Contents

1 Pipeline Architecture

The platform implements a multi-strategy cryptocurrency trading system with the following stages: data processing, signal generation, predictor construction, portfolio construction, backtesting, and performance evaluation.

1.1 Data Pipeline

1.1.1 Price Matrix Construction

Given raw OHLCV data for N symbols over T time periods, construct the price matrix:

$$P \in \mathbb{R}^{T \times N}, \quad P_{t,i} = \text{close price of asset } i \text{ at time } t$$

Data preprocessing includes removing duplicates, handling missing values via forward-filling, and resampling to hourly frequency.

1.2 Signal Generation Strategies

The platform implements three complementary signal strategies:

1.2.1 Momentum Strategy

Captures trending behavior using past returns:

$$\text{Signal}_{t,i}^{\text{mom}} = \frac{P_{t-1,i} - P_{t-1-L,i}}{P_{t-1-L,i}}$$

where L is the lookback period (typically 168 hours = 1 week). The shift by 1 period prevents look-ahead bias.

1.2.2 Mean Reversion Strategy

Exploits price extremes using z-score of returns:

$$r_{t,i} = \frac{P_{t,i} - P_{t-1,i}}{P_{t-1,i}}$$

$$\mu_{t,i} = \frac{1}{L} \sum_{k=0}^{L-1} r_{t-k,i}$$

$$\sigma_{t,i} = \sqrt{\frac{1}{L} \sum_{k=0}^{L-1} (r_{t-k,i} - \mu_{t,i})^2}$$

$$\text{Signal}_{t,i}^{\text{mr}} = -\frac{r_{t-1,i} - \mu_{t-1,i}}{\sigma_{t-1,i}}$$

The negative sign implements mean reversion: oversold assets (negative z-score) produce positive signals.

1.2.3 EWMA Crossover Strategy

Identifies trend changes via exponential moving average divergences:

$$\text{EWMA}_{t,i}^{\text{fast}} = \alpha_{\text{fast}} P_{t,i} + (1 - \alpha_{\text{fast}})\text{EWMA}_{t-1,i}^{\text{fast}}$$

$$\text{EWMA}_{t,i}^{\text{slow}} = \alpha_{\text{slow}} P_{t,i} + (1 - \alpha_{\text{slow}})\text{EWMA}_{t-1,i}^{\text{slow}}$$

where $\alpha = 2/(\mathit{span} + 1)$. The signal is volatility-normalized:

$$\text{Signal}_{t,i}^{\text{ewma}} = \frac{\text{EWMA}_{t-1,i}^{\text{fast}} - \text{EWMA}_{t-1,i}^{\text{slow}}}{\sigma_{t-1,i}^{\text{rolling}}}$$

where $\sigma_{t,i}^{\text{rolling}}$ is the rolling standard deviation over a specified window.

1.3 Predictor Construction

1.3.1 Cross-Sectional Z-Score Transformation

For comparability across strategies, raw signals are cross-sectionally standardized:

$$\bar{s}_t = \frac{1}{N} \sum_{i=1}^{N} \text{Signal}_{t,i}$$

$$\sigma_t = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\text{Signal}_{t,i} - \bar{s}_t)^2}$$

$$Z_{t,i} = \frac{\text{Signal}_{t,i} - \bar{s}_t}{\sigma_t}$$

1.3.2 Quantile Filtering

Filter signals by cross-sectional quantiles to select top/bottom performers:

$$F_{t,i} = \begin{cases} Z_{t,i} & \text{if } Z_{t,i} \geq Q_t^{\text{top}} \text{ (long)} \\ Z_{t,i} & \text{if } Z_{t,i} \leq Q_t^{\text{bottom}} \text{ (short)} \\ 0 & \text{otherwise} \end{cases}$$

where Q_t^{top} and Q_t^{bottom} are the top and bottom quantile thresholds (e.g., 0.8 and 0.2).

1.3.3 Re-Z-Score on Active Names

After filtering, re-standardize on active (non-zero) positions only:

Let $A_t = \{i : F_{t,i} \neq 0\}$ be the active set at time t:

$$\bar{f}_t = \frac{1}{|A_t|} \sum_{i \in A_t} F_{t,i}$$

$$\text{Signal}_{t,i}^{\text{final}} = \begin{cases} \frac{F_{t,i} - \bar{f}_t}{\sigma_{A_t}} & \text{if } i \in A_t \\ 0 & \text{otherwise} \end{cases}$$

1.4 Portfolio Construction

Combine multiple strategies using weighted averaging:

$$Z_{t,i}^{\text{combined}} = \sum_{s=1}^{S} w_s \cdot Z_{t,i}^{(s)}$$

where w_s are strategy weights satisfying $\sum_s w_s = 1$, and $Z_{t,i}^{(s)}$ is the unfiltered z-scored signal from strategy s.

The combined signal undergoes filtering and re-z-scoring (same process as above) to produce final portfolio weights $W_{t,i}$.

1.5 Target Engineering

Compute forward returns for evaluation:

$$\text{Target}_{t,i} = \frac{P_{t+h,i} - P_{t,i}}{P_{t,i}}$$

where h is the forward period (typically 1 hour).

1.6 Backtesting

1.6.1 Timing Convention

Weights determined at close of day t are held through day $t + 1$, realizing returns from close t to close $t + 1$.

Asset returns:

$$R_{t,i} = \frac{P_{t,i} - P_{t-1,i}}{P_{t-1,i}}$$

Portfolio return at time t:

$$R_t^{\text{portfolio}} = \sum_{i=1}^{N} W_{t-1,i} \cdot R_{t,i}$$

1.6.2 Turnover (portfolio rebalancing)

$$\text{Turnover}_t = \frac{1}{2} \sum_{i=1}^{N} |W_{t,i} - W_{t-1,i}|$$

1.6.3 Transaction costs

$$\text{Cost}_t = \text{Turnover}_t \times (\text{TC}_{\text{bps}} + \text{Slippage}_{\text{bps}})$$

1.6.4 Net return

$$R_t^{\text{net}} = R_t^{\text{portfolio}} - \text{Cost}_t$$

1.6.5 Cumulative return

$$\text{CR}_t = \prod_{k=1}^{t} (1 + R_k^{\text{net}})$$

1.7 Performance Metrics

1.7.1 Annualized Return (geometric mean)

$$R_{\text{annual}} = \left(\prod_{t=1}^{T} (1 + R_t^{\text{net}}) \right)^{\frac{8760}{T}} - 1$$

where $8760 = 365 \times 24$ hours per year.

1.7.2 Annualized Volatility

$$\sigma_{\text{annual}} = \sqrt{8760} \cdot \sqrt{\frac{1}{T} \sum_{t=1}^{T} (R_t^{\text{net}} - \bar{R})^2}$$

1.7.3 Sharpe Ratio

$$\text{Sharpe} = \frac{R_{\text{annual}} - R_f}{\sigma_{\text{annual}}}$$

1.7.4 Maximum Drawdown

$$\text{MDD} = \min_t \left(\frac{\text{CR}_t - \max_{k \le t} \text{CR}_k}{\max_{k \le t} \text{CR}_k} \right)$$

1.7.5 Calmar Ratio

$$\text{Calmar} = \frac{R_{\text{annual}}}{|\text{MDD}|}$$

1.7.6 Win Rate

$$\text{WR} = \frac{\sum_{t=1}^{T} \nVdash(R_t^{\text{net}} > 0)}{T}$$

1.7.7 Profit Factor

$$\text{PF} = \frac{\sum_{t:R_t^{\text{net}}>0} R_t^{\text{net}}}{|\sum_{t:R_t^{\text{net}}<0} R_t^{\text{net}}|}$$