

# Credit Scoring Methodology Document

## Objective

The objective is to assign a credit score between 0 and 100 to each wallet address based solely on its historical transaction behavior. The goal is to distinguish between trustworthy (human-like, protocol-aligned) and risky (bot-like, exploitative) behaviors.

## Data Sources

Three raw .json files containing Compound V2 Ethereum transaction data were used, specifically focusing on deposit events. Each transaction included:

- account.id: Wallet address
- amountUSD: Deposit amount in USD
- timestamp: Unix timestamp of the transaction

## Feature Engineering

From the raw logs, the following wallet-level features were derived:

1. Total Deposit
  - Cumulative USD value of all deposits by a wallet.
2. Number of Deposits
  - Total count of deposit transactions per wallet.
3. Activity Span (Days)
  - Duration between a wallet's first and last deposit (in days).
4. Deposit Frequency per Day
  - Average number of deposits per day:  
$$\text{frequency} = \text{total deposits} / \text{activity span}$$

These behavioral proxies reflect traits like consistency, commitment, and volatility, which are useful for evaluating responsible or suspicious usage.

## Modeling Approach

### Rule-Based Scoring System

A rule-based approach was chosen, as it fits within the constraints of unsupervised learning and prioritizes explainability.

Justification:

- Explainability: Simple and interpretable rules aligned with protocol goals.
- No Overfitting: Does not depend on data patterns from external or labeled sources.
- Compliance: No pretrained models or third-party scores used, ensuring integrity.

### Scoring System

Wallets receive scores from 0 to 100 using the following logic:

#### 1. Base Score

- All wallets begin with a base score of 50.

#### 2. Rewards

- +20: Wallet has  $\geq 3$  unique transaction types (high diversity)
- +10: Total deposit above the median across all wallets
- +10: Deposit frequency  $> 0.5$  deposits/day (consistent activity)

#### 3. Penalties

- -10: Only 1 transaction type used (low diversity)
- -5 (each): For every associated liquidation event
- -10: Activity span less than 7 days (short-lived behavior)

#### 4. Score Clipping

- Final score is bounded between 0 and 100

## Output

#### 1. Top Wallets CSV

- File: top\_1000\_wallets.csv
- Contains the top 1,000 wallets by credit score
- Columns:

- account.id: Wallet address
- score: Assigned credit score

## 2. Wallet Behavior Analysis

- A one-page summary comparing five high-scoring and five low-scoring wallets.
- Highlights:
  - Behavioral trends
  - Anomalies
  - Observed patterns of trustworthiness vs risk

## Summary

This methodology delivers a transparent, rule-based credit scoring system tailored to wallet behaviors on the Compound V2 protocol. It uses clearly defined behavioral signals to promote fairness and identify responsible vs. exploitative activity — without relying on machine learning or labeled data.