FITE4801 Interim Report

Account-based risk assessment helper for DeFi lending

Sistine Sichen YU	3035772949
Edith Huanyan GAO	3035772846

Date of Submission: Jan/21/2024

Table of Contents

Abstract

Overcollateralization in Defi lending is expected to wane and under-collateralization will flourish due to the growing need to unlock liquidity and attract fresh capital. To reduce the lending risk in under-collateralization, a reliable credit scoring system for borrowers needs to be established. Existing scoring frameworks both in traditional lending and blockchain-based lending have been researched. This project proposes a new model that can measure the credit risk level of accounts on the blockchain by examining the potential liquidation probability of accounts and making the model accessible through an interactive webpage.

This project can offer DeFi lending users and institutions unbiased references of the risk level of potential borrowers. It is hoped that this project can help build trust in the DeFi lending economy, thus promoting the development of under-collateralization and long-term prosperity of DeFi lending.

This report offers a thorough examination of the project's context, assesses the advancements achieved thus far, elucidates the challenges encountered, and presents the concluding remarks in the final section. The model optimization, along with the risk score calculation, has been ongoing. The subsequent phase will concentrate on the website construction.

Acknowledgment

We would like to extend our sincere appreciation and gratitude to Dr. Kam Pui CHOW for his guidance on our Final Year Project. He has given us precious suggestions on data and model selections, as well as advice on cryptocurrency markets.

Abbreviation

API	Application programming interface
FICO	Fair Isaac Corporation
DeFi	Decentralized Finance
KNN	K-Nearest Neighbors algorithm
LP	Lending Protocol
LR	Logistic Regression
WOE	Weight of Evidence
IV	Information Value

List of Figures

List of Tables

I. Introduction

Chapter 1 is the introduction to the whole project. Section 1.1 discusses the background information, including 1.1.1 DeFi Lending Overview and 1.1.2 Liquidation and Default Risk. Sequentially, sections 1.2 and 1.3 demonstrate the motivation and objective. After that, section 1.4 discusses the current research work, and section 1.5 clarifies the research gap between the DeFi Lending Risk landscape and our project.

1.1 Background

The following sections will give a brief overview of the Decentralized Finance market and lending protocols.

1.1.1 DeFi Lending Overview

The market value of Decentralized Finance is estimated at 16.33 billion USD in 2023, with the growth rate forecasted at 46% from 2023 to 2030. [1] As an evolutionary development in the DeFi ecosystem, DeFi lending is one of the largest DeFi markets with a TVL(Total Value Locked) of $14.95 billion, accounting for 40% of total TVL in DeFi as of June 2023. The top lending protocols (LPs) are AAVE, Compound, and JustLend. [2] In a normal DeFi lending flow, users have to deposit or lock their crypto-asset holdings in the liquidity pool beforehand as collaterals, and the liquidity position grants users the power to borrow crypto-assets with lending terms highly driven by the supply-demand dynamics. [3]

1.1.2 Liquidation Mechanisms

There is an assortment of factors engendering loan default and collateral liquidation. In AAVE, the index called health factor is devised to measure the collateralized positions of borrower accounts. If it drops below 1, liquidators can execute the liquidation calls towards those accounts, which involve paying back the debt and receiving discounted collaterals. [4] From the health factor calculation, it can be inferred that the volatility of collateral price plays an important role. The spike or plunge of crypto collaterals can induce massive liquidation in Defi protocols, intensifying the fragility of the market. One note-worthy stress-test event is the 'Black Thursday' event on 12 March 2020, when the 20% drop in Ether caused 16 million liquidations in LPs. [5]

Within the Compound Protocol, the possibility of liquidation looms over negative liquidity accounts. The liquidity assessment of these accounts involves multiplying the supplied balance by the collateral factor prevailing in the market, and subsequently summing the results while deducting any borrowed balances. Notably, both asset borrowing and asset withdrawal actions diminish Liquidity and augment the account's vulnerability. [6]

1.1.3 Liquidation and Default Risk

Beyond the market environment and account position, past user behavior can also be taken into consideration when making loan and liquidation decisions. It has been researched that accounts that have a default history hold only a small percentage of the participants' population, yet they are responsible for a large portion of the liquidated debts. Figure 1 shows that for May 2022, 13% of borrowers that had been liquidated could take 35% of the liquidations in the June crash in that year. [7] Hence, it can be inferred that the myriad risks are accounts-oriented, and account-based risk is more of the focus.

Figure 1 *Liquidated debts and liquidated borrowers*

1.2 Motivation

Undercollateralized lending is an inevitable trend to sustain the development of DeFi lending. This can be initiated by lending terms customization based on borrowers' creditworthiness. However, depending on the inherent design of current LPs, all borrowers are treated equally regardless of their historical transaction history or their repayment ability. Borrowers are not incentivized to keep themselves trustworthy, thus it is hard to push forward the under-collateralization transformation. To assess and minimize the undercollateralized lending risk, a model that can analyze all-around account-based information is needed to grade the borrower's risk level and provide references for loan approvals, interest rates, and various collateral requirements.

1.3 Objective

This project aims to mitigate heterogeneous DeFi lending risk through machine-learning-based assessment methods and web tool development. The project's target users, LPs and lenders, can use the platform prototype to query various lending risk-related information. The project consists of two layers: on-chain data analytics and refinement of credit scoring machine learning models, as well as the web tool design and development. The scope is limited to the leading LPs first, including AAVE and Compound, and expanded to other networks if possible. The project is expected to offer a more resilient, scalable, and responsive solution to DeFi lending risk control, thus maintaining the stability and security of the DeFi market.

1.4 Review of current work

The following three sections will discuss existing traditional banking credit rating model and blockchain data-based data analytics models.

1.4.1 FICO

FICO is a traditional loan credit scoring standard established in 1956, and is used by 90% of top lenders in traditional lending. It is a straightforward calculative formula with 5 components: payment history(35%), amount of debt(30%), length of credit history(15%), new credit(10%), credit mix(10%). Specifically, new credit indicates whether several credit accounts are opened in a short amount of time. The credit mix represents a combination of credit cards, retail accounts, installment loans, and mortgage loans. [8]

1.4.2 Octan Network

Octan is a social reputation ranking and on-chain data analytics model that makes use of the Google PageRank algorithm. It can be used by marketing agencies and DeFi protocols to qualify, classify and segment users, and filter out bots or cloned accounts. It extracts data by unifying user data across multiple blockchains, and implements categorization based on different user personas. Users have to mint its unique soulbound token(SBT) which is available on BNB Chain and its second-layer chains to carry their reputation scores in order to prove their trustworthiness. [9]

1.4.3 Account protection score by Block Analitica

It is a heuristic model that looks into multiple facets of account positions in AAVE, Compound, and MakerDAO, including current state and historical behavior related to liquidation protection actions. It utilizes an intuitive yes or no decision tree, as when some factors of an account meet certain numerical conditions, the strategy automatically lays the account into a category. It can empirically predict the likelihood of an account being liquidated shortly by categorizing accounts into low risk, medium risk, and high risk. [10]

1.5 Research Gap

The models researched all have their strengths and potential, but many of them are lacking in adaptability to the DeFi Lending to provide quality credit reference on DeFi loan default probability. For FICO, the factors considered for calculation range from personal bio-metadata to mortgage loan history, which are agnostic on blockchain as all the identities are pseudo-anonymous. Octan Network is a powerful project aimed at unleashing DeFi lending liquidity, but it has the issues of limited scope of application in BNB chains and too complicated procedures of minting SBT tokens which have barely intrinsic values. As for the account protection score, heuristics may lack certain properties that a machine learning model has, which may be too easy to crack and fake for users after they know the categorizing criteria. To solve these emerging issues, the project aims to build a machine-learning-based cross-chain credit scoring model that only depends on on-chain records and protocol data and is free and convenient for reference.

1.6 Outline

The Section 2 will elucidate the methodological aspects encompassing data collection and cleansing, machine learning model training, and webpage development. The third section will scrutinize the completed task, exhibit initial findings, and deliberate on encountered challenges.

II. Methodology

The subsequent sections, Sections 2.1 to 2.5, will elucidate the methodological aspects encompassing data collection and cleansing, machine learning model training, and webpage development.

2.1 Data Collection and Preprocessing

The raw data for blockchain addresses and DeFi transactions were obtained from two prominent lending protocols, namely AAVE and Compound. During the data collection process, it is discovered that both AAVE and Compound no longer maintain their own APIs and have instead entrusted this task to the Graph, an application designed to organize and serve blockchain data. The research is conducted using the Graph Playground to obtain the raw protocol data. Additionally, supplementary information is acquired by utilizing the Etherscan API, 0xWeb API, and the web3 library.

The project benefits from the inherent characteristics of immutability and transparency of LP data, rendering it a dependable source for the project. Another source of data is derived from third-party entities, although their authenticity may be comparatively weaker than that of direct data sources. Consequently, the project prioritizes the utilization of LPs' primary data and employs third-party data as a cross-validation approach.

Preprocessing steps also encompass addressing potential issues of dataset imbalance and redundancy. Techniques such as resampling and normalization are employed to tackle these challenges.

2.2 Model Training and Estimation

The task of determining whether a borrower will face liquidation can be framed as a binary classification problem, which can be effectively addressed using machine learning models designed for classification tasks. A wide range of models crossing from conventional ones like Logistic Regression and K Neighbors Classifiers to more advanced ensembled ones like XGB classifiers and Adaboost classifiers are attempted. The project involves testing these models, selecting a benchmark model based on its performance, and subsequently fine-tuning the parameters and hyperparameters to ensure suitability for DeFi lending scenarios.

2.3 Evaluation Metric Design and Result Interpretation

The Area Under Curve (AUC) serves as a measure of model accuracy, with a value closer to 1 indicating higher accuracy. Additionally, there exist numerous evaluation metrics to be considered, including but not limited to precision, recall, F1-score, false positive rate (FPR), and ROC curve (in the case of imbalanced datasets). It is imperative to carefully examine these metrics and establish an appropriate evaluation criterion. After selecting the best model, the predicted probability of default can serve as the input of the scoring card mechanism to calculate the user's risk level.

2.4 Website development and account risk visualization

The user can input the address or select from the address sample list. The laptop will send the request to the server and process it, then it will look up in the database and return the response back to the user.

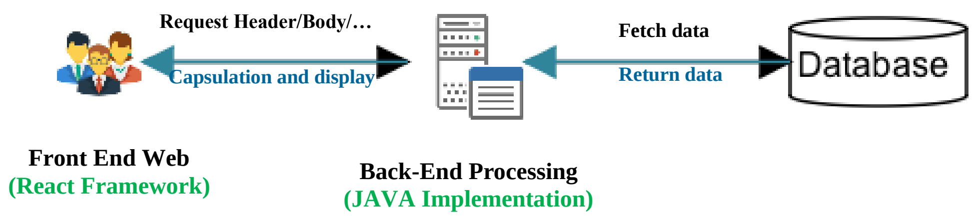

Figure 2 Query and Response Model

2.4.1 Webpage Development

Upon accessing the address score page, the user will be presented with a comprehensive overview consisting of summary and data visualization sections. The summary section showcases the credit scores assigned to the accounts directly derived from the machine learning algorithm. In the data visualization section, users will have access to charts and trading history, facilitating a detailed understanding of the analysis pertaining to addressing liquidation risk.

2.4.2 Back-End development

It is designed to fetch the data from the database effectively and preprocess it before returning the result to the user page. It should conduct basic operations and interact with other APIs if needed.

III. Current Progress

In this section, the project progress up to the submission of this report, will be discussed.

3.1 Data collection

There isn’t a credible comprehensive dataset featuring the information needed available online, so we established a customized framework for data collection and dataset formatting. After studying the calculating components of the prevalent traditional finance scoring model FICO[8], AAVE account protection score by Block Analytica[10], and an exploratory data analysis on DeFi lending users' health[11], 16 features that might be influencing factors are selected. As in Table 1, the account data information is split into 2 categories, account lending pattern, and account transaction history. The account lending pattern data is fetched from the subgraphs on the Graph platform through paginated queries, with some feature data calculated with subsequent treatment like 'averageBorrowAmont'. The account lending pattern data is downloaded and further processed from the Etherscan API. Finally, the account health factor of AAVE is collected by interacting with on-chain smart contracts using the function getUserAccountSummary(). While for Compound, the user health is fetched from the hosted service on the Graph platform. The dataset excludes inactive users who have no outstanding debts and invalid health factors.

Account Lending Pattern Features	**Account Transaction History Features**
borrowCount,depositCount,liquidateCount,liquidationCount,repayCount,withdrawCount,closedPositionCount,openPositionCount,positionCount,maxBorrowAmountUSD,totalBorrowAmountUSD,earliestBorrowTime,averageBorrowAmountUSD,	etherBalance,transactionsCount,accountCreationTime

Table 1 *Model Training Features*

We increment the health factor decision boundary of bad borrowers from 1 to 1.05, as 5% of the liquidation range is already a high-risk loan indicator and can be liquidated in case of minor market fluctuations. [12] We acquired the account information dataset AAVE v2/v3 with a size of 7979, and the dataset size of the Compound v2 protocol is 4683. The manual process of data collection ensures data integrity and timeliness.

3.2 Feature selection

Weight of Evidence(WOE) and Information Value(IV) are powerful techniques of binning in variable transformation and selection, which are widely used in credit scoring to measure the distribution of good and bad borrowers. The formula of WOE and IV is shown in Figure 4. After applying the technique to the DeFi borrower account features, we filter out comparatively more efficient predictor features with IV larger than 0.02 to use in the succeeding modeling. We also use Pearson Correlation to remove variables with strong correlations larger than 0.8.

$$WoE = \left[\ln \frac{Relative\ Frequency\ of\ Goods}{Relative\ Frequency\ of\ Bads}\right] * 100$$

$$IV = \sum(DistributionGood_i - DistributionBad_i) * WoE_i$$

Figure 3 *WOE IV calculation*

3.3 Insights from Exploratory Data Analysis

Due to the disparate liquidation strategies employed by the two protocols, we have segregated the data processing for AAVE and Compound, despite employing an identical dataset collection process. The subsequent two subsections will delve into the examination of the data before the training of models, with the aim of elucidating the underlying rationale.

3.3.1 AAVE decreasing default ratio in 2022 and 2023

Through our observations, we have noted a diminishing ratio of defaults, referencing the Figure 4, which is obtained by dividing the number of defaults (1) by the number of fulfillments (0). The DeFi market was significantly impacted by the decline in cryptocurrency prices and disruptions within decentralized exchanges (DEX). During the years 2022 and 2023, the DeFi lending sector experienced a contraction in market capitalization, potentially leading to a decrease in the influx of new accounts into the lending market. The prices of Ethereum (ETH) and Bitcoin (BTC) notably surged throughout 2021 and 2022, primarily due to their widespread usage as collateral cryptocurrencies. This surge, however, resulted in a devaluation of collateral assets held against debts, consequently negatively affecting the market liquidity and overall health of accounts. Nonetheless, as collateral prices subsequently recovered, the financial well-being of participants also improved. [13]

Furthermore, AAVE's annual report indicates that practical risk management measures have been implemented in the lending market, which may explain the observed increase in the proportion of healthy accounts over the past two years. [14]

3.3.2 Compound Position Closed Ratio Feature

The health status of Compound users is compared with the AAVE. The Information Value of features varies between AAVE and Compound, which necessitates the separate handling of data. The closed ratio feature is calculated by the closed position over the total position history, representing the debts successfully repaid. A higher closed position rate generally indicates a lower level of financial strain or a stronger ability to repay. This feature holds the highest Information Value (IV) ranking among all features in the Compound Protocol. Comparatively, as shown in the Table 2, Compound accounts tend to exhibit smaller open positions, larger total positions, and higher closed ratios when compared to AAVE. AAVE's protocol-based liquidation system mitigates the likelihood of borrower liquidation, while Compound incentivizes liquidators to initiate liquidation transactions, resulting in penalties for borrowers. Consequently, borrowers on the Compound Protocol are inclined to prioritize early debt repayment and uphold their level of exposure. This discrepancy may account for the higher closed ratio and lower prevalence of open debts observed in the Compound. [15]

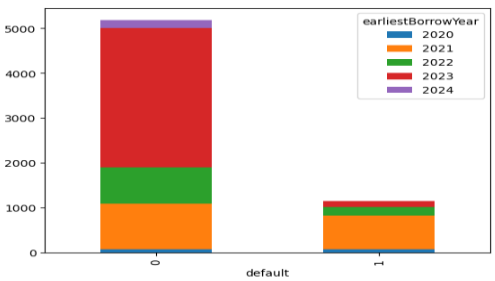

Figure 4 *Account Creation Year and Default Distribution*

	Open Position		Closed Ratio	
Protocol	Compound	AAVE	Compound	AAVE
Total Count	4683	3869	4683	3869
Mean	2.67	3.58	0.50	0.36
Standard Deviation	1.85	2.09	0.32	0.29
Min	0	0	0	0
25%	2	2	0.25	0
50%	2	3	0.57	0.33
75%	3	4	0.78	0.60
Max	17	23	1	1

Table 2 *Open Position and Closed Ratio Comparison between Compound and AAVE (v2 both)*

3.4 Model training and estimation

The credit scoring card mechanism is researched which includes default odds prediction and linear transformation of odds for credit scoring. Default status prediction is a supervised, binary classification problem given that we have available ground truth labels and two outcomes. Due to the inherent distinctive characteristics in AAVE and Compound datasets, two sets of models are trained separately.

For AAVE, in the data preprocessing, categorical data is labeled, and outliers are removed. The dataset is also balanced using SMOTE. It has been studied that classifier ensembles can significantly outperform single classifiers. [16] For the conventional machine learning model, we choose Logistic Regression because of its simplicity and interpretability, along with the k Neighbors Classifier and the Decision Tree Classifier. For ensembled models, we train the random forest model with hyperparameter tuning, XGB classifier with bagging, Adaboost and Gradient Boosting. Finally, a voting classifier is devised to combine the 3 best-performing models trained before, which are Random Forest, Adaboost and bagging XGB. The result aligns with the study, with the ensembled model having the highest accuracy of 84%, and the recall is 0.66. Recall is considered a major estimation metric because false negatives are more costly than false positives in this project's scenario. Considering the dataset is rebalanced, the model performs much better than random guesses and can provide some intuition about the account default risk. Additionally, the model also ranks the feature importance, showing that liquidationCount and earliestBorrowYear are the two most significant features, which is equivalent to the ranking of feature IV.

During the Compound model training, the Logistic Regression technique achieved the highest level of performance, with an accuracy of approximately 78% after undergoing dimensionality reduction. Moreover, the KNN method obtained a performance of around 74% accuracy. Additional data about the Compound Protocol is currently being processed, and our team intends to utilize this data to train the ensembled model in the forthcoming days.

3.5 Challenges and difficulties

The following subsections will go into the challenges we have encountered.

3.5.1 Etherscan API limited maximum query rate

Ethereum on-chain data queries are demanded to fulfill the account transaction history features. However, the free version of the official Etherscan API offers a limited query rate of 5 calls/sec, which is too slow for such a cumbersome data size. To solve this issue, we utilized another on-chain data query API called infura, but traded off some features that require further processing of transaction history like the number of 7-day transactions.

3.5.2 Compound Data Quality Concern

Throughout the process of data collection, it was observed that the Compound Protocol exhibits a lack of data maintenance practices, and there is inadequate disclosure regarding the calculation of its features. Notably, the V3 data Subgraph deployed on Ethereum is no longer accessible following the migration of the majority of transactions to Compound V3 protocols. To overcome this limitation and enhance the accuracy of the model, alternative approaches involving querying another subgraph deployed on Arbitrum One and the web3 library are being explored to query the protocol and obtain additional data.

IV. Future Work

The subsequent two subsections will comprehensively cover the remaining tasks and outline the proposed schedule.

4.1 Continue Machine Learning Model Training (1 February)

The training of different algorithms is currently in progress. For the first layer of the credit rating model, which is a crucial component of the project, we aim to achieve 90% accuracy. After acquiring statistics for different models, we are considering ensemble algorithms. Our objective is to end the model training by the end of January.

4.2 Web Page and Back-End Development (10 March)

In preparation for potential query retrievals, it is crucial to store the outputs generated by the machine learning (ML) model. In particular, the predictions pertaining to the likelihood of account liquidation will be stored within the database. The development of the interactive back-end and webpage will be carried out simultaneously.

V. Discussion

5.1 Cold Start Problem

The machine learning model leverages users' transaction and borrowing history as a basis for assessing the creditworthiness of accounts. However, for newly established accounts, the model may face limitations in gathering adequate information, which can consequently affect the accuracy of its output. In such scenarios, it may be prudent to consider only utilizing the original health factor of Lending Protocols to justify the health condition of these accounts.

VI. Conclusion

The primary aim of this project is to address the issue of insolvency in the Decentralized Finance sector by implementing enhanced credit rating models. The incorporation of machine learning techniques is anticipated to enhance the efficacy of the rating process. Initial findings from our model training indicate that the borrowing and transaction history of accounts do provide information on their overall health and can reveal their level of responsible borrowing. The ongoing model training process aims to further enhance the performance of the model. The future timeline is to conclude the development of layer one by the end of January, with specific attention directed towards the development and stabilization of the webpage in February.

VII. References

[1] “Decentralized finance market size & share [2023 report],” Decentralized Finance Market Size & Share [2023 Report], https://www.grandviewresearch.com/industry-analysis/decentralized-finance-market-report

[2] “Protocol categories,” DefiLlama, https://defillama.com/categories (accessed Oct. 24, 2023).

[3] Why Decentralised Finance (DEFI) matters and the policy implications - OECD, https://www.oecd.org/daf/fin/financial-markets/why- decentralised-finance-defi-matters-and-the-policy-implications.htm.

[4] AAVE. (n.d.). Liquidations. Liquidations - Developers. https://docs.aave.com/developers/guides/liquidations

[5] E. Frangella, “Crypto black Thursday: The good, the bad, and the ugly,” Medium, https://medium.com/aave/crypto-black-thursday-the-good-the-bad-and-the-ugly-7f2acebf2b83.

[6] DEFI lending: Intermediation without information?, https://www.bis.org/publ/bisbull57.pdf

[7] Hugh, “Introducing arcx credit and the Defi Credit Score,” Introducing ARCx Credit and the DeFi Credit Score - by Hugh, https://arcx.substack.com/p/introducing-arcx-credit-and-the-defi.

[8] “What is a FICO score and why is it important?,” myFICO, https://www.myfico.com/credit-education/what-is-a-fico-score.

[9] “Home - Octan Network,” Octan Network - Establish Trust and Credibility in Web3, https://octan.network/ (accessed Oct. 24, 2023).

[10] J. Osolnik, “AAVE account protection score,” Medium, https://medium.com/block-analitica/aave-account-protection-score-3bf8b7bc6f8b (accessed Jan. 2024).

[11] Alberquilla, “Exploratory Data Analysis on DEFI Lending Users Health,” Medium, https://towardsdatascience.com/exploratory-data-analysis-on-defi-lending-users-health-d8f8ff7cabbc (accessed Jan. 2024).

[12] J. Chiu, E. Ozdenoren, K. Z. Yuan, and S. Zhang, “On the fragility of defi lending,” *SSRN Electronic Journal*, 2023. doi:10.2139/ssrn.4328481

[13] M. Shen, “DEFI is becoming less competitive a year after FTX’s collapse battered crypto,” Bloomberg.com, https://www.bloomberg.com/news/articles/2023-11-16/defi-is-becoming-less-competitive-a-year-after-ftx-s-collapse-battered-crypto?embedded-checkout=true (accessed Jan. 21, 2024).

[14] Gauntlet et al., “[ARFC] gauntlet <> AAVE renewal 2023,” Aave, https://governance.aave.com/t/arfc-gauntlet-aave-renewal-2023/15380 (accessed Jan. 21, 2024).

[15] Okx, “Aave vs compound: A DEFI lending platform comparison,” OKX, https://www.okx.com/cn/learn/aave-compound-defi-lending (accessed Jan. 21, 2024).

[16] S. Shi, R. Tse, W. Luo, S. D’Addona, and G. Pau, “Machine Learning-Driven Credit Risk: A Systemic Review,” *Neural Computing and Applications*, vol. 34, no. 17, pp. 14327–14339, 2022. doi:10.1007/s00521-022-07472-2