



UMHackathon 2025 : Balaena Quant Problem Statement

Problem Statement: Alpha Strategies Using HMM or ML

Objective

Develop a **Machine Learning (ML)** model that analyzes on-chain data from various sources (e.g., CryptoQuant, Glassnode, Coinglass) to generate an **alpha trading strategy** that maximizes profit. The model should effectively extract implicit indicators from noisy data to generate profitable trading signals.

It is highly recommended that your model incorporate characteristics of **Hidden Markov Models (HMMs)** to identify deterministic patterns in market movements. **Natural Language Processing (NLP)** attempts at analyzing textual information are also welcome.

Requirements

1. Data Sources & Processing
 - The model must process data from **multiple sources** such as CryptoQuant, Glassnode, Coinglass and other relevant platforms.
 - Data intervals must be ≤ 1 day (e.g., 4 hours, 10 minutes).
2. Model & Strategy Design
 - The ML model must identify **implicit market indicators** for alpha generation.
 - The model may integrate **HMMs** to enhance pattern recognition and regime detection.
 - It should optimize trading signals based on extracted features to **maximize returns**.
3. Trading Execution
 - The strategy should generate at least **3% trade signals per data row** to ensure adequate trading frequency.
 - Execution logic should be based on the **predicted market states** or **sentiment shifts**.
 - Trading fees of 0.06% must be accounted for
 - Data period used for the backtest should be several years, and the forward test should be at least one year.

Success Criteria

- **Sharpe Ratio (SR) ≥ 1.8** (Ensures risk-adjusted returns are sufficiently high)
- **Maximum Drawdown (MDD) $\geq -40\%$** (Limits downside risk exposure)
- **Trade Frequency $\geq 3\%$ per data row** (Ensures sufficient trading activity)