Wallet Risk Scoring Explanation

By: Nidhi Rana

---

1. Introduction

In this project, I calculated risk scores for different Ethereum wallet addresses. The goal was to understand how risky each wallet is based on their transactions with DeFi lending protocols.

---

2. Data Collection Method

I used the **Etherscan API** to get the transaction history of each wallet.
Here is how it worked:

- I took each wallet address from the given list.

- Used the Etherscan API to fetch all its transactions on Ethereum.

- Stored the transactions for further analysis.

I used Etherscan because **The Graph endpoint for Compound was deprecated**, and Etherscan provides general transaction data reliably.

---

3. Feature Selection Rationale

I created features that can reflect each wallet's financial behaviour:

1. **Total Supplied:**
   Amount the wallet supplied to the protocol. Higher supply generally reduces risk.

2. **Total Borrowed:**
   Amount borrowed by the wallet. Higher borrow could indicate more risk if repayment is not good.

3. **Repayment Ratio:**
   Calculated as (total repaid / total borrowed). Shows how well the wallet repays its debt.

4. **Active Borrow Positions:**
   If the wallet has any active borrowings, it increases its risk score slightly.

---

4. Data Preprocessing

- If a wallet had **no transactions**, I filled its features with zeros.

- I used **Min-Max Normalization** to scale the features between 0 and 1 for fair scoring.

---

5. Scoring Method

To calculate the **final risk score (between 0 and 1000)**:

1. I assigned weights to each feature:

   - Total Supplied: **10% weight**

   - Total Borrowed: **15% weight**

   - Repayment Ratio: **25% weight**

   - Active Borrow Positions: **10% weight**

2. The formula combined these weighted features.

3. Finally, I multiplied the combined score by **1000** to bring it into a readable range from **0 to 1000**.

---

6. Justification of Risk Indicators

- **Supplying funds reduces risk** because the user is giving liquidity.

- **Borrowing increases risk** as it can lead to defaults.

- **Repayment ratio is crucial** – a higher ratio means the wallet repays its loans well, reducing risk.

- **Active borrow positions show exposure**, so it slightly increases risk.

---

7. Final Output

The final output is a **CSV file** with:

Wallet_id	Score
0xabc...123	742
0xdef...456	613

Each wallet gets a **score between 0 to 1000**.

Higher scores mean **lower risk**. Lower scores indicate **riskier wallets**.

---

8. Limitations & Future Improvements

- The method used placeholder logic for borrow/repay detection (method IDs).

- In the future, I can integrate **Compound or Aave subgraph APIs** for exact decoded DeFi transactions.

- Could use **more complex ML models** for better accuracy