发交易指令。 3. 事件抽取：从新闻中自动识别出“公司-事件-影响”三元组，例如“A公司发布新产品B，预计将抢占C市场”。	从非结构化文本中快速提取结构化、可量化的关键事实，直接驱动算法交易和投资决策。
…（其他任务）	1. 关系抽取：构建“公司-高管-子公司”知识图谱，用于风险传导分析。 2. 命名实体识别：自动识别文本中提到的所有公司、人物、产品名称，用于信息关联。 3. 机器翻译：实时翻译全球财经新闻，打破信息壁垒，捕捉跨市场套利机会。	构建更深层次的金融知识网络，实现更广泛的全球市场信息覆盖。

01

数据模块

数据模块：处理哪些数据?

传统金融数据

- 行情数据（OHLCV）
- 财务数据（三大报表）
- 基本面指标

高频微观数据

- 逐笔成交
- 订单簿快照
- 订单流

另类数据

- 新闻资讯
- 社交媒体
- 卫星图像
- 供应链数据

另类数据案例：
2020年初，有对冲基金通过卫星图像监测中国城市夜间灯光和高速公路车流量，比官方数据更早发现经济复苏迹象

量化交易的一个痛点

量化交易的一个痛点
传统量化主要处理结构化数据（价格、财务数字），但市场上还有大量非结构化数据无法利用

结构化数据
- 股票价格：10.25元
- 市盈率：15.3倍
- 营收增速：23.5%

传统量化处理很好

非结构化数据 → 大模型
- 新闻："公司否认造假指控"
- 研报：几十页分析报告
- 公告：复杂的法律文书

大模型可以处理

大模型在数据模块的应用

复杂语义理解

"公司否认造假" → 理解是"否认"→ 正面

传统关键词法
"公司否认造假" → 检测到"造假"→ 负面

智能信息提取

输入新闻:
"苹果公司宣布以185美元回购1100亿股票"

大模型输出:
事件类型: 股票回购
金额: 1100亿美元
情感: 正面

注：非结构化数据变成结构化数据

这是大模型在金融领域最成熟的应用场景之一

案例：BloombergGPT

Bloomberg专门训练的金融大模型
2023年发布，是金融领域最权威的专用大模型
https://arxiv.org/abs/2303.17564

技术规格
- 参数量：500亿
- 训练数据：7000亿token
- 金融数据占比：51.9%
- 训练成本：约130万美元

性能表现
- 金融情感分析：比通用模型高9%
- 金融实体识别：比通用模型高8%
- 金融问答：比通用模型高9%

关键洞察
- 金融专用模型 > 通用模型（领域知识有价值）
- 数据质量至关重要（Bloomberg的优势在数据）

输入：原始新闻
"据路透社报道，苹果公司于2024年1月15日宣布，将以每股185美元的价格回购总价值1100亿美元的公司股票。这是苹果历史上最大规模的股票回购计划。CEO库克表示，这反映了公司对长期增长前景的信心。"

输出：关于回购的结构化信息
事件类型：股票回购
公司：苹果公司 (AAPL)
日期：2024-01-15
金额：1100亿美元
回购价格：185美元/股
信息来源：路透社
情感倾向：正面
管理层态度：乐观

价值：这种结构化信息可以直接输入量化模型，构建"回购事件因子"

大模型在数据模块的局限

大模型的局限

幻觉问题
可能编造不存在的信息

成本较高
处理海量新闻需要大量API调用

模型版本变化
同样输入可能输出不同

应对策略：必须有人工抽检机制，保留原文便于复核

02

选股模块

因子选股

因子就是筛选股票的"条件"或"指标"

实际筛选示例

目标：找"便宜又赚钱"的公司

股票	PE	ROE	营收增速	负债率	结果
A公司	12	18%	25%	45%	入选
B公司	8	22%	30%	75%	负债太高
C公司	25	20%	35%	40%	太贵了

5000只 → 筛出200只候选

什么是因子?

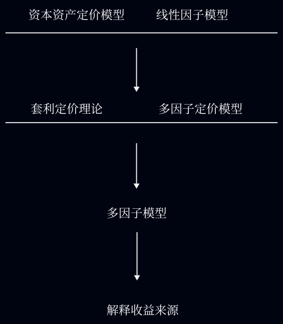
因子是一个能够解释股票收益差异的变量

类比理解：预测学生考试成绩
- 学习时间（勤奋因子）
- 智商（能力因子）
- 睡眠质量（状态因子）

经典因子示例

因子名称	计算方式	背后逻辑
市值因子	log(总市值)	小盘股长期跑赢大盘股
价值因子	1/市净率 (1/PB)	低估值股票长期表现更好
动量因子	过去12个月收益率	强者恒强
质量因子	ROE (净资产收益率)	盈利能力强的公司表现好
波动率因子	过去60日收益标准差	低波动股票风险调整收益更高

因子挖掘的起源



资本资产定价模型　线性因子模型

套利定价理论　多因子定价模型

多因子模型

解释收益来源

传统因子挖掘的挑战

人工效率低
研究员一个月可能只能设计和测试几个因子

受先验偏见影响
只测试"觉得有道理"的因子，可能错过真正有效的

难发现高阶关系
人脑难以想象"三个因子同时满足某条件时才有效"

因子拥挤
经典因子被大量使用，alpha被套利殆尽

大模型在选股模块的使用——辅助因子设计

研究员问："如何量化公司的'护城河'？"

AI建议：
- 毛利率稳定性
- 研发投入占比
- 客户集中度
- 专利数量...

辅助工具，但核心还需人工设计

大模型在选股模块的使用——挖掘选股因子

展望：端到端的因子挖掘（因子包含字参数黑盒中） → 特征挖掘

核心思想
让神经网络自动从原始数据中学习特征，不需要人工设计因子公式

架构示例（修改后）

输入
过去60天的OHLCV数据（300维）
↓
嵌入与位置编码层 将输入数据线性映射为嵌入向量，并加入位置编码以保留时序信息
↓
Transformer 通过多层自注意力机制和前馈网络，并行捕捉全序列的复杂依赖关系
↓
输出 一个“深度学习因子”值

劣势：黑箱，无法解释因子含义

大模型在选股模块的局限

局限与风险

黑箱问题

无法解释因子含义

过拟合风险
可能记住噪声而非规律

需要大量数据
深度模型参数多

03

预测/信号模块

预测/信号模块?

预测模块是策略的"决策核心"，负责基于因子生成具体的交易信号

点预测
股票A明天涨多少?

（回归问题）

方向预测
股票A明天涨还是跌?

（分类问题）

概率预测
股票A涨跌的概率各是多少?

（概率估计）

排序预测
哪些股票明天表现更好?

（排序问题）

预测的本质困难

信噪比极低
真正可预测的信号非常微弱

规律会变化
过去有效不代表未来有效

自我毁灭
规律被发现后会消失

大模型在预测模块的应用

实验性阶段 — 大模型直接预测股价效果有限

可行的应用：信息融合

让大模型整合多种信息，辅助决策：

整合分析：
- 技术指标显示超买
- 基本面估值偏高
- 新闻情绪偏正面

建议：谨慎看多，控制仓位

局限性

不擅长精确数值预测
可能"记住"历史信息（数据泄露）
输出的"信心"不可靠

辅助决策有价值，直接预测需谨慎

注：也是端到端，中间是黑箱

03

其他模块

组合优化 & 执行模块——"买多少"和"怎么买"

组合优化
预测告诉你"哪些好"，组合优化决定"买多少"

考虑因素：

- 风险分散（不要把鸡蛋放一个篮子）
- 风险对冲
- 持仓限制

……

执行模块
把决策变成真实的交易订单

核心挑战：

- 市场冲击：大单会推高价格
- 交易成本：频繁交易侵蚀收益

……

这两个模块需要精确计算,大模型应用有限

风控模块

事前风控

- 持仓限制
- 行业限制
- 杠杆上限

事中风控

- 实时监控
- 回撤预警
- 异常检测

事后风控

- 归因分析
- 复盘总结
- 策略优化

一个策略可能95%时间赚钱，但如果那5%的时间亏光本金，就毫无价值

案例1: LTCM
长期资本管理公司

- 诺贝尔奖得主创办
- 年化收益超40%
- 1998年俄罗斯危机破产
- 亏损46亿美元

案例2: Knight Capital
骑士资本

- 美国最大做市商之一
- 一个软件bug
- 45分钟亏损4.4亿美元
- 公司破产被收购

大模型在风控模块的应用

风险识别和报告生成是核心应用场景

风险事件识别

新闻:
"某公司被证监会立案调查"

AI分析:
风险类型：监管风险
严重程度：高
建议：立即减仓或清仓

风险报告生成

每日风险报告

- 最大回撤: 5.2%
- 重点关注：股票A收到问询函
- 建议：控制总仓位

……

自动生成可读的分析报告，节省人力

大模型风险识别示例

输入：4条风险相关新闻

局限与风险

某公司CEO因个人原因辞职
某公司被证监会立案调查
某公司回应传闻，否认存在财务造假
某公司收到交易所问询函

大模型输出：结构化风险分析

新闻事件	风险类型	风险等级	处理建议
新闻1	治理风险	**中等**	关注继任者
新闻2	监管风险	**高**	**建议清仓**
新闻3	声誉风险	**中低**	持续关注
新闻4	监管关注	**中等**	等待回复

幻觉问题
可能虚构或误判风险

极端事件数据稀少
真正的危机很少发生

责任界定
大模型判断错误谁负责?

04

回顾之前的问题

1、选什么股票？A股5000多只，怎么筛选？

2、估值高了还是低了？现在30块，是贵还是便宜？

3、监控什么消息？每天那么多新闻，哪些值得关注？

4、什么时候买？买多少？消息出来后该怎么操作？

5、什么时候卖？怎么不卖飞、也不坐过山车？

选什么股票

量化方案：因子选股

股票	PE	ROE	入选
A公司	12	18%	✅
B公司	25	20%	

大模型辅助（当作为投资者没有自己的交易体系，只有感性的认知，大模型可以帮我们量化）
你说："找新能源行业，市值中等，业绩增长的股票"

AI转化：
行业 = 新能源
100亿 < 市值 < 500亿
近3年净利润增速 > 0

是Google提供的备忘录应用，具有的界面和快速的备忘录功能。它与Google账号同步，方便用户在不同设备上访问备忘录。

估值高了还是低了？

量化方案：PE百分位

过去5年PE数据：

最低PE：15倍 (熊市)
最高PE：45倍 (牛市)
当前PE：18倍

当前处于历史10%分位
→ 相对便宜

对于大模型的使用很有局限性

你问：PE只有8倍，是不是很便宜？

AI分析：
PE低可能有几种原因：
1. 周期股在盈利高点
2. 市场预期业绩下滑
3. 公司有隐藏风险

⚠ 可能是"估值陷阱"

目前只能起到一个分析助手的作用

监控什么消息？

量化解决方案：事件驱动 + NLP

事件类型

业绩相关 财报、业绩预告
资本运作 增发、回购、减持
行业政策 补贴、监管变化
公司动态 大合同、新产品

大模型实时解读

公告原文："公司预计净利润同比增长30%-50%"

AI解析：

事件：业绩预告
方向：正面
情绪评分：+0.8

买卖决策与持仓管理

买多少?

仓位 = 基准 信号强度 风险调整

示例:

基准：10%
信号：0.8
风险：0.7

→ 仓位：5.6%

持有关注?

监控仪表盘:

- 盈亏比例
- 相对强弱
- 资金流向
- 新消息

何时平仓?

规则化止盈止损:

- 固定止损：-8%
- 移动止损：回撤10%
- 目标止盈：+20%
- 逻辑破坏：清仓

大模型可以提供建议，汇总信息、辅助决策分析，但核心规则仍需人工设定

03

关于大模型应用的局限性的探讨

大模型在各模块的参与程度

模块	参与深度	大模型核心作用
数据模块	★★★★★	文本处理、信息提取（处理新闻、公告等非结构化数据）
选股模块	★★★★	情绪因子生成、辅助因子设计
预测模块	★★	多源信息整合、自动生成推理与报告
组合模块	★	应用有限（仍以传统优化方法为主）
执行模块	★	应用有限（属于低延迟优化问题）
风控模块	★★★★★	文本风险识别、自动生成风控报告

1. 处理非结构化数据
新闻、公告、研报 → 结构化信息

2. 理解复杂语义
"否认造假" → 正面，而非负面

3. 识别风险事件
快速分类和评估风险严重程度

4. 辅助人工决策
生成报告、提供多角度分析

现在大模型在量化领域应用的基本特点

更多的是在各个阶段作为分析助手，没有真正融入到模块的功能中

更多是定性的建议，即使能给出定量的回答，也无法倒找其根据

黑箱：不可解释

◆ 角色定位：
外挂式助手
更多的是在各个阶段作为分析助手
没有真正融入到模块的功能中

◆ 输出特性：定性为主
更多是定性的建议
即使能给出定量的回答，也无法倒找其根据

◆ 本质问题：黑箱不可解释
无法追溯决策逻辑
无法定位错误来源
无法验证、无法信任

量化交易需要： 可追溯、可解释、可验证、高可靠

传统大模型与自成长模型之间对比

对比维度	传统大模型（脑子糊涂）	自成长模型（脑子清楚）
可解释性	黑箱决策，无法追溯原因	参数可解释，每个选项可追溯
知识更新	全量训练，6-12个月周期	增量学习，推理中即时训练
错误修正	无法精确定位，只能重新训练	精准定位修正，只改错误部分
知识冲突	灾难性遗忘，新知识覆盖旧知识	冲突识别解决，新旧知识共存
训练成本	数千万美元，需要大量算力	降低550倍训练
上下文限制	Token限制，2K-128K tokens	结构记忆，理解融入结构
可靠性	不知道边界，乱编也很自信	知道能力边界，区分确信与猜测

核心差异：可解释性是一切的基础
理解就是结构，符号化的理解结构就是可解释的结构

突破与展望

可解释性 → 一切的基础

经济：降低成本

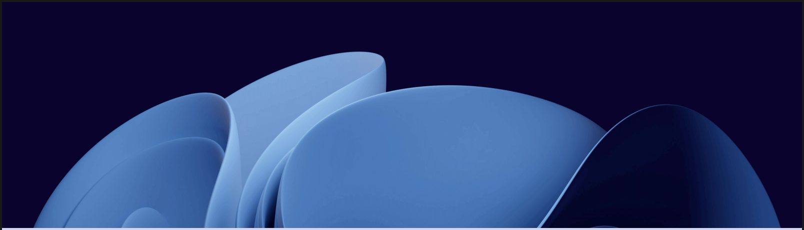
可靠：可追溯、无幻觉

自我进化：持续学习

标准化：可协作

从"外挂助手"升级为"核心引擎" → 大模型真正融入策略决策流程
从"定性建议"升级为"可追溯可执行的决策" → 每个交易信号都能归因
从"静态模型"升级为"动态进化" → 策略在实盘中持续学习优化

愿景：让AI拥有可解释的清醒的"大脑"



致谢