Wallet Risk Scoring Framework
For Compound Protocol V2/V3

1. Data Collection Method

Sources

✓ Primary Data:

- Ethereum blockchain (via Etherscan API/Web3 providers)
- Compound Subgraph (The Graph) for protocol-specific events

✓ Secondary Data (if needed):

- DeFi Llama/Risk APIs (cross-protocol exposure)
- On-chain analytics (Nansen, Arkham for labeled addresses)

Key Metrics Collected

Metric	Description
Transactions	All historical Compound interactions
Liquidations	Count of LiquidateBorrow events
Borrow/Supply	Frequency, amounts, collateral types
Health Factor	Min/max proximity to liquidation threshold
Cross-Protocol	Exposure to other DeFi apps (Aave, etc.)

Scalability

- **Batch processing** for large wallet sets
- **Caching** to reduce API calls
- **Rate limiting** to avoid throttling

2. Feature Selection & Rationale

Core Risk Indicators

Feature	Risk Relevance	Weight
Liquidation Count	Direct evidence of under-collateralization	30%
Borrow Frequency	High borrowing → leverage risk	20%
Collateral Diversity	Over-reliance on volatile assets (e.g., single-token collateral)	15%
Health Factor Trends	How close the wallet was to liquidation (even if avoided)	15%
Transaction Velocity	Sudden spikes → potential hacking/panic selling	10%
Cross-Protocol Risk	Exposure to high-risk strategies (e.g., leverage farming)	10%

3. Scoring Methodology

Risk Scale: 0-1000

- **0-300** = Low risk
- **301-600** = Moderate risk
- **601-1000** = High risk

Formula

Risk Score =
(Liquidation Score × 0.30) +
(Borrowing Score × 0.20) +
(Collateral Score × 0.15) +
(Health Factor Score × 0.15) +
(Transaction Velocity Score × 0.10) +
(Cross-Protocol Score × 0.10)

Normalization

- Each feature scaled **0-1** before weighting.
- **Example:** A wallet with **5 liquidations** (max observed = 10) → 5/10 = 0.5.

4. Justification of Risk Indicators

Indicator	Why It Matters
Liquidations	Proof of past collateral failures
Borrowing Frequency	Frequent borrowers = higher leverage risk
Collateral Diversity	Single-asset collateral → higher volatility risk
Health Factor Trends	Warns of near-liquidations before they happen
Transaction Velocity	Abnormal activity → potential exploit/hack
Cross-Protocol Risk	Using risky protocols (e.g., leverage farms) compounds systemic risk

5. Validation & Scalability

Validation

✓ **Backtesting:** Compare scores against historical liquidations.
✓ **Known Risk Wallets:** Test against flagged addresses (e.g., hacked wallets).
✓ **Manual Audits:** Human review of sample wallets.

Scalability Enhancements

- **Parallel processing** for 10,000+ wallets
- **Tiered analysis:** Quick scan → deep dive for high-risk wallets
- **Dynamic updates:** Adjust weights if Compound changes parameters