

InelidaMarketScan — Wissenschaftliche Analyse der Finanzmärkte mit Machine Learning

Quantitative Finanzforschung: ICT, Ichimoku, XGBoost & Stochastische Analysis Von **Didier Vally / Reuniware Systems**
Version 1.0.6 — 2026-07-19 □ PyPI: <https://pypi.org/project/inelida-marketscan> | GitHub: <https://github.com/reuniware/InelidaMarketScanner>

Zusammenfassung

Dieses Dokument präsentiert die Ergebnisse von 6 Monaten quantitativer Finanzforschung im algorithmischen Handel. Das InelidaMarketScan-Framework kombiniert technische ICT-Analyse (Inner Circle Trader) und Multi-Timeframe-Ichimoku mit einer XGBoost Machine-Learning-Pipeline (123 Features) und stochastischen Analysefunktionen (Ornstein-Uhlenbeck, Hurst, GARCH, Monte Carlo, Kelly). Modell v18 erreicht 64,0% Kreuzvalidierungsgenauigkeit bei 1.210 historischen Trades. Die Analyse von 27 aufeinanderfolgenden Modellen zeigt, dass Ichimoku-Features (Kijun/Tenkan-Abstände, T/K-Kreuzungen) prädiktiver sind als ICT-Muster (FVG, Order Blocks). Die Kapitalsimulation mit $ML\% \geq 50\%$ -Filterung zeigt einen Profitfaktor von 2,27 bei In-Sample-Daten, während die Out-of-Sample-Validierung eine Obergrenze von 1,173 identifiziert. Dieses Dokument dient als umfassende wissenschaftliche Referenz zur Schnittstelle von ICT, Ichimoku und Machine Learning.

1. Einleitung

1.1 Kontext und Problemstellung

Algorithmischer Handel basiert auf der Identifikation wiederkehrender Muster in Finanzzeitreihen. Die von Michael Huddleston entwickelten ICT-Konzepte (Inner Circle Trader) beschreiben institutionelles Marktverhalten: Liquiditätssweeps, Fair Value Gaps, Order Blocks, Market Structure Shifts. Parallel dazu liefert der Ichimoku Kinko Hyo Indikator eine Multi-Timeframe-Betrachtung des Marktgleichgewichts. Dieses Projekt erforscht die Fusion dieser beiden Paradigmen durch einen automatisierten Scanner mit XGBoost-Modell.

1.2 Forschungsziele

1. Entwicklung eines automatischen Scanners, der ICT und Ichimoku über 8 Timeframes (M5 bis MN) kombiniert.
2. Aufbau einer ML-Pipeline mit

123 benannten Features aus dem Scanner. 3. Validierung der Vorhersagekraft durch historisches Backtesting (2025-2026) und Out-of-Sample-Tests. 4. Integration quantitativer Finanzfunktionen (OU, Hurst, GARCH, Monte Carlo, Kelly). 5. Erreichen eines Profitfaktors > 1,5 unter realen Bedingungen.

2. Theoretische Grundlagen von ICT

ICT-Konzepte modellieren das Verhalten von Finanzinstitutionen (Banken, Hedgefonds), die Preisniveaus manipulieren, um Positionen aufzubauen. Die folgenden 8 Konzepte werden automatisch vom Scanner erkannt:

ICT-Konzept	Beschreibung und Erkennung
Liquiditätssweeps	Kurzer Bruch eines Schlüsselnieaus (Asian High/Low, London High/Low) zur Auslösung von Stops vor der Umkehr. Erkannt auf H1 mit Umkehrbestätigung.
Fair Value Gaps (FVG)	Lücke zwischen 3 aufeinanderfolgenden Kerzen, in der kein Handel stattfand. Erkannt auf 6 TFs (M5 bis D1) mit Mitigationsprozentsatz.
Order Blocks (OB)	Letzte Kerze vor einem gerichteten Impuls — Zone, in der Institutionen Orders platzierten. Erkannt via ATR(14) auf 6 TFs.
Breaker Blocks (BRK)	Gebrochener Order Block, der sich in entgegengesetzte Unterstützung/Widerstand umkehrt. Abgeleitet von bestehenden OBs.
Market Structure Shift (MSS/BOS)	Änderung der Marktstruktur: Break of Structure und Change of Character auf 6 TFs.
Volumen HVN/LVN	High Volume Node und Low Volume Node basierend auf MT5
Reversal Pipeline	Tick-Volumen. Erkannt auf 6 TFs. Vollständige Sweep → Umkehr → Bestätigungskette mit Post-Sweep-Bar-Zähler.
Fibonacci-Extensions	+1,618 bis +4,0 Extensions basierend auf der asiatischen Range für TP-Ziele.

3. Multi-TF Ichimoku-Architektur

Der Diamond-Scanner analysiert 8 Timeframes gleichzeitig. Für jeden TF werden 5 Ichimoku-Komponenten berechnet:

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

4. Diamond-Scanner-Architektur

Der Diamond-Scanner aggregiert 109 Scoring-Kriterien in 10 Kategorien:

Kategorie	Kriterien	Beispiele
Ichimoku H1-D1	24	Kijun/Tenkan-Abstände, T/K-Kreuzungen, Kumo-Positionen, Flat Bars
Ichimoku W1-MN	8	Kijun W1/MN, Kumo W1/MN, Wolkenfarbe, Doppelspeicher
FVG (6 TFs)	12	FVG bear/bull M5 bis D1, Mitigations-%, Zeitpriorität
Order Blocks (6 TFs)	12	OB bear/bull M5 bis D1, Position ÜBER/UNTER/AN
MSS/BOS (6 TFs)	18	Regime, BOS, CHoCH auf M5 bis D1
Volumen HVN/LVN (6 TFs)	12	Volumenspitze, HVN, LVN auf M5 bis D1
Breaker Blocks (6 TFs)	12	BRK bear/bull M5 bis D1
Asian/London Sweeps	4	AH/AL-Sweep, London H/L-Sweep
Reversal Pipeline	5	Sweep→Umkehr→Bestätigung, Post-Sweep-FVG, Retest
M30/M15/M5 Kijun	2	Kijun M30/M15/M5 Richtungskreuzung

Maximale Punktzahl: 109 Kriterien (107 ohne MN)

5. Machine-Learning-Pipeline

Die ML-Pipeline transformiert Scanner-Ergebnisse in Gewinnwahrscheinlichkeit durch 3 Stufen:

Stufe	Werkzeug	Details
1	<code>ml_labeler.py</code>	Feature-Extraktion aus Diamond-Scan (DB / manuelles CSV / Backtest). Erzeugt gelabeltes CSV mit Outcome-Spalte (1=gewonnen, 0=verloren).
2	<code>ml_trainer.py</code>	XGBoost-Training mit 5-facher Kreuzvalidierung, <code>scale_pos_weight</code> für unausgeglichene Klassen, Hyperparameter-Grid-Search.
3	<code>ml_predictor.py</code>	Echtzeit-Inferenz: <code>extract_features()</code> → <code>predict_proba()</code> → ML% im Live-Scan angezeigt. Vereinfachter MLPredictor: 1 <code>joblib.load</code> statt 5.

6. XGBoost-Modell v18 — Konfiguration

Modell v18 ist das aktive Produktionsmodell. Trainiert auf 1.210 historischen Trades mit 123 benannten Features (inklusive GARCH und Stochastik).

Hyperparameter	Wert
<code>max_depth</code>	6
<code>learning_rate</code>	0,03
<code>n_estimators</code>	200
<code>subsample</code>	0,8

Hyperparameter	Wert
min_child_weight	3
gamma	0,1
reg_alpha	0,5 (L1)
reg_lambda	1,0 (L2)
scale_pos_weight	auto (basierend auf Win/Loss-Verhältnis)

123 benannte Features umfassen: Scores/Bias (6), Kijun 6 TFs (18), Tenkan 6 TFs (18), T/K-Kreuzungen (6), Kumo (6), M5/M15/M30-Kreuzungen (12), Flat Bars (2), Aktive FVGs (1), Order Blocks (6), Sweeps (5), Volumen (6), MSS/BOS (6), Breaker Blocks (6), Kompression (1), Stochastik (8: OU, Hurst, Monte Carlo, GARCH), Zeitlich (3), Anlagetyp (3).

7. Modellhistorie

27 XGBoost-Modelle wurden seit Projektbeginn trainiert und evaluiert. Hier die Historie der Hauptversionen:

Modell	Daten	Features	Trades	CV	OOS PF
v1	6 Hauptpaare	36	852	61,0%	—
v4	13 Symbole	111	1.210	62,9%	1,26*
v7	13 Symbole	25 beste	2.923	60,1%	1,173
v13	4 Modelle/Anlage	25	2.923	53-64%	4,000
v15	2025-2026	25 + OU/Hurst	5.009	59,8%	—
v18	13 Symbole	123 (GARCH)	1.210	64,0%	0,503 F1

8. Feature-Wichtigkeit (v4)

Die Feature-Importance-Analyse zeigt, welche Kriterien am meisten Gewicht in den Modellentscheidungen haben:

Rang	Feature	Gewinn	Erkenntnis
1	is_dxy	14,2	Der Dollar folgt einer radikal anderen Logik als Forex-Paare

Rang	Feature	Gewinn	Erkenntnis
2	rr	10,0	Das theoretische Risk/Reward-Verhältnis sagt das tatsächliche Ergebnis voraus
3	tkx_h1	9,2	Die H1 Tenkan/Kijun-Kreuzung ist der beste Einzelprädiktor
4	kj_h1	6,9	Abstand zur H1 Kijun erfasst das kurzfristige Gleichgewicht
5	sl	6,4	Das Stop-Loss-Niveau ist ebenso wichtig wie der Take-Profit
6	day_wednesday	6,4	Mittwoch (FOMC) zeigen statistisch abweichendes Verhalten
7	d_kj_h4	6,3	Abstand zur H4 Kijun zeigt den zugrunde liegenden Trend an
8	d_kj_d1	6,0	Abstand zur D1 Kijun verankert den Makrokontext
9	tkx_h4	6,0	H4 T/K-Kreuzung bestätigt den Session-Trend

10	is_forex	5,9	FX-Paare haben eigenes Verhalten vs. In- dizes/Metalle
----	----------	-----	---

9. Out-of-Sample-Validierung

Die OOS-Validierung (Training Januar-Juni 2026, Test Juli 2026) ist der entscheidende Test. Modell v4 zeigt einen In-Sample-PF von 2,27 bei ML% $\geq$ 50%-Schwelle, aber dies enthält Data Leakage (das Modell sah 80% der Testdaten während des Trainings). Der wahre OOS-PF von v7 beträgt 1,173 — ein statistisch signifikanter, aber bescheidener Vorteil. Modell v13 (pro Anlagetyp) erreicht PF=4,000 bei 15 Trades, aber die Stichprobe ist zu klein für Zuverlässigkeit. Fazit: ML bietet einen echten Vorteil (PF > 1,0), aber die aktuelle Obergrenze liegt bei $\sim$ 1,17 mit einheitlicher Architektur.

10. Kapitalsimulation — ML%-Filterung

Simulation auf 1.210 historischen Trades mit Anfangskapital von 10.000 € und 1% Risiko pro Trade:

ML%-Schwelle	Trades	Gewinnrate	PF (RR)	G&V (10k€)
Kein Filter	1.210	37,6%	1,02	+1.442 €
ML% $\geq$ 45%	661	54,3%	1,91 □	+27.542 €
ML% $\geq$ 50%	504	59,7%	2,27 □□	+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. Quantitative Finanzanalyse & Stochastische Analysis

6 quantitative Finanzfunktionen wurden in Scanner und ML-Pipeline integriert. Diese Funktionen berechnen Risiko-, Volatilitäts- und Marktregime-Metriken:

Funktion	Zweck	ML-Feature	v18 Rang
Ornstein-Uhlenbeck	Erkennt Preisüberdehnung vom langfristigen Mittelwert. Score 0-100%: >80% = wahrscheinliche Umkehr.	ou_score	Nicht getestet
Hurst-Exponent	Klassifiziert das Marktregime: $H < 0,45$ = RANGE, $H > 0,55$ = TREND.	hurst_h	Nicht getestet
GARCH(1,1)	Modelliert bedingte Volatilität. 4 Features: Persistenz, Langfrist-Vol, Halbwertszeit, Prognose-Vol.	garch_persistence#14	
Monte Carlo	Simuliert 10.000 Preispfade. Berechnet P(SL erreicht) und P(TP erreicht).	mc_p_sl, mc_p_tp	Nicht getestet
Kelly-Kriterium	Optimale Positionsgröße pro Anlagetyp. Gewinn=0,00 (redundant mit optimiertem SL/TP).	kelly_full	#124
FTMO-Ruin	Ruinwahrscheinlichkeit für Prop-Firm-Konto mit 10% Drawdown über 100 simulierte Trades.	ftmo_ruin_prob	Nicht getestet

12. Kelly-Kriterium nach Anlagetyp

Das Kelly-Kriterium verwendet reale Gewinnraten pro Anlagetyp (aus dem Backtest) statt eines einheitlichen `win_rate=0,40`. Ergebnisse:

Anlagetyp	Gewinnrate	Kelly Voll	Kelly Halb
DXY	73%	46,0%	23,0%
XAUUSD	42%	13,0%	6,5%
Forex	31%	3,5%	1,8%
Indizes	38%	7,0%	3,5%
Krypto	35%	2,5%	1,3%

13. Vereinfachter MLPredictor (19.07.2026)

Der MLPredictor lud zuvor dasselbe Modell 5-mal (1× pro Vorhersagemethode). Die Vereinfachung vom 19.07.2026 reduzierte dies auf einen einzigen `joblib.load()`. Ergebnis: 30 Codezeilen entfernt, Ladezeit durch 4 geteilt, und die redundante `_ASSET_MODEL_MAP` eliminiert. Modell v18 ist nun das einzige einheitliche Modell für alle Anlagetypen.

14. Bugfixes — 19.07.2026

Eine gründliche Codeüberprüfung am 19.07.2026 identifizierte und behob 4 Bugs:

#	Schwere	Datei	Beschreibung
1	KRITISCH	<code>diamond_scan</code>	Im Code nach <code>return in ml_available</code> — Symbole wurden nie aktiviert. <code>initialize()</code> setzte <code>initialized</code> nicht auf <code>True</code> .

#	Schwere	Datei	Beschreibung
2	HOCH	stochastic_analysis.py	GARCH-Wenn kein gültiges (α, β) -Paar gefunden, stiller Fallback mit erfundenen Parametern. $best_ll \leq -1e10$ -Prüfung hinzugefügt.
3	MITTEL	stochastic_analysis.py	Monte-Carlo-Log-Return-Volatilitäts-Fallback wurde berechnet, aber nie verwendet, wenn OU-Sigma null war.
4	NIEDRIG	stochastic_analytics.py	FIMO pnl_avg/roi_avg außerhalb des try-Blocks — Risiko von NameError bei Code-Erweiterung.

15. Endgültige Entscheidungsmatrix

Die Entscheidungsmatrix kombiniert Diamond-Score, ML%, Sicherheitsregeln und Makrokontext:

Bedingung	Aktion	Begründung
ML% $\geq$ 60% + TKx H1 ausgerichtet + kein HIGH NEWS	□ GO	Dreifache Bestätigung: ML, Ichimoku, Makro
ML% $\geq$ 60% + TKx H1 Konflikt	□ WARTEN	ML wird durch T/K H1-Kreuzung widersprochen
ML% 50-60% + TKx H1 ausgerichtet	□ VORSICHT	Moderater ML-Vorteil, braucht zusätzliche Bestätigung

Bedingung	Aktion	Begründung
ML% < 50%	□ KEIN GO	Modell hat keinen statistischen Vorteil
HIGH NEWS DAY + BULL	□ WARTEN	Makro-Tage (≥ 2 Großereignisse) brechen Korrelationen
Mittwoch (FOMC-Tag) + Forex	□ Hohes Risiko	Mittwoche zeigen statistisch abweichendes Verhalten

16. Fazit und Ausblick

Das InelidaMarketScan-Projekt zeigt, dass eine ML-Pipeline mit Ichimoku-Features einen statistisch signifikanten Vorteil generieren kann (OOS PF = 1,173). Die 4 kritischen Bugs, die am 19.07.2026 behoben wurden, stärken die Zuverlässigkeit des Live-Scanners. Die 27 aufeinanderfolgenden Modelle offenbarten wichtige Erkenntnisse: Ichimoku-Features sind prädiktiver als ICT-Muster, der Anlagetyp ist ein Hauptprädiktor, und stochastische Funktionen liefern komplementäres Volatilitätssignal. Der nächste Leistungssprung erfordert einen größeren Datensatz (2024-2026, >5.000 Trades/Anlagetyp) und potenziell Deep-Learning-Architekturen (LSTM, Transformer).

Referenzen

- 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

Quantitative Finanzforschungsarbeit. Reproduktion mit Zitat gestattet.

☐ **Gemeinschaft**

Treten Sie dem Discord-Server **Trading Pro** bei, um sich mit der Community auszutauschen: <https://discord.gg/u9JApXqhXy>

Didier Vally / Reuniware Systems — InelidaMarketScan v1.0.6
2026-07-19