Explore cryptocurrency trading strategies：

From Factor Models and XGBoost to Generative AI

Team Members

Wenjing Li, Yanbing Zhao,

Yihui Wang, Lu Zhuo, Hanshuo Wang

University of Nottingham

10 Cryptocurrency Classification Framework

Market Share

Representing 75% -85% of the market share

Filtering logic

Liquidity and reliability

Mitigate manipulation risks

Non-Turing complete architecture	Turing-complete crypto commodities	Pegged to off-chain assets (USD)
Store of Value BTC, LTC	**Smart Contracts** ETH, SOL, TRX	**Fiat Stablecoins** USDT, USDC
Settlement XRP	**Exchange** BNB	**Meme** DOGE
Alliance chain institutional B2B focus	Hybrid asset (CeFi + DeFi)	Emotional drive meme premium

Multi-Dimensional Factor System

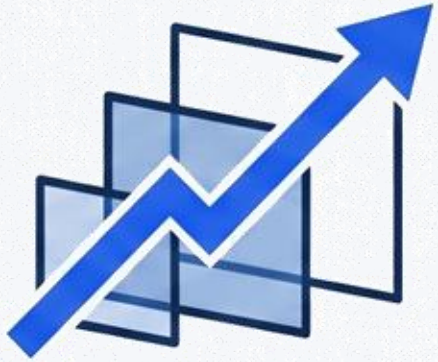
Price

- **Closing Price (PRC):** Intraday info clearing equilibrium.
- **Maximum daily price (MAXDPRC):** Impacts expectations & liquidity.

Size Trading Activity

- **Market Cap (MCAP):** Linked to size premium.
- **Price Volume (PRCVOL):** Reflects market depth.
- **Standard Deviation of Price Volume (STDPRCVOL):** Liquidity uncertainty.

Momentum



Short-frequency **momentum over 1, 5, 10, 20-day** windows to capture trend persistence in high-volatility crypto.

Sentiment

Crypto Fear&Greed Index: A behavioral indicator quantifying irrational fluctuations driven by noise trading.

Data Sources & Benchmark Index

Primary Data Source

Provider: CoinMarketCap

- **Rationale:** The most widely recognized and authoritative platform globally for tracking dynamic cryptocurrency market information.
- **Scope:** Includes comprehensive data on historical pricing, market capitalization, and granular trading volumes.

Accessed at:
https://coinmarketcap.com

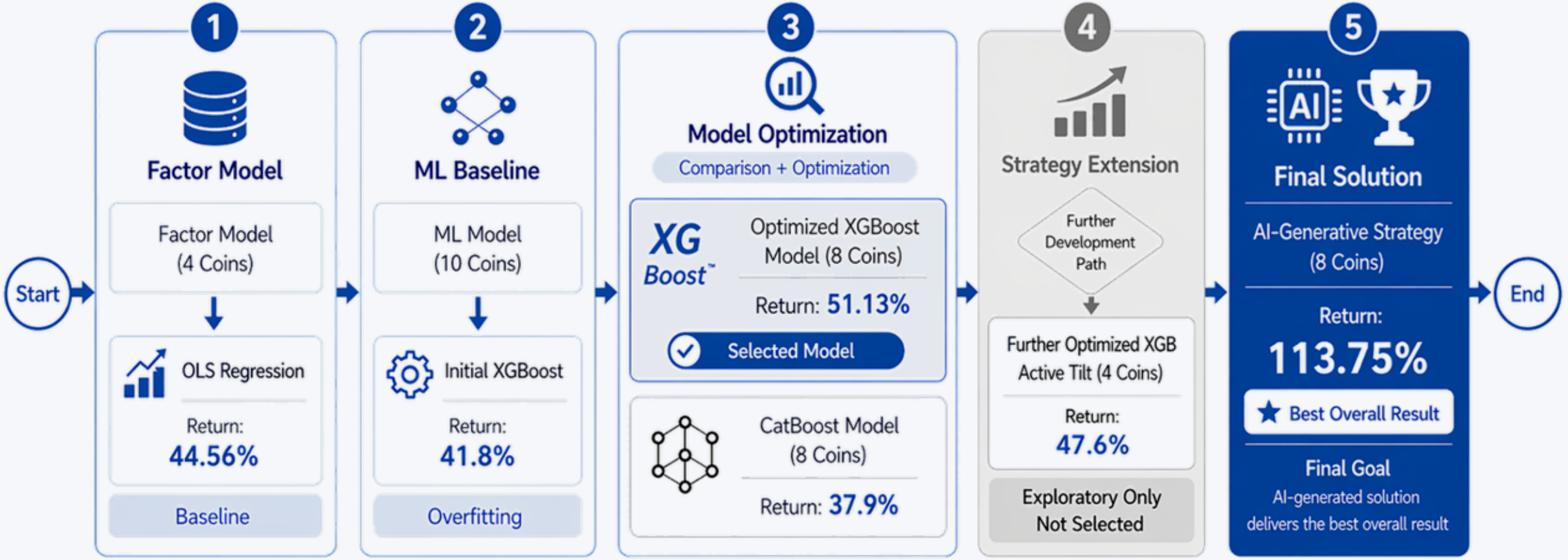
Benchmark Selection

S&P Cryptocurrency Top 30 Equal Weight Index

- **Advantage:** Effectively reduces the severe market capitalization concentration typically dominated by BTC and ETH.
- **Outcome:** Allows for a scientifically balanced and accurate measurement of active excess returns generated by the strategies.

Accessed at:
https://www.spglobal.com/spdji/en/indices/digital-assets/sp-cryptocurrency-top-30-equal-weight-index/#overview

Research Framework & Technical Roadmap



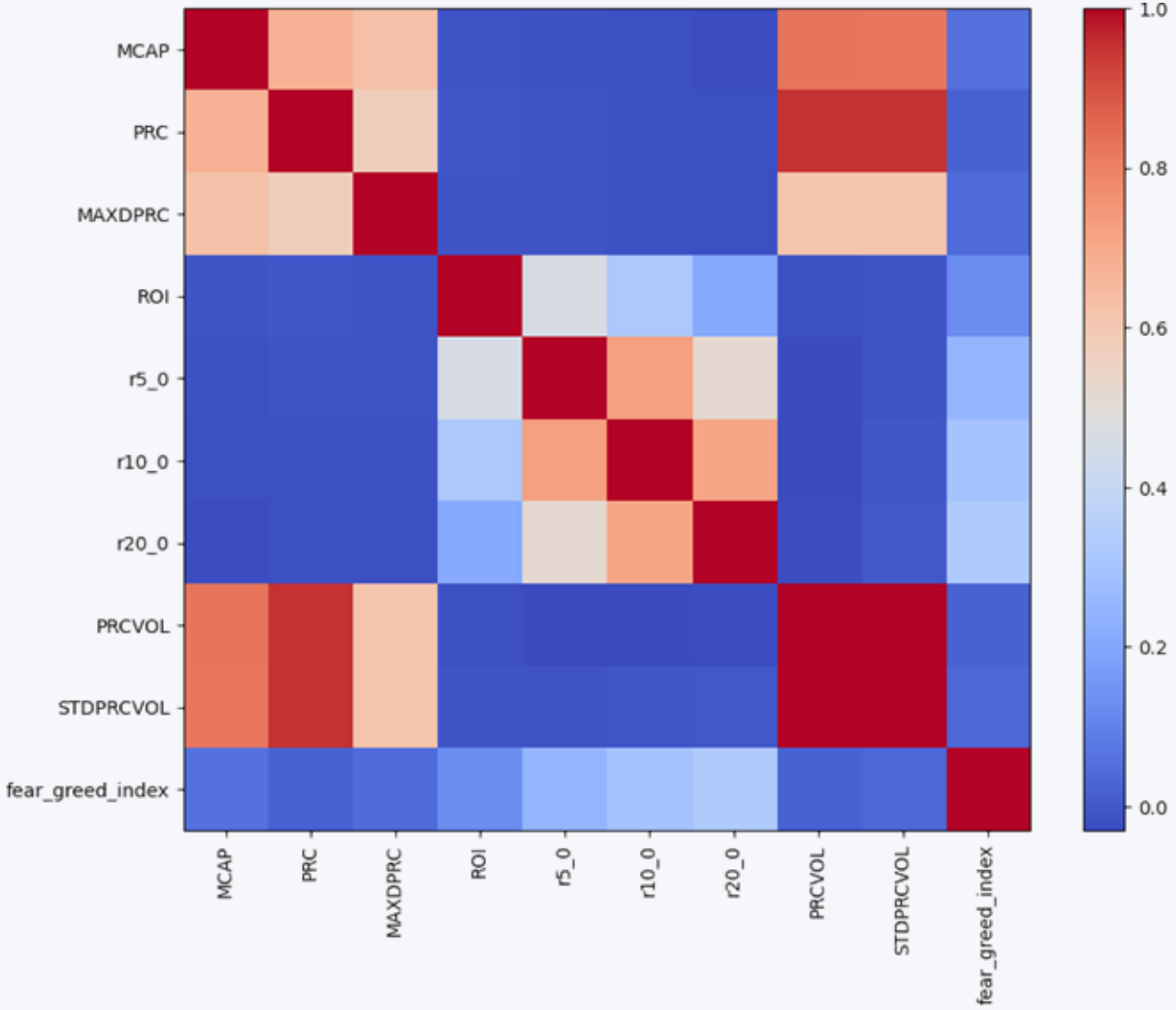
Note: Return refers to cumulative out-of-sample backtest return.

Model Design

1. From basic factor regression to machine learning model (XGBoost)
2. Continuous comparison and iterative optimization
3. Ultimately achieving the highest return (113.75%)

Speaker：Hanshuo Wang

Exploratory Data Characteristics & Analysis



High Collinearity Cluster

- Size& Trading activity : 0.82
- Price & Trading activity: 0.95
- Momentum: **short-term trend persistence**

OLS coefficient estimates are unstable.

Independent Signal

- CFGI & Price: 0.02
- CFGI & Size: 0.06
- CFGI & Trading activity: 0.03
- CFGI & Momentum: 0.3

Sentiment provides unique predictive value.

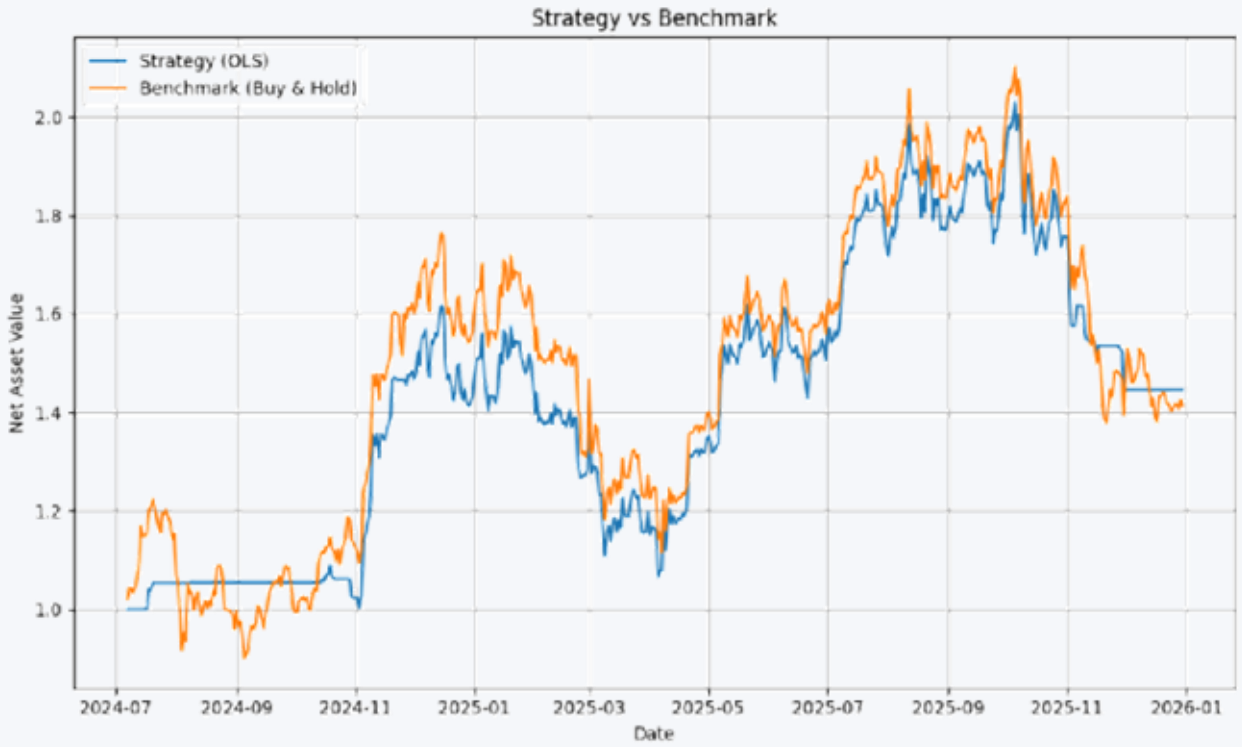
Non-Normal Returns

- Daily returns exhibit **leptokurtic** (Kurtosis 6.13) and **fat-tailed** distributions

Confirm the limitations of OLS

OLS Benchmark Model: Linear Relationship Capture

Performance of strategy back testing based on OLS



0.020

R^2 (Variation Explained)

2.256 (0.004)

F (Prob F)

Core factor I: MCAP Factor

coef=-0.1911 (P=0.007): Strong "small-cap effect" identified: Lower market cap assets outperform.

Core factor II: Fear & Greed

coef=0.0002 (P=0.018): Sentiment-driven pricing: Periods of high greed correlate with higher returns.

+3.18% / 0.81

Excess Return/Sharpe Ratio

vs. 0.72 (Benchmark)

53.42%

forecast accuracy

vs. 50% (Random)

core findings

The OLS linear model is difficult to fit the highly nonlinear and dynamically changing return characteristics of the cryptocurrency market, thus having limited model explanatory power and low prediction accuracy.

XGBoost Baseline: The Overfitting Lesson

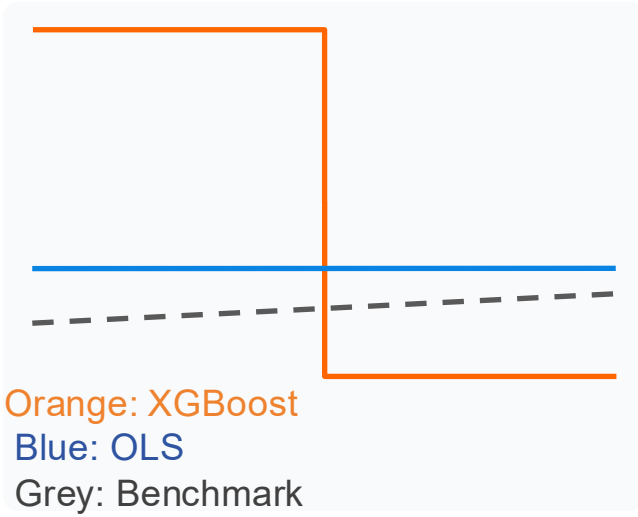
Confusion Matrix (Test Set)

True Positive **1238**	False Positive **1224**
False Negative **1440**	True Negative **1436**

Accuracy 50.32%

ROC AUC 0.5001

Strategy Returns Comparison



Orange: XGBoost
Blue: OLS
Grey: Benchmark

Total Return:
XGBoost 41.82%
OLS 44.56%

Max Drawdown:
XGBoost -46.52%
OLS -34.04%

Root Cause Analysis

Overly Complex Model
n_estimators = 5000 (excessive depth)

→ **Noise Over Signal**
Model memorized high-frequency market noise

Logical flaw in Asset Pool design
Including both USDT and USDC stablecoins

→ **Covering up the truth**
offer NO predictable excess returns but instead interferes with signal learning.

XGBoost Optimization: Systematic Reconstruction

Four-Dimension Optimization

1 Asset Purification

Remove USDT/USDC low-volatility assets

2 Time-Series Validation

60/10/30 split, validation prevents leakage

3 Dynamic Thresholding

Optimize accuracy, F1, recall, precision

4 Benchmark Restructuring

S&P Crypto Top 30 as external benchmark

Key Parameters

max_depth **3**	n_estimators **300**
learning_rate **0.03**	colsample_bytree **0.8**

Optimized Performance

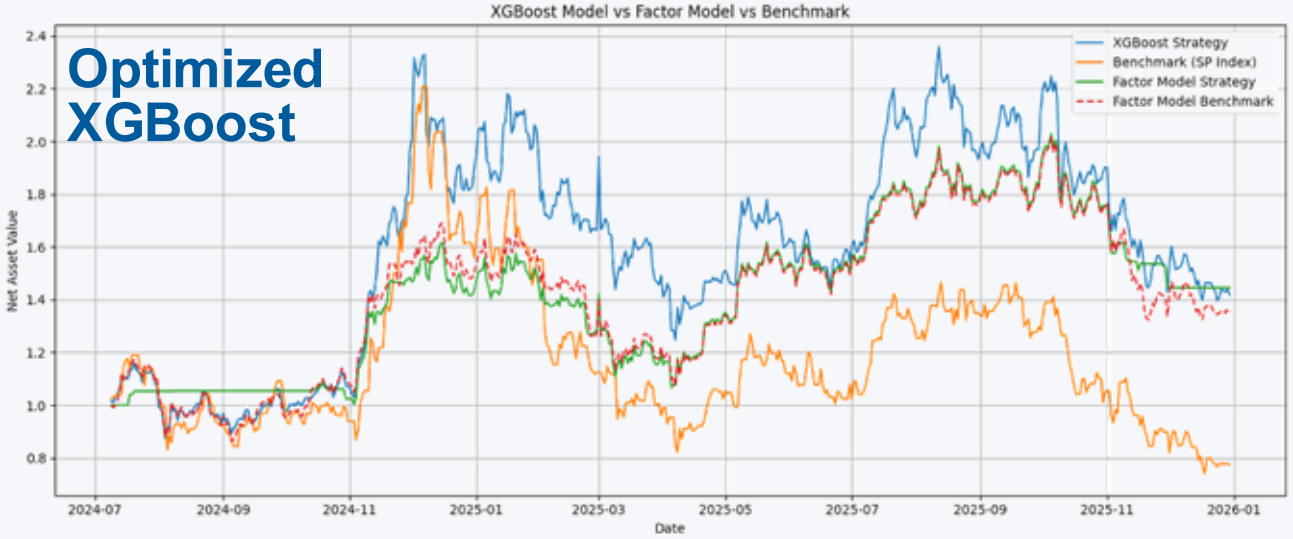
Benchmark Comparison:

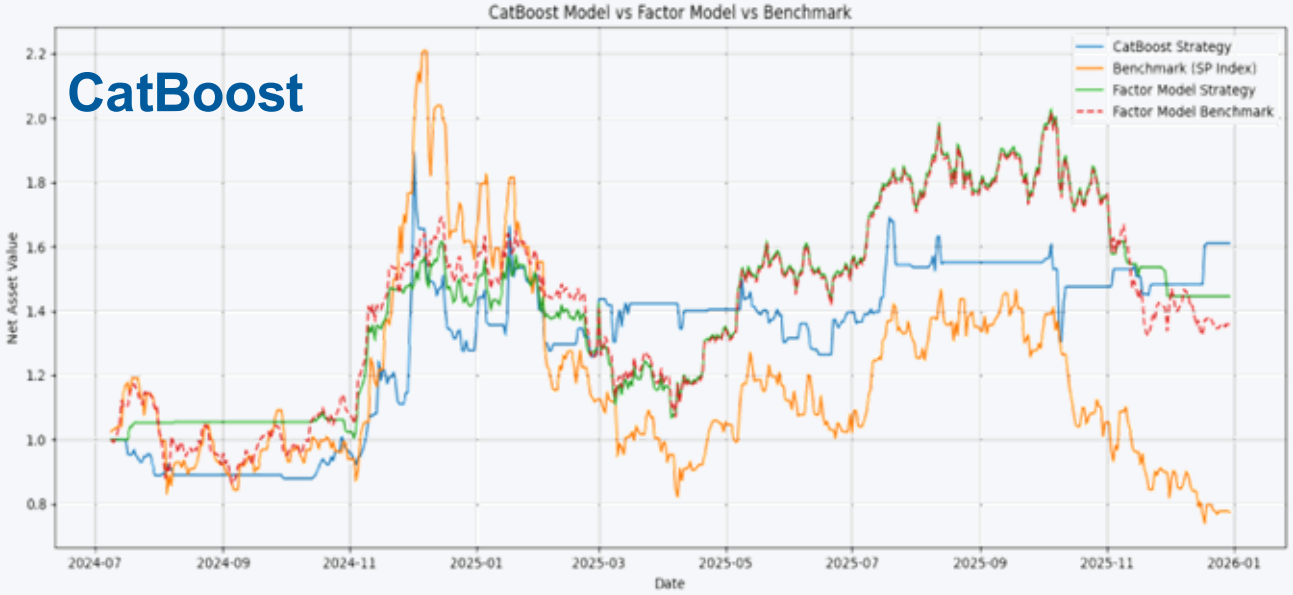
Metric	Optimized XGBoost	Benchmark S&P
Annualized Volatility	***0.6090***	0.7168
Sharpe Ratio	***0.5468***	-0.2333
Max Drawdown	***-0.4652***	-0.6652

Total Return ***53.13%***	Annualized Return ***33.31%***
Sharpe Ratio ***0.55***	Win Rate ***51.94%***

Speaker：Lu Zhuo

XGBoost vs CatBoost: Model Validation





Model Comparison

- **CatBoost model underperforms during high-volatility market periods.**
- **XGBoost sustains a steady upward trajectory throughout market fluctuations.**
- **XGBoost proves more robust for dynamic crypto market conditions, both in practice and academic consensus.**

Further Optimized XGBoost: 4 Coins Active-Tilt

Strategy Design

Four-coin active tilt on **BTC, ETH, BNB, SOL** with market-cap base + XGBoost signals.

Position Sizing

Top-K=2 selection, market-cap weighting

Rebalancing

Periodic rebalancing, individual weight caps

Theoretical Basis

Active Portfolio Management (Grinold & Kahn, 2000)

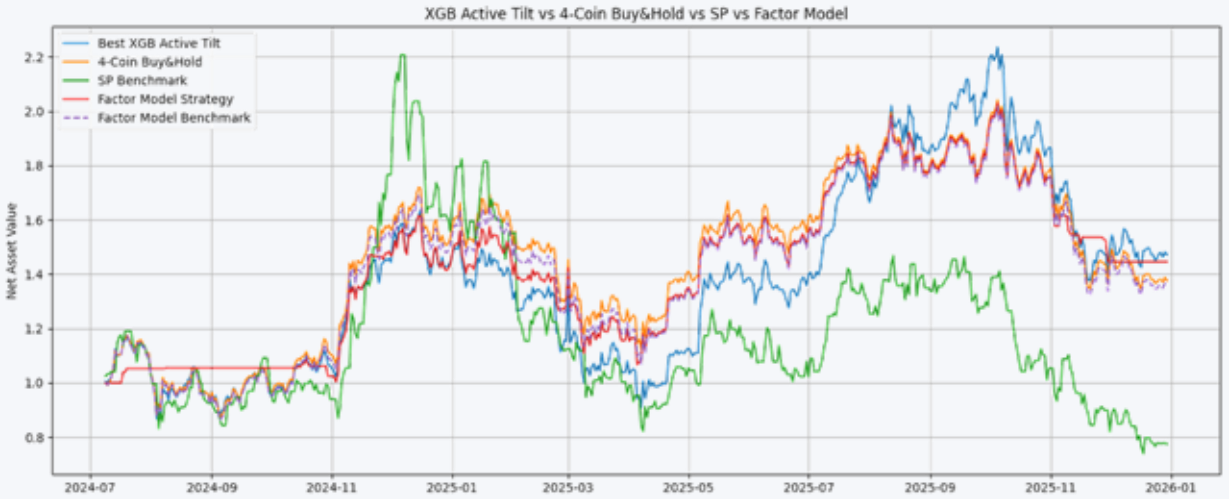
Attribution Analysis

Limitation 1

Restricted asset pool limits cross-sectional differentiation

Limitation 2

Complex rules weaken signal transmission efficiency



Performance Results

Total Return	Annualized Return
47.21%	**29.93%**
vs. Factor Model 44.56%	Higher absolute return

Sharpe Ratio	Max Drawdown
0.53	**- 44.45%**
vs. Factor Model 0.69	Worse than benchmark

Critical Finding

Higher absolute returns but lower risk-adjusted returns — complexity harmed efficiency

Speaker：Yanbing Zhao

AI-Generated Solution: The Power of Minimalist Execution

Core Philosophy

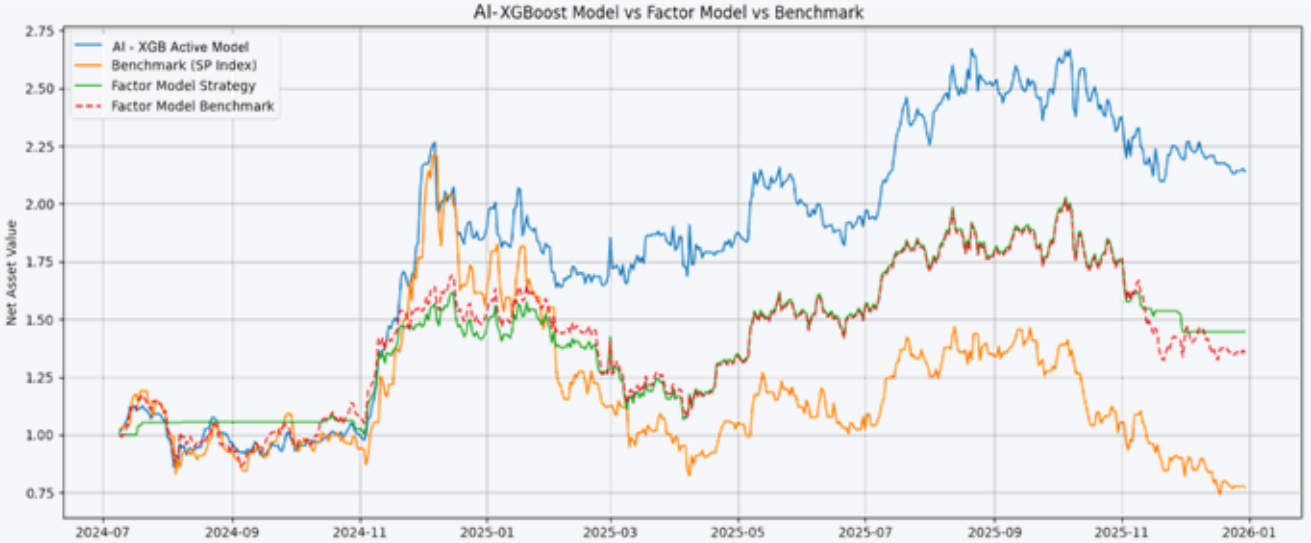
Strip redundant constraints, adopt intuitive equal-weight long-only rule.

Execution Logic

Equal-weight longs in predicted appreciating coins, no position if no signal.

AI Integration

Auto-identifies and standardizes prediction dataframes.



Why It Works

- ✓ **Signal purity:** Direct mapping from predictions to positions
- ✓ **No dilution:** Complex rules removed
- ✓ **Efficiency focus:** Better transmission, not more complexity

Breakthrough Performance

Total Return	Annualized Return
1.1375	66.95%

Max Drawdown	Sharpe Ratio
- 27.74%	1.3774 Highest in study

↑ Risk-Adjusted Excellence

Improvement from enhanced risk-adjusted performance , not higher risk.

Core Findings & Full Model Comparison

Three Core Findings

1 Signal Transmission Efficiency

Efficient signal-to-position mapping matters more than prediction accuracy alone.

2 Anti-Overfitting Defense

In high-noise markets, shallow tree ensembles with regularization provide stronger generalization.

3 Minimalist Execution

Removing redundant constraints allows model signals to be translated into positions more directly.

No	Metric	Factor Model	XGBoost	Optimized XGBoost	XGB Active Tilt	AI-Generative XGBoost	SP Benchmark
1	Total Return	0.445623	0.41818	0.531282	0.472085	**1.137529**	-0.237577
2	Annualized Return	0.283471	0.26692	0.333068	0.299334	**0.66949**	-0.167239
3	Annualized Volatility	**0.410473**	0.606644	0.609071	0.563691	0.48607	0.716828
4	Sharpe Ratio	0.690596	0.439995	0.546846	0.531025	**1.377352**	-0.233304
5	Max Drawdown	-0.340351	-0.46517	-0.46517	-0.444493	**-0.277372**	-0.665195
6	Win Rate	0.374768	0.517625	0.519409	**0.54731**	0.486137	0.349353

Overarching Conclusion:

In machine-learning-driven quantitative trading, greater complexity does not guarantee superiority.

Sustainable outperformance comes **from removing redundant constraints** and **adopting a minimalist execution framework**.

Investment Proposal: $1M Allocation Strategy

Conservative Investor ("Barbell" Approach)

Daily Allocation:

- Hold 50% of the $1 million in USDC/USDT (stablecoin reserve).
- Invest the remaining 50% equally among core coins predicted to rise by the Generative AI.

Key Strategy Benefit:

Uses the stablecoin reserve to effectively dilute overall drawdown risk.

Aggressive Investor (Generative-AI XGBoost Strategy)

Daily Allocation:

- Invest 100% of funds equally among core coins predicted to rise by the Generative-AI Strategy.
- Hold the entire $1 million in cash if there are no positive signals.

Key Strategy Benefit:

Maximizes gains by executing pure predictive signals protecting capital during market downturns.

Thank you for listening !

Team Members

Wenjing Li, Yanbing Zhao,

Yihui Wang, Lu Zhuo, Hanshuo Wang

University of Nottingham