

InelidaMarketScan — Analisi Scientifica dei Mercati Finanziari con Machine Learning

Ricerca in Finanza Quantitativa: ICT, Ichimoku, XGBoost e Calcolo Stocastico Di Didier Vally / Reuniware Systems Versione 1.0.6 — 2026-07-19
□ PyPI: <https://pypi.org/project/inelida-marketscan> | GitHub: <https://github.com/reuniware/InelidaMarketScanner>

Sommario

Questo documento presenta i risultati di 6 mesi di ricerca in finanza quantitativa applicata al trading algoritmico. Il framework InelidaMarketScan combina l'analisi tecnica ICT (Inner Circle Trader) e Ichimoku multi-timeframe con una pipeline di Machine Learning XGBoost (123 feature) e funzioni di calcolo stocastico (Ornstein-Uhlenbeck, Hurst, GARCH, Monte Carlo, Kelly). Il modello v18 raggiunge il 64,0% di accuratezza in cross-validazione su 1.210 trade storici. L'analisi di 27 modelli successivi rivela che le feature Ichimoku (distanze Kijun/Tenkan, incroci T/K) sono più predittive dei pattern ICT (FVG, Order Blocks). La simulazione di capitale con filtraggio $ML\% \geq 50\%$ mostra un profit factor di 2,27 su dati in-sample, mentre la validazione out-of-sample identifica un tetto a 1,173. Questo documento costituisce un riferimento scientifico completo sull'intersezione tra ICT, Ichimoku e Machine Learning.

1. Introduzione

1.1 Contesto e Problematica

Il trading algoritmico si basa sull'identificazione di pattern ricorrenti nelle serie temporali finanziarie. I concetti ICT (Inner Circle Trader) sviluppati da Michael Huddleston descrivono il comportamento istituzionale dei mercati: sweep di liquidità, Fair Value Gaps, Order Blocks, Market Structure Shifts. Parallelamente, l'indicatore Ichimoku Kinko Hyo fornisce una lettura multi-timeframe dell'equilibrio di mercato. Questo progetto esplora la fusione di questi due paradigmi attraverso uno scanner automatizzato con modello XGBoost.

1.2 Obiettivi di Ricerca

1. Sviluppare uno scanner automatico che combini ICT e Ichimoku su 8 timeframe (M5 a MN).
2. Costruire una pipeline ML con 123 feature nominali estratte dallo scanner.
3. Validare il potere predittivo tramite backtest storico (2025-2026) e test out-of-sample.
4. Integrare funzioni di finanza quantitativa (OU, Hurst, GARCH, Monte Carlo, Kelly).
5. Raggiungere un profit factor $> 1,5$ in condizioni reali.

2. Fondamenti Teorici ICT

I concetti ICT modellano il comportamento delle istituzioni finanziarie (banche, hedge fund) che manipolano i livelli di prezzo per accumulare posizioni. Gli 8 concetti seguenti sono rilevati automaticamente dallo scanner:

Concetto ICT	Descrizione e Rilevamento
Sweep di Liquidità	Rottura breve di un livello chiave (Asian High/Low, London High/Low) per attivare gli stop prima dell'inversione. Rilevato su H1 con conferma di inversione.
Fair Value Gaps (FVG)	Divario tra 3 candele consecutive dove il prezzo non ha negoziato. Rilevato su 6 TF (M5 a D1) con percentuale di mitigazione.
Order Blocks (OB)	Ultima candela prima di un impulso direzionale — zona dove le istituzioni hanno piazzato ordini. Rilevato via ATR(14) su 6 TF.
Breaker Blocks (BRK)	Order Block rotto che si inverte in supporto/resistenza opposto. Derivato dagli OB esistenti.
Market Structure Shift (MSS/BOS)	Cambiamento nella struttura di mercato: Break of Structure e Change of Character su 6 TF.
Volume HVN/LVN	High Volume Node e Low Volume Node basati sul tick volume MT5. Rilevato su 6 TF.
Reversal Pipeline	Catena completa sweep → inversione → conferma con contatore di barre post-sweep.
Estensioni di Fibonacci	Estensioni +1,618 a +4,0 basate sul range asiatico per obiettivi di TP.

3. Architettura Ichimoku Multi-TF

Lo scanner Diamond analizza 8 timeframe simultaneamente. Per ogni TF, vengono calcolati 5 componenti Ichimoku:

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

4. Architettura dello Scanner Diamond

Il Diamond Scanner aggrega 109 criteri di scoring in 10 categorie:

Categoria	Criteri	Esempi
Ichimoku HI-D1	24	Distanze Kijun/Tenkan, incroci T/K, posizioni Kumo, Flat bar
Ichimoku W1-MN	8	Kijun W1/MN, Kumo W1/MN, colore nuvola, Doppia memoria
FVG (6 TF)	12	FVG bear/bull M5 a D1, % mitigazione, priorità temporale
Order Blocks (6 TF)	12	OB bear/bull M5 a D1, posizione SOPRA/SOTTO/A
MSS/BOS (6 TF)	18	Regime, BOS, CHoCH su M5 a D1
Volume HVN/LVN (6 TF)	12	Picco volume, HVN, LVN su M5 a D1
Breaker Blocks (6 TF)	12	BRK bear/bull M5 a D1
Sweep Asian/London	4	Sweep AH/AL, sweep London H/L
Reversal Pipeline	5	Sweep→inversione→conferma, FVG post-sweep, retest
Kijun M30/M15/M5	2	Incrocio direzionale Kijun M30/M15/M5

Punteggio massimo: 109 criteri (107 senza MN)

5. Pipeline di Machine Learning

La pipeline ML trasforma i risultati dello scanner in probabilità di vincita attraverso 3 fasi:

Fase	Strumento	Dettagli
1	ml_labeler.py	Estrazione feature dalla scansione Diamond (DB / CSV manuale / backtest). Genera CSV etichettato con colonna outcome (1=vinto, 0=perso).
2	ml_trainer.py	Addestramento XGBoost con cross-validazione 5-fold, scale_pos_weight per classi sbilanciate, grid search iperparametri.
3	ml_predictor.py	Inferenza in tempo reale: extract_features() → predict_proba() → ML% visualizzato nella scansione live. MLPredictor semplificato: 1 joblib.load invece di 5.

6. Modello XGBoost v18 — Configurazione

Il modello v18 è il modello di produzione attivo. Addestrato su 1.210 trade storici con 123 feature nominali (incluse GARCH e stocastico).

Iperparametro	Valore
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	auto (basato su rapporto win/loss)

123 feature nominali includono: punteggi/bias (6), Kijun 6 TF (18), Tenkan 6 TF (18), incroci T/K (6), Kumo (6), incroci M5/M15/M30 (12), Flat bar (2), FVG attivi (1), Order Blocks (6), Sweep (5), Volume (6), MSS/BOS (6), Breaker Blocks (6), Compressione (1), Stocastico (8: OU, Hurst, Monte Carlo, GARCH), Temporale (3), Tipo di asset (3).

7. Cronologia dei Modelli

27 modelli XGBoost sono stati addestrati e valutati dall’inizio del progetto. Ecco la cronologia delle versioni principali:

Modello	Dati	Feature	Trade	CV	OOS PF
v1	6 principali	36	852	61.0%	—
v4	13 simboli	111	1.210	62.9%	1.26*
v7	13 simboli	25 migliori	2.923	60.1%	1.173
v13	4 modelli/asset	25	2.923	53-64%	4.000
v15	2025-2026	25 + OU/Hurst	5.009	59.8%	—
v18	13 simboli	123 (GARCH)	1.210	64.0%	0.503 F1

8. Importanza delle Feature (v4)

L’analisi dell’importanza delle feature rivela quali criteri pesano di più nelle decisioni del modello:

Rank	Feature	Guadagno	Insight
1	is_dxy	14.2	Il dollaro segue una logica radicalmente diversa dalle coppie forex
2	rr	10.0	Il rapporto risk/reward teorico predice il risultato reale
3	tkx_h1	9.2	L'incrocio Tenkan/Kijun H1 è il miglior predittore singolo
4	kj_h1	6.9	La distanza dalla Kijun H1 cattura l'equilibrio a breve termine
5	s1	6.4	Il livello di stop-loss è importante quanto il take-profit
6	day_wednesday	6.4	I mercoledì (FOMC) mostrano un comportamento statisticamente distinto
7	d_kj_h4	6.3	La distanza dalla Kijun H4 indica il trend sottostante
8	d_kj_d1	6.0	La distanza dalla Kijun D1 ancora il contesto macro
9	tkx_h4	6.0	L'incrocio T/K H4 conferma il trend di sessione
10	is_forex	5.9	Le coppie FX hanno un comportamento proprio vs indici/metalli

9. Validazione Out-of-Sample

La validazione OOS (addestramento Gennaio-Giugno 2026, test Luglio 2026) è il test decisivo. Il modello v4 mostra un PF in-sample di 2,27 alla soglia $ML\% \geq 50\%$, ma questo include data leakage (il modello ha visto l'80% dei dati di test durante l'addestramento). Il vero PF OOS del v7 è 1,173 — un edge statisticamente significativo ma modesto. Il modello v13 (per asset) raggiunge PF=4,000 su 15 trade, ma il campione è troppo piccolo per essere affidabile. Conclusione: il ML fornisce un edge reale ($PF > 1,0$), ma il tetto attuale è ~1,17 con architettura unificata.

10. Simulazione di Capitale — Filtraggio ML%

Simulazione su 1.210 trade storici con capitale iniziale di 10.000 € e rischio 1% per trade:

Soglia ML%	Trade	Tasso di Successo	PF (RR)	P&L (10k€)
Nessun filtro	1.210	37.6%	1.02	+1.442 €
ML% $\geq$ 45%	661	54.3%	1.91 $\square$	+27.542 €
ML% $\geq$ 50%	504	59.7%	2.27 $\square\square$	+25.842 €
ML% $\geq$ 55%	327	67.3%	2.80	+19.242 €
ML% $\geq$ 60%	170	78.8%	3.59	+9.340 €
ML% $\geq$ 70%	71	88.7%	1.64	+514 €

11. Finanza Quantitativa e Calcolo Stocastico

6 funzioni di finanza quantitativa sono state integrate nello scanner e nella pipeline ML. Queste funzioni calcolano metriche di rischio, volatilità e regime di mercato:

Funzione	Scopo	Feature ML	Rank v18
Ornstein-Uhlenbeck	Rileva la sovraestensione del prezzo dalla media a lungo termine. Score 0-100%: >80% = probabile inversione.	ou_score	Non testato
Esponente di Hurst	Classifica il regime di mercato: $H < 0.45$ = RANGE, $H > 0.55$ = TREND.	hurst_h	Non testato
GARCH(1,1)	Modella la volatilità condizionale. 4 feature: persistenza, vol_lungo_termine, emivita, vol_previsionale.	garch_persistence	#14
Monte Carlo	Simula 10.000 traiettorie di prezzo. Calcola P(SL raggiunto) e P(TP raggiunto).	mc_p_sl, mc_p_tp	Non testato
Criterio di Kelly	Dimensionamento ottimale delle posizioni per tipo di asset. Guadagno=0.00 (ridondante con SL/TP già ottimizzati).	kelly_full	#124
Rovina FTMO	Probabilità di rovina di un conto prop firm con drawdown 10% su 100 trade simulati.	ftmo_ruin_prob	Non testato

12. Criterio di Kelly per Tipo di Asset

Il Criterio di Kelly utilizza tassi di successo reali per tipo di asset (dal backtest) invece di un win_rate=0.40 uniforme. Risultati:

Tipo di Asset	Tasso di Successo	Kelly Pieno	Kelly Metà
DXY	73%	46.0%	23.0%
XAUUSD	42%	13.0%	6.5%
Forex	31%	3.5%	1.8%
Indici	38%	7.0%	3.5%
Cripto	35%	2.5%	1.3%

13. MLPredictor Semplificato (19/07/2026)

Il MLPredictor caricava precedentemente lo stesso modello 5 volte (1× per metodo di predizione). La semplificazione del 19/07/2026 l’ha ridotto a un singolo `joblib.load()`. Risultato: 30 righe di codice eliminate, tempo di caricamento diviso per 4, e la `_ASSET_MODEL_MAP` ridondante eliminata. Il modello v18 è ora il modello unificato unico per tutti i tipi di asset.

14. Correzione di Bug — 19/07/2026

Una revisione approfondita del codice il 19/07/2026 ha identificato e corretto 4 bug:

#	Gravità	File	Descrizione
1	CRITICO	diamond_scanner.py	Codice morto dopo <code>return</code> in <code>ml_available</code> — i simboli non venivano mai attivati. <code>initialize()</code> non impostava <code>_initialized = True</code> .
2	ALTO	stochastic_analytics.py	GARCH: se nessuna coppia (α, β) valida trovata, fallback silenzioso con parametri inventati. Aggiunto controllo <code>best_ll <= -1e100</code> .
3	MEDIO	stochastic_analytics.py	Monte Carlo: il fallback di volatilità (log-returns) era calcolato ma mai usato quando il sigma OU era zero.
4	BASSO	stochastic_analytics.py	FTMO: <code>pn1_avg/roi_avg</code> fuori dal blocco <code>try</code> — rischio di <code>NameError</code> se codice aggiunto tra calcolo e <code>return</code> .

15. Matrice Decisionale Finale

La matrice decisionale combina punteggio Diamond, ML%, guardrail e contesto macro:

Condizione	Azione	Giustificazione
ML% $\geq$ 60% + TKx HI allineato + no HIGH NEWS	☐ GO	Tripla conferma: ML, Ichimoku, macro
ML% $\geq$ 60% + TKx HI conflitto	☐ ATTENDERE	Il ML è contraddetto dall'incrocio T/K HI
ML% 50-60% + TKx HI allineato	☐ CAUTELA	Edge ML moderato, necessita conferma aggiuntiva
ML% < 50%	☐ NO GO	Il modello non ha edge statistico
HIGH NEWS DAY + BULL	☐ ATTENDERE	I giorni macro (≥ 2 eventi maggiori) rompono le correlazioni
Mercoledì (giorno FOMC) + Forex	☐ Rischio Alto	I mercoledì mostrano comportamento statisticamente distinto

16. Conclusione e Prospettive

Il progetto InelidaMarketScan dimostra che una pipeline ML addestrata su feature Ichimoku può generare un edge statistico significativo (PF OOS = 1,173). I 4 bug critici corretti il 19/07/2026 rafforzano l'affidabilità dello scanner live. I 27 modelli successivi hanno rivelato intuizioni chiave: le feature Ichimoku sono più predittive dei pattern ICT, il tipo di asset (is_dxy, is_forex) è un predittore principale, e le funzioni di calcolo stocastico forniscono segnale complementare di volatilità. Il prossimo salto di prestazioni richiederà un dataset più grande (2024-2026, >5.000 trade/asset) e potenzialmente architetture deep learning.

Riferimenti

- 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

Opera di ricerca in finanza quantitativa. Riproduzione consentita con citazione.

☐ Comunità

Unisciti al server Discord **Trading Pro** per connetterti con la community:
<https://discord.gg/u9JApXqhXy>