Regime-adaptive XGBoost

2018-12-31 to 2023-12-29 · 1,259 observations

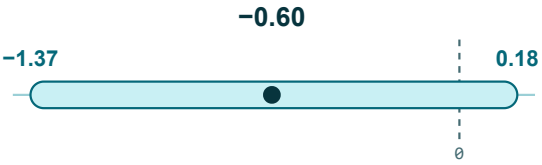
FINDING Descriptive replay

The NVDA replay lost 29.80% net versus a gain of 84.46% for URTH over 1,259 sessions. Net Sharpe was −0.598. Whole-block Viterbi changed 53 test labels, and 551 prediction dates had a sizing regime that differed from their prediction regime.

RETURN AND RISK

METRIC	STRATEGY	BENCHMARK
Compound annual growth	−6.84%	13.04%
Sharpe	−0.60	0.71
Maximum drawdown	−36.53%	−34.01%

Sharpe ratio interval



95.00% block-bootstrap interval. Standard error 0.39.

HOW IT WAS PRODUCED

Dividend- and split-adjusted OHLC for NVDA, QQQ, SPY and URTH. Data begin in late 2016 and the evaluation spans 31 December 2018 to 29 December 2023.

65 expanding windows, a per-window HMM, 300-tree XGBoost and hysteresis selected on validation. Open execution with 5 bps commission and 5 bps slippage. The replay recorded 349 trades and USD 6,908.04 in costs.

The engine reconciled net returns to equity, used explicit 0% annual RF to match the notebook convention, and used total-return URTH without costs.

A separate 14-asset replay used the notebook's GRI statistic. AMD and GOOGL exceeded its descriptive threshold of 0.60.

EVIDENCE SCOPE

EVIDENCE LEVEL **Descriptive replay** Unsigned research evidence

- On 30 May 2019 the state was 0 using the whole block through 10 June and 2 using the data available through that date. Later folds overwrite stored states used for position sizing.
- The final training labels use validation closes. The first position is sized using the session close, with the trade filled at the open. The result is descriptive replay evidence.
- The source Sharpe (−0.439) omits costs from daily returns and first-day P&L. The return series and account equity differ on 349 dates. Short borrow costs are absent.
- The notebook metric labeled WFE compares two halves of the same simulated equity curve. The canonical evidence covers NVDA only.

Regime-adaptive XGBoost

2018-12-31 to 2023-12-29 · 1,259 observations

RETURN AND RISK

METRIC	STRATEGY	BENCHMARK
Compound annual growth	**−6.84%**	13.04%
Arithmetic annual return	**−6.49%**	14.34%
Cumulative return	**−29.80%**	84.46%
Annual volatility	**10.86%**	20.34%
Sharpe	**−0.60**	0.71
Sortino	**−0.80**	0.98
Calmar	**−0.19**	0.38
Maximum drawdown	**−36.53%**	−34.01%
Positive periods	**42.73%**	55.28%
Peak to trough periods	**1,183**	27
Recovery periods	**—**	134
Skewness	**−0.54**	−0.80
Excess kurtosis	**9.82**	14.56

AGAINST THE BENCHMARK

METRIC	ESTIMATE
Arithmetic annual difference	−20.83%
Tracking error	22.95%
Information ratio	−0.91
Beta	0.01
Correlation	0.01
Annual Jensen alpha	−6.57%

UNCERTAINTY AND VERIFICATION

METRIC	ESTIMATE
Sharpe interval	−1.37 … 0.18
Information ratio interval	−1.81 … −0.17
Active return HAC t	−2.19
Two-sided HAC p	0.03
Block length	5
Valid resamples	1,000

CONFIGURATION

METRIC	ESTIMATE
Benchmark	URTH
Annual risk-free rate	0.00%
Periods per year	252
Calendar	XNYS