

InelidaMarketScan — 機械学習による金融市場の科学的分析

定量ファイナンス研究: ICT、一目均衡表、XGBoostと確率解析 著者 Didier Vally / Reuniware Systems バージョン 1.0.6 — 2026-07-19 □ PyPI: <https://pypi.org/project/inelida-marketscan> | GitHub: <https://github.com/reuniware/InelidaMarketScanner>

要約

本書は、アルゴリズム取引に適用された6ヶ月間の定量ファイナンス研究の成果を提示する。InelidaMarketScanフレームワークは、ICT（インナーサークルトレーダー）テクニカル分析とマルチタイムフレーム一目均衡表を、XGBoost機械学習パイプライン（123特徴量）と確率解析関数（Ornstein-Uhlenbeck, Hurst, GARCH, Monte Carlo, Kelly）と組み合わせる。v18モデルは1,210件の過去取引で64.0%の交差検証精度を達成。27の連続モデルの分析により、一目均衡表の特徴量（Kijun/Tenkan距離、T/Kクロス）がICTパターン（FVG, Order Blocks）よりも予測力が高いことが明らかになった。ML% $\geq 50\%$ フィルタリングによる資本シミュレーションではインサンプルで利益因子2.27を示し、アウトオブサンプル検証では上限1.173を特定した。本書はICT、一目均衡表、機械学習の交点に関する包括的な科学的リファレンスである。

1. はじめに

1.1 背景と問題提起

アルゴリズム取引は、金融時系列における繰り返しパターンの識別に依存している。Michael Huddlestonによって開発されたICT（インナーサークルトレーダー）概念は、機関投資家の市場行動を説明する：流動性スweep、フェアバリューギャップ、オーダーブロック、市場構造シフト。並行して、一目均衡表指標は市場均衡のマルチタイムフレームの読み取りを提供する。本プロジェクトは、自動スキャナーとXGBoostモデルを通じてこれら2つのパラダイムの融合を探索する。

1.2 研究目的

- 8つの時間枠（M5からMN）でICTと一目均衡表を組み合わせた自動スキャナーを開発する。
- スキャナーから抽出された123の名前付き特徴量を持つMLパイプラインを構築する。
- 過去バックテスト（2025-2026）とアウトオブサンプルテストで予測力を検証する。
- 定量ファイナンス関数（OU, Hurst, GARCH, Monte Carlo, Kelly）を統合する。
- 実条件下で利益因子 > 1.5 を達成する。

2. ICT理論的基礎

ICT概念は、ポジションを蓄積するために価格水準を操作する金融機関（銀行、ヘッジファンド）の行動をモデル化する。以下の8つの概念がスキャナーによって自動検出される：

ICI概念	説明と検出
流動性スweep	反転前にストップを発動させるためのキーレベル(アジア高値/安値、ロンドン高値/安値)の短時間の突破。H1で検出し反転確認。
フェアバリューギャップ(FVG)	価格が取引されなかった3本の連続ローソク足間のギャップ。6つの時間枠(M5からD1)で検出し、埋め戻し率を表示。
オーダーブロック(OB)	方向性インパルスの前の最後のローソク足 — 機関が注文を出したゾーン。ATR(14)を用いて6つの時間枠で検出。
ブレーカーブロック(BRK)	破られたオーダーブロックが反対のサポート/レジスタンスに反転。既存のOBから派生。
市場構造シフト(MSS/BOS)	市場構造の変化:6つの時間枠でのストラクチャーブレイクとキャラクターチェンジ。
出来高HVN/LVN	MT5のティック出来高に基づく高出来高ノードと低出来高ノード。6つの時間枠で検出。
リバーサルパイプライン	スweep → 反転 → 確認の完全なチェーンとスweep後バーカウンター。
フィボナッチ拡張	TPターゲットのためのアジアレンジに基づく+1.618から+4.0の拡張。

3. マルチTF一目均衡表アーキテクチャ

Diamondスキャナーは8つの時間枠を同時に分析する。各TFについて5つの一目均衡表コンポーネントが計算される:

時間枠	Tenkan	Kijun	Kumo	T/K Cross	Flat History
M5	☐	☐	—	☐	—
M15	☐	☐	—	☐	—
M30	☐	☐	—	☐	—
H1	☐	☐	☐	☐	☐
H4	☐	☐	☐	☐	☐
D1	☐	☐	☐	☐	☐
W1	☐	☐	☐	☐	☐
MN	☐	☐	☐	☐	☐

4. Diamondスキャナーアーキテクチャ

Diamondスキャナーは10カテゴリーで109のスコアリング基準を集約する:

カテゴリー	基準数	例
一目均衡表 H1-D1	24	Kijun/Tenkan距離、T/Kクロス、Kumo位置、フラットバー
一目均衡表 W1-MN	8	Kijun W1/MN、Kumo W1/MN、雲の色、ダブルメモリー
FVG(6TF)	12	M5からD1のFVGベア/ブル、埋め戻し%、時間優先度
オーダーブロック(6TF)	12	M5からD1のOBベア/ブル、位置上/下/同
MSS/BOS(6TF)	18	M5からD1のレジーム、BOS、CHoCH
出来高HVN/LVN(6TF)	12	M5からD1の出来高スパイク、HVN、LVN
ブレーカーブロック(6TF)	12	M5からD1のBRKベア/ブル
アジア/ロンドンスイープ	4	AH/ALスイープ、ロンドンH/Lスイープ
リバーサルパイプライン	5	スイープ→反転→確認、スイープ後FVG、再テスト
M30/M15/M5 Kijun	2	Kijun M30/M15/M5方向性クロス

最大スコア: 109 基準 (MNなしで107)

5. 機械学習パイプライン

MLパイプラインは3段階でスキャナー結果を勝率に変換する:

段階	ツール	詳細
1	ml_labeler.py	Diamondスキャンからの特徴量抽出 (DB/手動CSV/バックテスト)。outcome列付きラベル付きCSVを生成 (1=勝ち、0=負け)。 5分割交差検証、不均衡クラス用
2	ml_trainer.py	scale_pos_weight、ハイパーパラメータグリッドサーチによるXGBoost訓練。 リアルタイム推論:
3	ml_predictor.py	extract_features() → predict_proba() → ML%をライブスキャンに表示。簡素化されたMLPredictor: 5回ではなく1回のjoblib.load。

6. XGBoostモデルv18 — 設定

v18モデルはアクティブな本番モデル。GARCHと確率特徴量を含む123の名前付き特徴量で1,210件の過去取引に訓練。

ハイパーパラメータ	値
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	自動(勝敗比に基づく)

123の名前付き特徴量には以下が含まれる:スコア/バイアス(6)、Kijun 6TF(18)、Tenkan 6TF(18)、T/Kクロス(6)、Kumo(6)、M5/M15/M30クロス(12)、フラットバー(2)、アクティブFVG(1)、オーダーブロック(6)、スweep(5)、出来高(6)、MSS/BOS(6)、プレーカーブロック(6)、圧縮(1)、確率(8:OU, Hurst, Monte Carlo, GARCH)、時間(3)、資産タイプ(3)。

7. モデル履歴

プロジェクト開始以来、27のXGBoostモデルが訓練・評価された。主要バージョンの履歴は以下の通り:

モデル	データ	特徴量	取引	CV	OOS PF
v1	6主要ペア	36	852	61.0%	—
v4	13シンボル	111	1,210	62.9%	1.26*
v7	13シンボル	25ベスト	2,923	60.1%	1.173
v13	4モデル/資産	25	2,923	53-64%	4.000
v15	2025-2026	25 + OU/Hurst	5,009	59.8%	—
v18	13シンボル	123 (GARCH)	1,210	64.0%	0.503 F1

8. 特徴量重要度 (v4)

特徴量重要度分析により、どの基準がモデル決定で最も重みを持つかが明らかになる:

順位	特徴量	ゲイン	洞察
1	is_dxy	14.2	ドルは外国為替ペアとは根本的に異なるロジックに従う
2	rr	10.0	理論的リスク/リワード比が実際の結果を予測する
3	tkx_h1	9.2	H1 Tenkan/Kijunクロスが最良の単一予測因子
4	kj_h1	6.9	H1 Kijunまでの距離が短期均衡を捉える
5	s1	6.4	ストップロスレベルはテイクプロフィットと同様に重要
6	day_wednesday	6.4	水曜日 (FOMC) は統計的に異なる行動を示す
7	d_kj_h4	6.3	H4 Kijunまでの距離が基調トレンドを示す
8	d_kj_d1	6.0	D1 Kijunまでの距離がマクロコンテキストを固定する
9	tkx_h4	6.0	H4 T/Kクロスがセッショントレンドを確認する
10	is_forex	5.9	FXペアは指数/金属とは異なる独自の行動を持つ

9. アウトオブサンプル検証

OOS検証 (2026年1月-6月訓練、2026年7月テスト)が決定的テストである。v4モデルはML% ≥ 50%閾値でインサンプルPF 2.27を示すが、これにはデータリークが含まれる(モデルは訓練中にテストデータの80%を見た)。v7の真のOOS PFは1.173 — 統計的に有意だが控えめな優位性。v13モデル(資産別)は15取引でPF=4.000を達成するが、サンプルが小さすぎて信頼できない。結論: MLは真の優位性 (PF > 1.0)を提供するが、統一アーキテクチャでの現在の上限は約1.17である。

10. 資本シミュレーション — ML%フィルタリング

初期資本10,000ユーロ、取引ごとリスク1%で1,210件の過去取引をシミュレーション:

ML%閾値	取引	勝率	PF(RR)	損益 (1万€)
フィルターなし	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. 定量ファイナンスと確率解析

6つの定量ファイナンス関数がスキャナーとMLパイプラインに統合された。これらの関数はリスク、ボ

ラティリティ、市場レジームの指標を計算する：

関数	目的	ML特徴量	v18順位
Ornstein-Uhlenbeck	長期平均からの価格の過剰延伸を検出。スコア0-100%:>80% = 反転の可能性。	ou_score	未テスト
Hurst指数	市場レジームを分類:H < 0.45 = レンジ、H > 0.55 = トренд。	hurst_h	未テスト
GARCH(1,1)	条件付きボラティリティをモデル化。4特徴量:持続性、長期ボラティリティ、半減期、予測ボラティリティ。	garch_persistence	#14
Monte Carlo	10,000の価格パスをシミュレーション。P(SL到達)とP(TP到達)を計算。	mc_p_sl, mc_p_tp	未テスト
Kelly基準	資産タイプ別の最適ポジションサイズ。ゲイン=0.00(最適化済みSL/TPと重複)。	kelly_full	#124
FTMO破産	10%ドローダウンでのプロップファーム口座の破産確率。100取引シミュレーション。	ftmo_ruin_prob	未テスト

12. 資産タイプ別 Kelly基準

Kelly基準は一律のwin_rate=0.40ではなく、バックテストからの実際の勝率を使用する。結果：

資産タイプ	勝率	Kellyフル	Kellyハーフ
DXY	73%	46.0%	23.0%
XAUUSD	42%	13.0%	6.5%
外国為替	31%	3.5%	1.8%
指数	38%	7.0%	3.5%
暗号通貨	35%	2.5%	1.3%

13. 簡素化されたMLPredictor(2026年7月19日)

MLPredictorは以前、同じモデルを5回ロードしていた(予測メソッドごとに1回)。2026年7月19日の簡素化により、単一のjoblib.load()に削減された。結果:30行のコード削除、ロード時間が4分の1に、冗長な_ASSET_MODEL_MAPが排除された。v18モデルは現在、すべての資産タイプに対する唯一の統一モデルである。

14. バグ修正 — 2026年7月19日

2026年7月19日の徹底的なコードレビューにより4つのバグが特定され修正された：

#	重大度	ファイル	説明
1	致命的	diamond_scanner.py	ml_available内のreturn後のデッドコード — シンボルが一度もアクティブ化されなかった。initialize()が_initialized = Trueを設定していなかった。
2	高	stochastic_analytics.py	GARCH: 有効な(α, β)ペアが見つからない場合、架空のパラメータでサイレントフォールバック。best_ll <= -1e100チェックを追加。
3	中	stochastic_analytics.py	Monte Carlo: 対数収益率ボラティリティフォールバックが計算されていたが、OUSigmaがゼロの場合に使用されていなかった。
4	低	stochastic_analytics.py	FTMO: pnl_avg/roi_avgがtryブロック外 — 計算とreturnの間にコードが追加された場合のNameErrorリスク。

15. 最終決定マトリックス

決定マトリックスはDiamondスコア、ML%、安全装置、マクロコンテキストを組み合わせる：

条件	アクション	根拠
ML% $\geq$ 60% + TKx HI 整合 + HIGH NEWSなし	☐ GO	三重確認: ML、一目均衡表、マクロ
ML% $\geq$ 60% + TKx HI 競合	☐ 待機	MLがT/K HIクロスと矛盾
ML% 50-60% + TKx HI 整合	☐ 注意	中程度のML優位性、追加確認が必要
ML% < 50%	☐ NO GO	モデルに統計的優位性なし
HIGH NEWS DAY + BULL	☐ 待機	マクロ日 (≥ 2 主要イベント) は相関を破壊
水曜日 (FOMC日) + 外国為替	☐ 高リスク	水曜日は統計的に異なる行動を示す

16. 結論と展望

InelidaMarketScanプロジェクトは、一目均衡表特徴量で訓練されたMLパイプラインが統計的に有意な優位性 (OOS PF = 1.173) を生成できることを実証する。2026年7月19日に修正された4つの致命的バグは、ライブスキャナーの信頼性を強化する。27の連続モデルは重要な洞察を明らかにした: 一目均衡表特徴量はICTパターンよりも予測力が高く、資産タイプ (is_dxy、is_forex) が主要予測因子であり、確率解析関数 (GARCH, OU, Hurst) が補完的なボラティリティシグナルを提供する。次のパフォーマンス飛躍には、より大きなデータセット (2024-2026、資産あたり>5,000取引) と潜在的にはディープラーニングアーキテクチャ (LSTM、Transformer) が必要である。

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

定量ファイナンス研究著作。引用による複製許可。Didier Vally / Reuniware
Systems — InelidaMarketScan v1.0.6 2026-07-19