Kadabra: A Systematic Cryptocurrency Trading Framework

Channel Islands
CALIFORNIA STATE UNIVERSITY

Danil Zanozin • Michael Soltys, Ph.D. • COMP 499

Introduction

This project formalizes the full lifecycle of alpha discovery and deployment in a single proprietary framework, designed and implemented end-to-end by the author. Kadabra supports any trading idea operating at frequencies above one minute (no HFT) and is strategy-agnostic by design: statistical arbitrage, cross-exchange triangular arbitrage, blockchain-analytics-driven signals, traditional finance metrics, and machine-learning models can all be expressed through the same interface and routed through the same risk and execution layers.

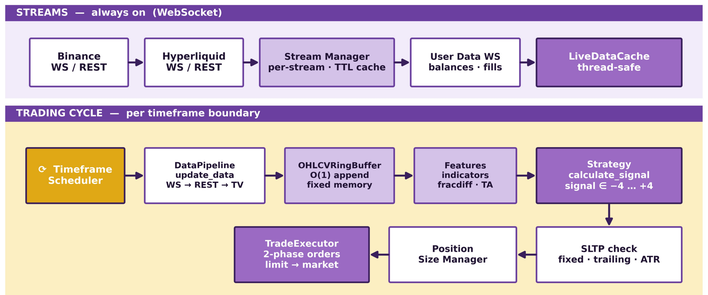
The motivation came from years spent producing open-source market indicators on TradingView, several of which reached the platform's front page and put the author in conversation with industry practitioners. The conclusion from those conversations was unambiguous:
in a maturing market a systematic approach is no longer a preference - it is a requirement.
Anything else is gambling.

Methodology

Kadabra is split into two cooperating subsystems:
1 - an always-on live execution engine,
2 - an offline research / validation pipeline.

Live Execution — Data to Trade

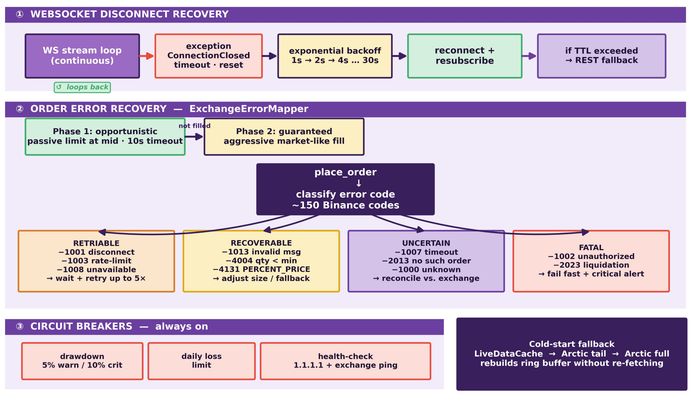
Two independent clocks: WebSocket streams update the cache continuously; the trading cycle fires only on timeframe boundaries.



WebSocket streams from exchanges feed a thread-safe data cache and a numpy-backed ring buffer (O(1) appends).
At each timeframe boundary, a data pipeline produces features (Indicators, statistics, market-structure metrics), the strategy emits a discrete signal in {−4 … +4}, and a two-phase order executor places the trade.
Failure handling is explicit: a 150-error-code exchange error map classifies every API response into
RETRIABLE / RECOVERABLE / UNCERTAIN / FATAL

Live Recovery — How the System Heals Itself

Two failure surfaces, one safety net: connection loss is absorbed by the stream layer; order errors are classified and routed to a recovery action.



A strategy candidate must clear a multi-test robustness gate before promotion to live trading.
The gate combines four independent checks:

-Probability of Backtest Overfitting (PBO) via Combinatorially Symmetric Cross-Validation (Bailey & López de Prado, 2014).
The in-sample/out-of-sample data is split into S=16 contiguous chunks, every C(16,8) = 12,870 partition is evaluated, and PBO is the fraction of partitions in which the in-sample winner underperforms out-of-sample. Lower is better.

-Deflated Sharpe Ratio (DSR) correcting the observed Sharpe for the number of independent trials run during optimization, removing selection bias that would otherwise inflate apparent skill.

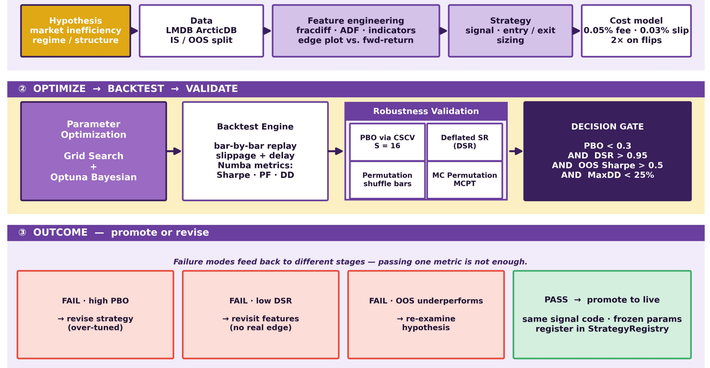
-Monte Carlo Permutation Test (MCPT) - the bar order is shuffled many times to test whether the strategy's profit factor exceeds what is achievable on permuted (signal-free) data.

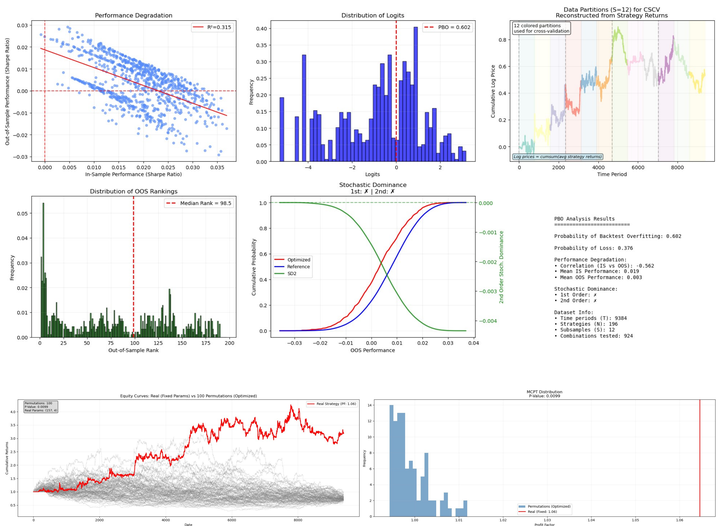
-Out-of-sample Sharpe & MaxDD evaluated on a strict time-series hold-out segment never seen by the optimizer.

A candidate is promoted to live only when all four checks pass simultaneously.
The gate is intentionally aggressive - most candidates are rejected, which is the correct behavior in a regime where optimistic bias is the default failure mode.

Strategy Research — From Hypothesis to Live

Multi-criteria gate separates real edges from overfitted ones.
PBO, DSR, and permutation tests must all clear before a strategy is promoted to live.

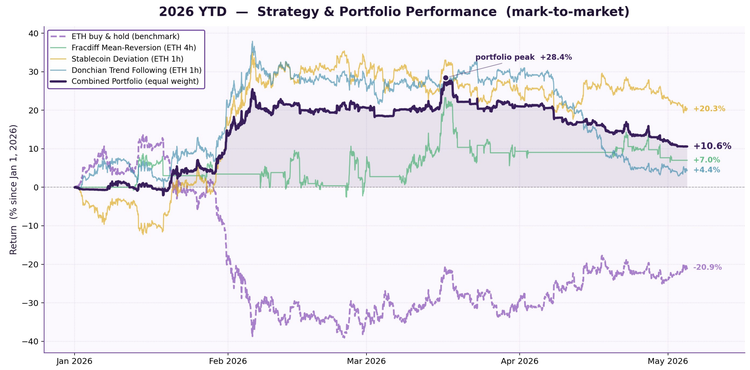




Results

Three strategies cleared the validation gate and are running live with real capital allocation since early 2026:

- Stablecoin Deviation (ETH 1h) - mean-reversion on transient ETH/USDT depegs at the venue level
- Fractional-Differencing Mean-Reversion (ETH 4h) – fracdiff preserves long memory while restoring stationarity, enabling reversion entries on residuals against an adaptive band
- Donchian Trend Following (ETH 1h) - channel breakout with ATR-based stops



The portfolio returned +10.6% YTD against an ETH benchmark of 20.9% . Drawdowns were contained relative to the benchmark, and the equal-weight combination materially reduced single-strategy variance.

Future Work

- Strategy-idea expansion - cross-exchange triangular arbitrage, funding-rate-driven carry, on-chain flow signals (CEX inflows/outflows, stablecoin mint events), and order-book microstructure features as inputs to ML models.
- Multi-asset portfolio - extend beyond ETH to a basket of liquid majors with correlation-aware sizing.
- Allocation models beyond equal-weight - risk-parity, volatility targeting, and Kelly-fractional sizing across the strategy basket

Literature Cited

- Bailey, D. H., & López de Prado, M. (2014). The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting and Non-Normality. Journal of Portfolio Management, 40(5).
- Bailey, D. H., Borwein, J., López de Prado, M., & Zhu, Q. J. (2014). The Probability of Backtest Overfitting. Journal of Computational Finance.
- López de Prado, M. (2018). Advances in Financial Machine Learning. Wiley.
- Masters, T. (2018). Permutation and Randomization Tests for Trading System Development.

Conclusions

- Systematic trading is far harder than it appears from the outside. Most published "edges" do not survive proper out-of-sample testing or selection-bias correction. Kadabra exists to make rejection cheap and promotion rigorous.
-A good idea is necessary but not sufficient. Execution quality (latency, slippage, recovery from API failures) and risk management (position sizing, stop logic, exposure caps) are not afterthoughts - they are first-class components, sized and tested with the same rigor as the alpha itself.
-The validation gate is a filter, not a guarantee. Passing PBO, DSR, MCPT, and OOS thresholds reduces the probability of overfit but does not eliminate it. Live performance remains the only honest test.
-2026 YTD results are consistent with the design hypothesis - positive return in a negative-benchmark regime is good indicator, but true confidence may require many more quarters to establish statistical significance.