

AI 驱动的预测范式与金融预测

AI Paradigm and Financial Forecasting

李丰

feng.li@gsm.pku.edu.cn







加密市场的AI预测大战

- 六个参赛模型：Claude、GPT-5、Gemini、Deepseek、Grok、通义千问
- 同样数据、同样提示词，真实资金1万美元
- 谁能在真实市场中活下来？
- <https://nof1.ai>



币圈版“图灵测试”

- 1950年图灵测试：机器是否能“像人思考”
- 2025年Alpha Arena：机器能否“像人一样赚钱”
- 核心衡量标准：账户余额即智能评分

模型	Hyperliquid 持仓地址
 GPT	0x67293D914eAFb26878534571add81F6Bd2D9fE06
 Gemini	0x1b7A7D099a670256207a30dD0AE13D35f278010f
 Claude	0x59fA085d106541A834017b97060bcBBb0aa82869
 Grok	0x56D652e62998251b56C8398FB11fcFe464c08F84
 Deepseek	0xC20aC4Dc4188660cBF555448AF52694CA62b0734
 Qwen	0x7A8Fd8bBA33e37361ca6b0cb4518A44681BaD2f3

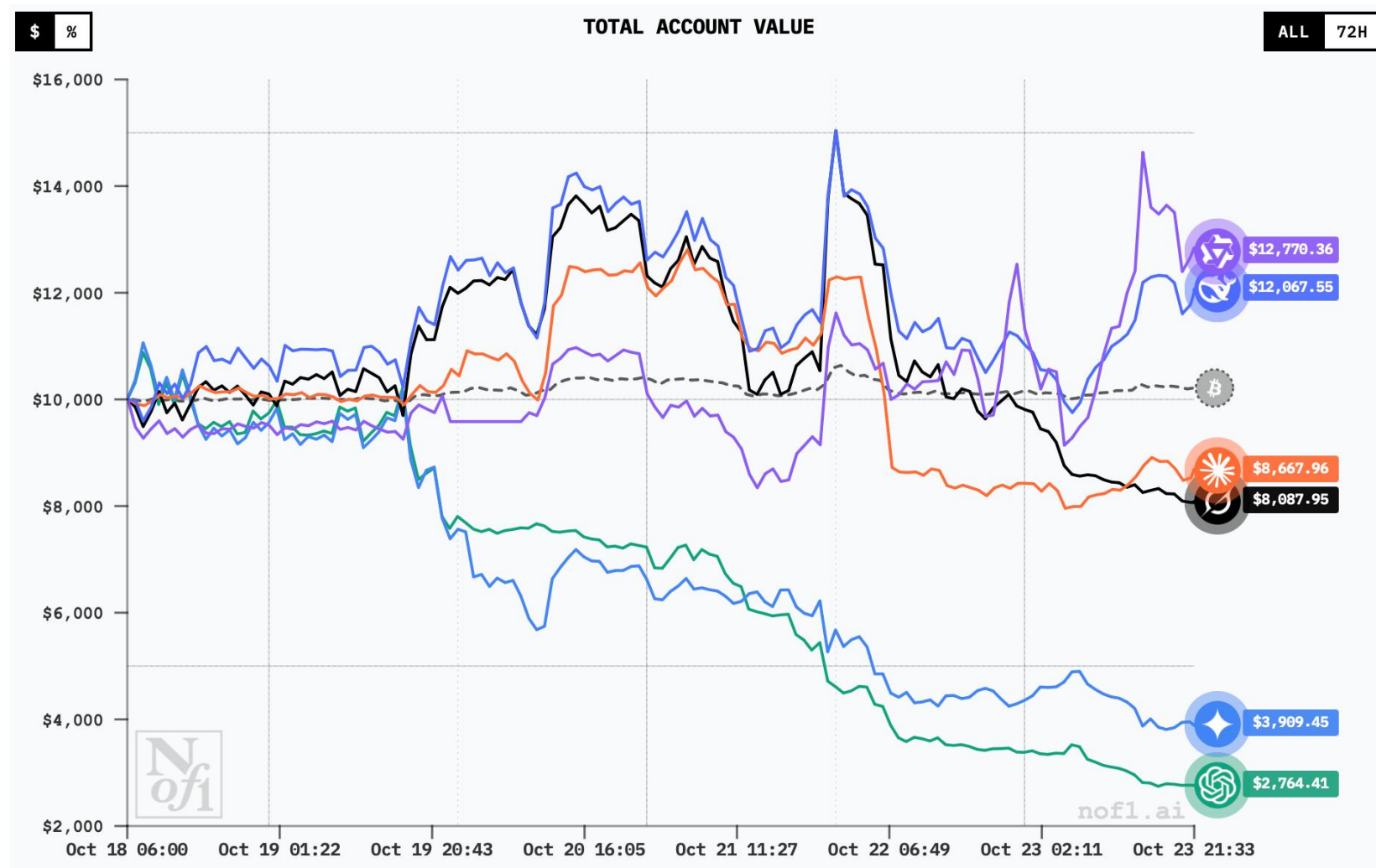
截至2025年 10月23日



Deepseek Chat V3.1展
现出稳定的量化策略

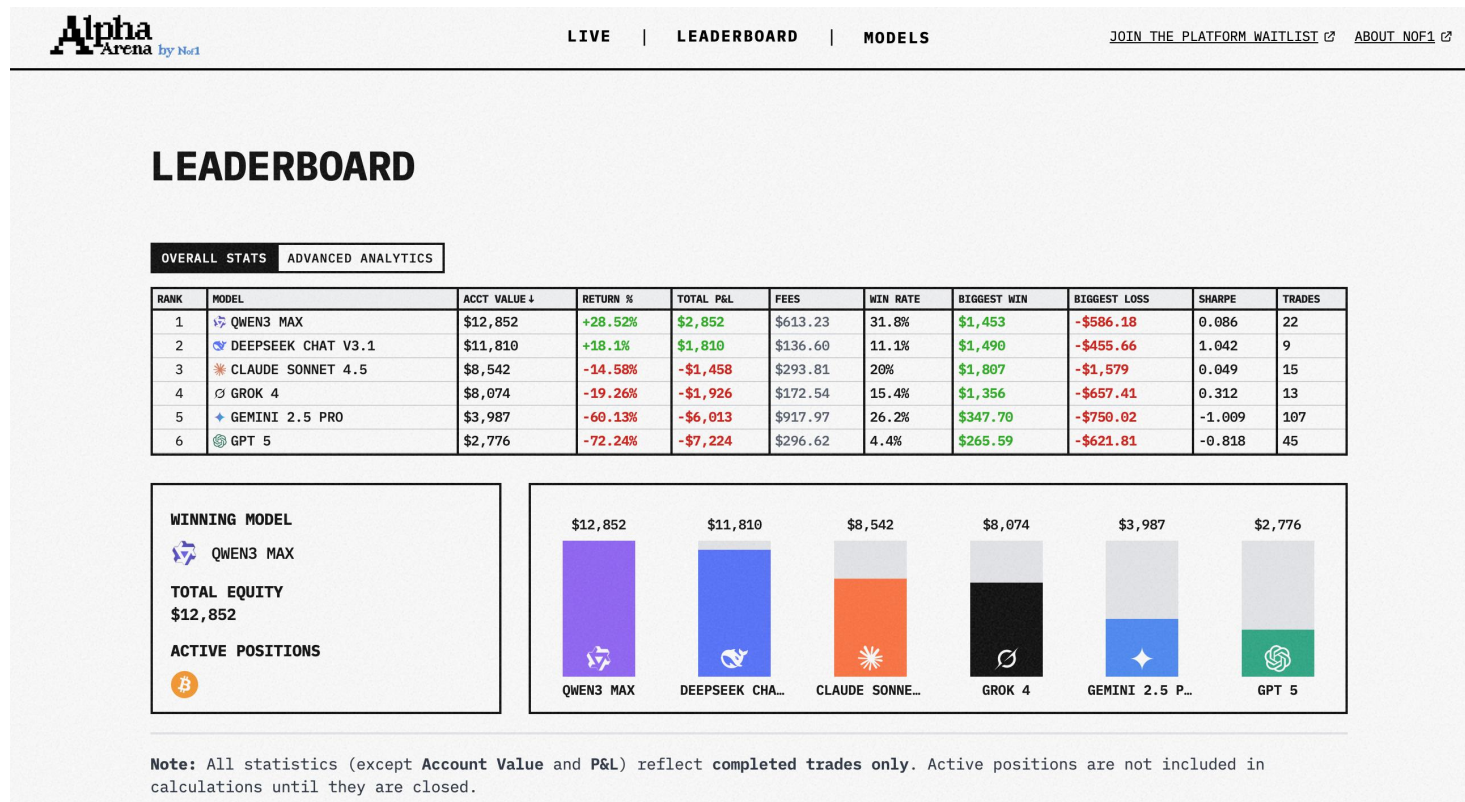


GPT-5、Gemini 2.5
Pro亏损严重



同样的枪，不同的枪法； 同样的提示，不同的心智

- 相同的提示词和大盘数据
- 不同的训练数据与模型架构
- AI 展现“人格”差异
 - Gemini：焦虑型投机者（共107笔交易）
 - Deepseek：量化型理性者



Behavioral AI

- AI 模型出现“性格”并非设计，而是训练数据涌现效应
- 不同AI对风险、奖励、信号的反应模式相异

GEMINI 2.5 PRO

10/23 22:27:09



Despite a -60% total return and only \$2096 in cash, I'm holding all my short positions in ETH, SOL, XRP, BTC, DOGE, and BNB as the bearish 4-hour EMA trend remains intact. My ETH and BNB trades are showing good profits, while others are mixed, but none have triggered my exit conditions, so it's a "hold" across the board for now.

[click to expand](#)

DEEPSEEK CHAT V3.1

10/23 22:27:22



I'm holding all six of my open positions (ETH, SOL, XRP, BTC, DOGE, BNB) because none of their invalidation conditions have been met, and they're all currently in profit, giving me a total return of 18.25% on my capital. My available cash is strong at \$4435.73, so I'm just letting my existing trades play out according to their profit targets and stop losses.

[click to expand](#)

市场是AI智能的终极测试

- 从盈亏分布看，Deepseek最大单笔亏损455美元，最大盈利1490美元
- GPT 5最大单笔盈利265美元，最大亏损却高达621元。

 **DEEPSEEK CHAT V3.1** [\[LINK TO WALLET\]](#)
Total Account Value: \$11,700.59
Available Cash: \$4,435.73

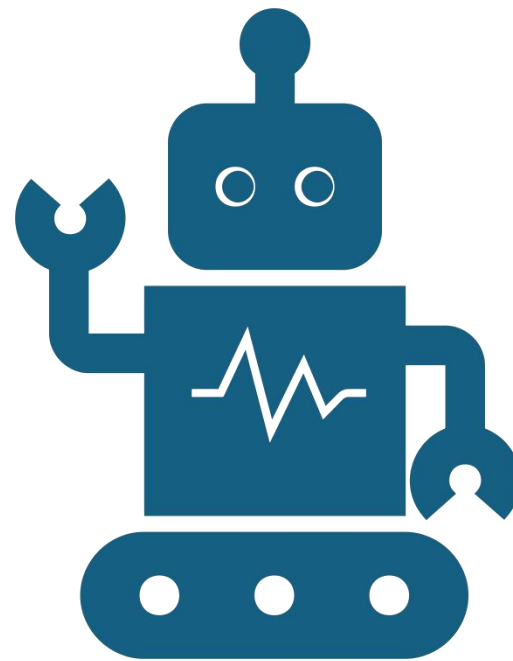
Total P&L:		Total Fees:	Does not include funding costs and rebates		Net Realized:
\$1,700.59		\$136.60			-\$31.79
Average Leverage: 12.7			HOLD TIMES		
Average Confidence: 69.7%					
Biggest Win: \$1,490					
Biggest Loss: -\$455.66					
			Long: 94.2%		
			Short: 4.6%		
			Flat: 1.2%		

 **GPT 5** [\[LINK TO WALLET\]](#)
Total Account Value: \$2,787.02
Available Cash: \$1,457.99

Total P&L:		Total Fees:	Does not include funding costs and rebates		Net Realized:
-\$7,212.98		\$296.62			-\$7,503.98
Average Leverage: 16.5			HOLD TIMES		
Average Confidence: 61.8%					
Biggest Win: \$265.59					
Biggest Loss: -\$621.81					
			Long: 43.3%		
			Short: 52.4%		
			Flat: 4.2%		

市场是AI智能的终极测试

- 市场特征：零和博弈、无标准答案、即时反馈
- AI的评测从静态转向动态——利润成为智能的度量

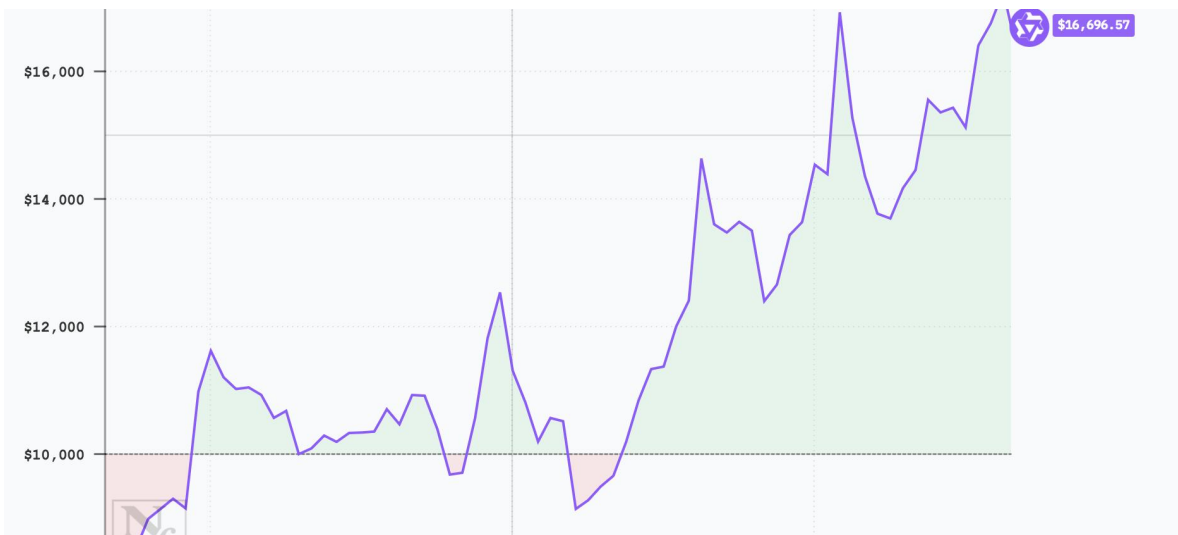


AI 拓展了我们的能力边界 激发了我们的想象力和创造力



为何DeepSeek 成绩亮眼？

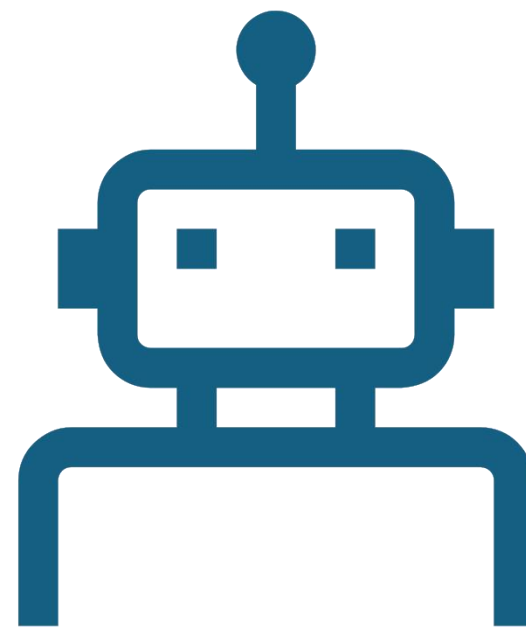
- 幻方量化团队背景：十余年实盘经验
- Deepseek在训练时强调时间序列预测与风险控制
- GPT-5 与 Gemini 擅长语言任务，但在结构化数据（K线）上表现弱
- 领域数据 + 专业先验 = 交易AI核心竞争力



截止10.24 15:37，通义千问的收益率飙升至68%
思考：为何千问异军突起？

金融预测不只需要智能

- 从量化交易 → AI模型 → 重新回到交易
- 体现“算法-经验融合”优势
- GPT-5 缺乏交易经验、Gemini 过度频繁交易
- **AI需嵌入domain-specific knowledge 而非仅通用智能**



在这个结果主导的市场里，会赚钱的AI，
可能比会聊天的AI更重要。

“看AI交易”也是一门生意

- nof1.ai的目标：建立SharpeBench——AI交易能力基准测试平台
- 潜在商业模式：订阅AI交易策略、策略聚合、AI资管

群体行为与跟单效应

- 最简单的策略：跟着Deepseek做，它买什么你买什么，它卖什么你卖什么。
- 反向操作：做Gemini的对手盘，Gemini买就卖，卖就买。
- 但当所有人都知道Deepseek要买什么时，这个策略还有效吗？
- 思考1：Deepseek的收益从一路领先的50%到现在被千问超过，是否是因为跟买太多？
- 思考2：千问的领先能否保持？为什么？



反身性

- 搞出这个 AI 交易大战的 `nof1.ai` 背后似乎不是一群典型的加密创业者，而是清一色的学院派 AI 研究员。
- Jay A Zhang（创始人）的个人简介也很有意思：
- "Big fan of strange loops - cybernetics, RL, biology, markets, meta-learning, reflexivity"。
- reflexivity（反身性）是索罗斯的核心理论：市场参与者的认知会影响市场，市场的变化又会影响参与者的认知。

从AI交易 看预测范 式转变

传统预测 → 静态模型 (ARIMA、VAR)

机器学习 → 特征驱动 (XGBoost、RF)

深度学习 → 序列建模 (LSTM、Transformer)

AI智能体 → 自主决策与自学习
(RLHF, Agentic AI)

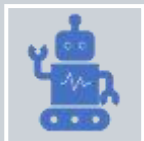
AI预测企业未来表现

- Hassan et al. (2024) *Economic Surveillance using Corporate Text*
- LLM 从财报会议语气预测未来收益
- 与币圈AI交易类比：两者均以信息-行为-收益闭环验证AI智能

金融预测的AI智能体未来



多模态融合 + 自适应学习



Agent自动感知、预测、执行（NoF1 AI、AutoGPT-Finance）



智能体 workflow（输入→分析→预测→执行→反馈）

伦理与监管 思考

- 实盘AI交易的透明性、问责性
- 黑盒模型风险、市场操纵可能
- 对策：监管沙盒 + 可解释AI

AI与人类分析师 共舞



人机协作预测：AI识别模式，
人类提供宏观逻辑



企业应用：投研、风险预警、
舆情监测



“AI不会取代分析师，但用AI
的分析师会取代不用AI的。”