

InelidaMarketScan — 基于机器学习的金融市场科学分析

量化金融研究: ICT、一目均衡表、XGBoost与随机计算 作者 Didier Vally /
Reuniware Systems 版本 1.0.6 — 2026-07-19 □ PyPI:
<https://pypi.org/project/inelida-marketscan> | GitHub:
<https://github.com/reuniware/InelidaMarketScanner>

摘要

本文呈现了6个月量化金融研究在算法交易中的应用成果。InelidaMarketScan框架结合了ICT(内部圈交易者)技术分析和多时间框架一目均衡表, 配合XGBoost机器学习管道(123个特征)和随机计算函数(Ornstein-Uhlenbeck、Hurst、GARCH、Monte Carlo、Kelly)。v18模型在1,210笔历史交易上达到64.0%的交叉验证准确率。对27个连续模型的分析表明, 一目均衡表特征(Kijun/Tenkan距离、T/K交叉)比ICT模式(FVG、Order Blocks)更具预测性。ML% ≥ 50%过滤的资本模拟显示样本内利润因子为2.27, 而样本外验证识别出上限为1.173。本文是关于ICT、一目均衡表与机器学习交汇的完整科学参考。

1. 引言

1.1 背景与问题陈述

算法交易依赖于识别金融时间序列中的重复模式。由Michael Huddleston开发的ICT(内部圈交易者)概念描述了机构市场行为: 流动性扫荡、公允价值缺口、订单块、市场结构转变。同时, 一目均衡表指标提供了市场均衡的多时间框架读数。本项目通过自动化扫描器与XGBoost模型探索这两种范式的融合。

1.2 研究目标

- 开发一个自动扫描器, 在8个时间框架(M5到MN)上结合ICT和一目均衡表。
- 构建一个具有123个命名特征的ML管道, 从扫描器中提取。
- 通过历史回测(2025-2026)和样本外测试验证预测能力。
- 集成量化金融函数(OU、Hurst、GARCH、Monte Carlo、Kelly)。
- 在实际条件下实现利润因子 > 1.5。

2. ICT理论基础

ICT概念模拟了操纵价格水平以积累头寸的金融机构(银行、对冲基金)的行为。扫描器自动检测以下8个概念:

ICI概念	描述与检测
流动性扫荡	短暂突破关键水平(亚洲高/低、伦敦高/低)以触发止损后再反转。在H1上检测并确认反转。
公允价值缺口(FVG)	连续3根K线之间价格未交易的缺口。在6个时间框架(M5到D1)上检测并显示填补百分比。
订单块(OB)	方向性冲动前的最后一根K线 — 机构下订单的区域。通过ATR(14)在6个时间框架上检测。
断路器块(BRK)	被突破的订单块反转为相反的支撑/阻力。从现有OB派生。
市场结构转变(MSS/BOS)	市场结构变化:结构突破和性格变化,在6个时间框架上检测。
成交量HVN/LVN	基于MT5 tick成交量的高成交量节点和低成交量节点。在6个时间框架上检测。
反转管道	完整的扫荡→反转→确认链,带有扫荡后K线计数器。
斐波那契扩展	基于亚洲区间的+1.618到+4.0扩展,用于止盈目标。

3. 多时间框架一目均衡表架构

Diamond扫描器同时分析8个时间框架。每个TF计算5个一目均衡表组件:

时间框架	Tenkan	Kijun	Kumo	T/K Cross	Flat History
M5	☐	☐	—	☐	—
M15	☐	☐	—	☐	—
M30	☐	☐	—	☐	—
H1	☐	☐	☐	☐	☐
H4	☐	☐	☐	☐	☐
D1	☐	☐	☐	☐	☐
W1	☐	☐	☐	☐	☐
MN	☐	☐	☐	☐	☐

4. Diamond扫描器架构

Diamond扫描器汇聚了10个类别中的109个评分标准:

类别	标准数	示例
一目均衡表 H1-D1	24	Kijun/Tenkan距离、T/K交叉、Kumo位置、平坦K线
一目均衡表 W1-MN	8	Kijun W1/MN、Kumo W1/MN、云颜色、双重记忆
FVG(6个TF)	12	M5到D1的FVG空头/多头、填补%、时间优先级
订单块(6个TF)	12	M5到D1的OB空头/多头、位置上方/下方/平
MSS/BOS(6个TF)	18	M5到D1的体制、BOS、ChoCH
成交量HVN/LVN(6个TF)	12	M5到D1的成交量尖峰、HVN、LVN
断路器块(6个TF)	12	M5到D1的BRK空头/多头
亚洲/伦敦扫荡	4	AH/AL扫荡、伦敦H/L扫荡
反转管道	5	扫荡→反转→确认、扫荡后FVG、重测
M30/M15/M5 Kijun	2	Kijun M30/M15/M5方向交叉

最高分: 109 标准(无MN为107)

5. 机器学习管道

ML管道通过3个阶段将扫描器结果转化为获胜概率:

阶段	工具	详情
1	ml_labeler.py	从Diamond扫描提取特征(数据库/手动CSV/回测)。生成带有outcome列的标记CSV(1=赢, 0=输)。
2	ml_trainer.py	XGBoost训练, 5折交叉验证, scale_pos_weight处理不平衡类别, 超参数网格搜索。
3	ml_predictor.py	实时推理: extract_features() → predict_proba() → ML% 在实时扫描中显示。简化的MLPredictor: 1次joblib.load而非5次。

6. XGBoost模型v18 — 配置

v18模型是活跃的生产模型。在1,210笔历史交易上训练, 具有123个命名特征(包括GARCH和随机特征)。

超参数	值
max_depth	6
learning_rate	0.03
n_estimators	200
subsample	0.8
min_child_weight	3
gamma	0.1
reg_alpha	0.5(L1)
reg_lambda	1.0(L2)
scale_pos_weight	自动(基于赢/输比率)

123个命名特征包括: 分数/偏差(6)、Kijun 6个TF(18)、Tenkan 6个TF(18)、T/K交叉(6)、Kumo(6)、M5/M15/M30交叉(12)、平坦K线(2)、活跃FVG(1)、订单块(6)、扫荡(5)、成交量(6)、MSS/BOS(6)、断路器块(6)、压缩(1)、随机(8: OU、Hurst、Monte Carlo、GARCH)、时间(3)、资产类型(3)。

7. 模型历史

自项目启动以来, 已训练和评估了27个XGBoost模型。以下是主要版本的历史:

模型	数据	特征	交易	CV	OOS PF
v1	6个主要货币对	36	852	61.0%	—
v4	13个符号	111	1,210	62.9%	1.26*
v7	13个符号	25个最佳	2,923	60.1%	1.173
v13	4个模型/资产	25	2,923	53-64%	4.000
v15	2025-2026	25 + OU/Hurst	5,009	59.8%	—
v18	13个符号	123(GARCH)	1,210	64.0%	0.503 F1

8. 特征重要性(v4)

特征重要性分析揭示了哪些标准在模型决策中权重最大：

排名	特征	增益	洞察
1	is_dxy	14.2	美元遵循与外汇对根本不同的逻辑
2	rr	10.0	理论风险/回报比预测实际结果
3	tkx_h1	9.2	H1 Tenkan/Kijun交叉是最佳单一预测因子
4	kj_h1	6.9	到H1 Kijun的距离捕捉短期均衡
5	s1	6.4	止损水平与止盈同等重要
6	day_wednesday	6.4	周三(FOMC)显示统计上独特的行为
7	d_kj_h4	6.3	到H4 Kijun的距离指示潜在趋势
8	d_kj_d1	6.0	到D1 Kijun的距离锚定宏观背景
9	tkx_h4	6.0	H4 T/K交叉确认交易时段趋势
10	is_forex	5.9	外汇对有独特行为vs指数/金属

9. 样本外验证

OOS验证(2026年1月-6月训练, 2026年7月测试)是决定性测试。v4模型在ML% ≥ 50% 阈值下显示样本内PF为2.27, 但这包含数据泄漏(模型在训练期间看到了80%的测试数据)。v7的真实OOS PF是1.173 — 统计上显著但适度的优势。v13模型(按资产)在15笔交易上达到PF=4.000, 但样本太小不可靠。结论:ML提供真实的优势(PF > 1.0), 但统一架构的当前上限约为1.17。

10. 资本模拟 — ML%过滤

在1,210笔历史交易上模拟, 初始资本10,000欧元, 每笔交易风险1%:

ML%阈值	交易	胜率	PF(RR)	盈亏(1万€)
无过滤	1,210	37.6%	1.02	+1,442 €
ML% ≥ 45%	661	54.3%	1.91 □	+27,542 €
ML% ≥ 50%	504	59.7%	2.27 □□	+25,842 €
ML% ≥ 55%	327	67.3%	2.80	+19,242 €
ML% ≥ 60%	170	78.8%	3.59	+9,340 €
ML% ≥ 70%	71	88.7%	1.64	+514 €

11. 量化金融与随机计算

6个量化金融函数已集成到扫描器和ML管道中。这些函数计算风险、波动率 and 市场体制指标：

函数	目的	ML特征	v18排名
Ornstein-Uhlenbeck	检测价格相对于长期均值的过度延伸。得分0-100%:>80% = 可能反转。	ou_score	未测试
Hurst指数	分类市场体制:H < 0.45 = 震荡, H > 0.55 = 趋势。	hurst_h	未测试
GARCH(1,1)	建模条件波动率。4个特征:持续性、长期波动率、半衰期、预测波动率。	garch_persistence	#14
Monte Carlo	模拟10,000条价格路径。计算P(SL触及)和P(TP触及)。	mc_p_sl, mc_p_tp	未测试
Kelly准则	按资产类型优化头寸规模。收益=0.00(与已优化的SL/TP冗余)。	kelly_full	#124
FTMO破产	自营公司账户在100笔模拟交易中10%回撤的破产概率。	ftmo_ruin_prob	未测试

12. 按资产类型的Kelly准则

Kelly准则使用来自回测的实际胜率, 而非统一的win_rate=0.40。结果:

资产类型	胜率	Kelly满额	Kelly半额
DXY	73%	46.0%	23.0%
XAUUSD	42%	13.0%	6.5%
外汇	31%	3.5%	1.8%
指数	38%	7.0%	3.5%
加密货币	35%	2.5%	1.3%

13. 简化的MLPredictor(2026年7月19日)

MLPredictor之前加载同一模型5次(每个预测方法1次)。2026年7月19日的简化将其减少为单次joblib.load()。结果:删除30行代码, 加载时间除以4, 冗余的_ASSET_MODEL_MAP被消除。v18模型现在是在所有资产类型的唯一统一模型。

14. 错误修复 — 2026年7月19日

2026年7月19日的彻底代码审查识别并修复了4个错误:

#	严重性	文件	描述
1	严重	diamond_scanner.py	ml_available中return后的死代码 — 符号从未被激活。initialize()未设置_initialized = True。
2	高	stochastic_analytics.py	GARCH: 如果未找到有效的 (α, β) 对, 使用虚构参数的静默回退。添加了best_ll <= -1e100检查。
3	中	stochastic_analytics.py	Monte Carlo: 对数收益率波动率回退已计算但在OU sigma为零时从未使用。
4	低	stochastic_analytics.py	FTMO: pnl_avg/roi_avg在try块外 — 如果在计算和返回之间添加代码, 有NameError风险。

15. 最终决策矩阵

决策矩阵结合了Diamond分数、ML%、安全护栏和宏观背景：

条件	行动	理由
ML% ≥ 60% + TKx HI 对齐 + 无HIGH NEWS	☐ GO	三重确认: ML、一目均衡表、宏观
ML% ≥ 60% + TKx HI 冲突	☐ 等待	ML被T/K HI交叉矛盾
ML% 50-60% + TKx HI 对齐	☐ 谨慎	适度的ML优势, 需要额外确认
ML% < 50%	☐ NO GO	模型无统计优势
HIGH NEWS DAY + BULL	☐ 等待	宏观日(≥2个重大事件)打破相关性
周三(FOMC日)+ 外汇	☐ 高风险	周三显示统计上独特的行为

16. 结论与展望

InelidaMarketScan项目证明, 基于一目均衡表特征训练的ML管道可以产生统计上显著的优势(OOS PF = 1.173)。2026年7月19日修复的4个关键错误增强了实时扫描器的可靠性。27个连续模型揭示了关键洞察: 一目均衡表特征比ICT模式更具预测性, 资产类型(is_dxy, is_forex)是主要预测因子, 随机计算函数(GARCH、OU、Hurst)提供补充的波动率信号。下一性能飞跃需要更大的数据集(2024-2026, 每资产>5,000笔交易)以及可能的深度学习架构(LSTM、Transformer)。

参考文献

- Huddleston, M. — Inner Circle Trader (ICT) Concepts, 2022-2024
- Chen, T. & Guestrin, C. — XGBoost: A Scalable Tree Boosting System, KDD 2016
- Bollerslev, T. — Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics, 1986
- Hurst, H.E. — Long-term Storage Capacity of Reservoirs, Trans. ASCE, 1951
- Uhlenbeck, G.E. & Ornstein, L.S. — On the Theory of Brownian Motion, Physical Review, 1930
- Kelly, J.L. — A New Interpretation of Information Rate, Bell System Technical Journal, 1956
- Hosoda, G. — Ichimoku Kinko Hyo, 1968

