Aave Wallet Credit Scoring

This project builds a machine learning-inspired scoring pipeline to assess the reliability of wallets interacting with the Aave V2 protocol. Each wallet receives a credit score between 0 and 1000, based solely on historical transaction behavior.

Goal

To assign credit scores to wallets based on actions
like deposit, borrow, repay, redeemunderlying, and liquidationcall and evaluate behavioral patterns of users in a DeFi protocol.

Methodology

Data Source

The dataset consists of ~100K transaction records in JSON format representing user interactions with the Aave V2 protocol.

Feature Engineering

We extracted the following features per wallet:

- Total number of transactions
- Count of deposits, borrows, repays, redeems, and liquidations
- Average transaction amount
- Repay ratio = repays / borrows
- Liquidation ratio = liquidations / total transactions

Scoring Logic

The final score combines positively and negatively weighted behaviors:

- Rewards for deposits, repays, and redeems
- Penalties for borrows and liquidations
- Higher scores for consistent repayments
- Additional adjustment using average transaction amount (log-transformed)

The final score is normalized to a 0-1000 range using Min-Max scaling and clipped within bounds.

Architecture

- R-based script using jsonlite, dplyr, ggplot2, and scales
- Reads and flattens raw JSON
- Cleans and transforms data
- Scores wallets and saves output as CSV

- Produces an aggregated credit score distribution chart

File Descriptions

- score_wallets.R: Main processing script
- wallet_credit_scores.csv: Output credit scores for each wallet
- credit_score_distribution_improved.png: Visual distribution of credit scores
- analysis.md: Detailed insights from the scoring output

How to Run

1. Place the JSON file in your desired path.
2. Update the path in score_wallets.R accordingly.
3. Run the R script.
4. Outputs will be saved in the specified output directory.

Dependencies

- R
- Packages: jsonlite, dplyr, ggplot2, scales

Author

Bhukya Abhinay

---

analysis.md

Score Distribution Summary

This document analyzes the results of scoring 3497 unique wallets based on Aave V2 transaction data.

Score Distribution (Range Buckets)

The score distribution was grouped into buckets with 100-point width:

Score Range	Wallet Count
0 - 99	20
100 - 199	48
200 - 299	132

300 - 399	221
400 - 499	415
500 - 599	569
600 - 699	643
700 - 799	696
800 - 899	556
900 - 1000	197

Total Wallets: 3497

Behavioral Analysis

Low Credit Score Wallets (0-200)

- Very low transaction counts
- High liquidation to transaction ratios
- Frequent borrowing with no repayment
- Possible signs of exploitation or passive abandonment

Mid-range Score Wallets (400-700)

- Balanced activity across borrowing and repaying
- Occasional liquidations observed
- Average transaction amounts clustered around typical values

High Score Wallets (800-1000)

- High repays relative to borrows
- Frequent deposits and redemptions
- Near-zero or zero liquidations
- Regular transaction volume with stable patterns

Conclusion

This scoring model captures responsible wallet behavior in DeFi. The analysis reveals strong behavioral clustering by score, making this a viable scoring prototype for future lending risk frameworks in decentralized finance systems.