Comparative Study of DEFI Lending Protocols using LSTM

Sankalp Bhoyar[1], Sagar Chakraborty[2], Pahuni Choudhary[3], Kisor Ray[4], Biswabandhu Jana[1]

Atal Bihari Vajpayee, Indian Institute of Information Technology and Management, Gwalior, India[1]
Budge Budge Institute of Technology, Kolkata, India[2]
Lakshmi Narain College of Technology, Bhopal, India[3]
Techno India University, Kolkata, India[4]

Corresponding author: Sankalp Bhoyar, Email: sankalp121314@gmail.com

Decentralized finance is a revolutionary change in the financial system, using blockchain technology to build a diverse and open network of financial services. By cutting out middlemen, DeFi promotes financial equality and reaches out to more people. Smart contracts can complete transactions without the need for traditional banks to lower costs and improve efficiency in lending, borrowing, trading, and yield farming activities. The proposed research compares four leading DeFi lending protocols: AAVE, MAKERDAO, COMPOUND, and VENUS Finance. We have used Long Short-Term Memory (LSTM) neural networks to analyze historical data and measure key parameters, such as lending and borrowing rates, Total Value Locked (TVL), Market Capitalization, and token price dynamics. We found that AAVE and COMPOUND exhibit similar mean rates but AAVE offers more precise predictions. MAKER provides potentially higher returns but with a higher degree of unpredictability. VENUS, despite its precise predictions, yields the lowest returns due to its lower mean lending rate. Overall, the approach enhances the understanding of the dynamics within the DeFi ecosystem, helping stakeholders to make informed decisions. Index Terms—Decentralized, Blockchain, LSTM, Time Series Analysis.

Keywords: DeFi, Long Short-Term Memory (LSTM), Blockchain, AAVE, Time Series Analysis.

in: *Advancements in Communication and Systems*. Ed. by Ashish Kumar Tripathi and Vivek Shrivastava. Computing and Intelligent Systems, SCRS, India., 2023, pp. 583–591. DOI: https://doi.org/10.56155/978-81-955020-7-3-52

1 Introduction

Blockchain technology has spawned a new form of financial services called decentralized finance (DeFi), which provides open, trustless, and permissionless access to various financial products [1]. DeFi has garnered considerable attention from both institutional and retail investors due to its potential to transform the traditional financial landscape. Lending protocols are one of the most important and popular DeFi applications, as they enable users to lend and borrow cryptocurrencies and digital assets with attractive interest rates and without intermediaries [2]. Lending protocols play a crucial role in the DeFi ecosystem by facilitating liquidity provision and interest generation for both lenders and borrowers.

The DeFi space is constantly evolving, and the competition among lending protocols is fierce. Some of the leading protocols are AAVE, COMPOUND, VENUS Finance, and MAKERDAO, each with its own distinctive features, governance models, and token economics. DeFi users and investors need a thorough understanding of the performance and risk factors of these protocols to make informed decisions about where to invest their assets [3]. There has been limited research on integrating machine learning or deep learning techniques into DeFi lending protocols. This research paper provides such an understanding by conducting a comparative study and performance analysis of these four lending protocols.

This paper aims to address the urgent need for a comprehensive understanding of the performance and risk factors of four leading DeFi lending protocols: AAVE, COMPOUND, VENUS Finance, and MAKERDAO. To achieve this goal, we have used Long Short-Term Memory (LSTM) neural networks, a state-of-the-art deep learning technique for time series data analysis, to forecast and evaluate key performance indicators of these protocols, such as interest rates, token prices, liquidity utilization, governance participation, and more. LSTM models are effective in predicting and understanding complex sequences, making them a useful tool for assessing the performance of DeFi protocols that involve time-dependent data [4]. We have analyzed key parameters such as lending and borrowing rates, TVL (Total Value Locked) value, Market Capitalization, and token price dynamic. We aim to provide valuable insights into the performance and risk characteristics of these protocols, as well as their potential for long-term sustainability. This research will provide useful guidance for both DeFi enthusiasts and institutional investors who want to explore the rapidly growing DeFi landscape.

2 Methods

2.1 Dataset Description

The dataset serves as a very crucial foundation for the insights. The dataset was not randomly taken or assembled, rather, it underwent a meticulous process of data collection, curation and refinement. Some key principles of this process were:

- Data Collection Source: In this research, data have been collected from various credible DeFi platforms, such as DeFiLIama, CoinGecko, and CoinMarketCap as no single platform is sufficient to provide all the required data [5-7]. Distinct time frames have been designated for the respective datasets under analysis, with the temporal boundaries as follows: AAVE (03/12/2020 to 11/08/2023), MAKER (03/12/2020 to 11/08/2023), COMPOUND (16/07/2020 to 11/08/2023), and VENUS (31/10/2021 to 11/08/2023).
- Data Validation: To spot and correct anomalies, outliers, or incorrect data points, a thorough validation process was put in place. Careful data cleaning was used to fix any data inconsistencies.
- Manual Review: For the purpose of ensuring the data accuracy and applicability, our dataset underwent manual review.

Table 1: A snapshot of the dataset

START	END	OPEN	HIGH	LOW	CLOSE	VOLUME	MCAP	TVL
11-08-2023	12-08-2023	66.9641	66.9641	65.3395	65.9525	128607245.9	958314811.6	2889435253
10-08-2023	11-08-2023	67.3022	69.0488	66.8022	67.0256	141533157.3	976190392.6	2781729962
09-08-2023	10-08-2023	66.3546	67.9084	65.8749	67.4406	137584647.2	966238600.9	2806623077
08-08-2023	09-08-2023	64.9758	66.8669	64.6616	66.1994	141752268.3	952561693.3	2832729978
07-08-2023	08-08-2023	66.061	66.977	63.3895	64.8849	149256791.4	945923789.4	2816843138
06-08-2023	07-08-2023	64.4217	66.0573	64.1478	65.8081	132846567.9	945577324.5	2829416506

A snapshot of the dataset is presented in Table I.

- **START**: Represents the start date of the period.
- **END**: Represents the end date of the period.
- **OPEN**: Price at which a particular asset started trading at the beginning of the period.
- **HIGH**: Highest price at which the asset is traded during the period.
- **LOW**: Lowest price at which the asset is traded during the period.
- **CLOSE**: Price at which the asset was trading at the end of the period.

- **VOLUME**: Number of assets that were traded during the period.
- **MCAP**: MCAP or Market Capitalization, represents the aggregate market value of a specific cryptocurrency or token. It's determined by multiplying the current market price of each coin or token by the total number of coins or tokens in circulation.
- **TVL**: This stands for Total Value Locked, a metric in DeFi that measures the total value of assets that are currently being staked in a specific protocol [8].

2.2 LSTM model

LSTM is a variant of recurrent neural network (RNN) architecture, specifically engineered to overcome the challenges faced by conventional RNNs in retaining and recognizing long-term dependencies in sequential data. Designed to counteract the vanishing gradient issue, LSTM incorporates memory cells equipped with self-gating mechanisms, enabling the network to selectively retain or discard information across lengthy sequences. This feature renders LSTMs particularly adept at tasks related to time series data, natural language processing, and speech recognition. The architecture's proficiency in preserving contextual information over prolonged durations, along with its ability to tackle vanishing and exploding gradient problems, has been established as a formidable instrument in the sphere of deep learning.

Due to LSTM's aptitude for time series data, it serves as our primary modeling framework. LSTM excels in capturing sequential dependencies and patterns in time series data, enabling accurate predictions based on past observations. Its unique ability to remember and incorporate information from earlier time steps distinguishes it, allowing the capture of temporal dynamics and dependencies in time series data. This is crucial in DEFI lending, where interest rates are influenced by past trends, current market conditions, and user behavior. LSTM's efficacy in handling time series data, particularly in DEFI lending rates, informed our deliberate choice over alternative modeling techniques.

Table 2: Descriptive statistics for the dataset

METRIC	AAVE	COMPOUND	MAKER DAO	VENUS
High Value	664.96	909.33	6244.44	16.26
Low Value	51.11	26.67	508.17	16.26
ACTUAL MEAN	189.76	191.47	1757.18	18.17

Table II outlines the statistical measures of DEFI protocols, including the high, low, and mean values of their performance. It provides insights into the variability and average performance of these protocols, presenting ranges between the highest and lowest recorded values, as well as the central tendency depicted by the mean values.

2.3 LSTM Model Performance

The use of LSTM neural network models to forecast rates within each DEFI lending protocol forms the basis of our analysis. These recurrent neural network-based models are remarkably good at capturing the complex temporal dependencies present in time series data.

Model Architecture: The LSTM model consists of a series of linked layers, each of which plays a particular part in the modelling process.

- Starting with an input layer created to ingest sequences of historical lending rate data, the model is introduced. Each sequence in this case has 200 time steps, giving a significant historical context.

- LSTM Layers: Our model makes use of a number of LSTM layers to accurately capture sequential dependencies. The layers of the LSTM are set up as follows: 100 units make up the first LSTM layer, which is set up to return sequences. As a result, it retains the data's temporal context, which makes it ideal for modelling long-distance dependencies. Sequences are also returned by the second LSTM layer, which has 50 units, making it easier to extract intermediate patterns from data. Sequences are not returned by the 50-unit, final LSTM layer. Instead, it condenses the data from the earlier layers into a representation to get it ready for the last prediction step.

- The output layer, which consists of a single neuron, is in charge of forecasting the lending rate for the future.

- The mean squared error (MSE) is used as the loss function. The squared difference between the model's projections and actual lending rates is measured by MSE.

- The 'Adam' optimizer, a popular option for training neural networks, is used to optimise the model. Adam can handle a range of optimisation issues because it adjusts the learning rate during training.

Table 3: Parameters of the LSTM model

Layers	Type	Output shape
layer_1	LSTM	(None, 200, 100)
layer_2	LSTM	(None, 200, 50)
layer_3	LSTM	(None, 50)
dense_1	Dense	(None, 1)

Table III shows the parameters of the LSTM model.

Training and Testing: The performance evaluation of our model is dependent on the training and testing phases:

- Training: Historical rate information from each DEFI protocol is used to train the model. To reduce the mean squared error, it accomplishes this by repeatedly adjusting its internal parameters.

- Testing: Using a different dataset, the model's performance is assessed after training. The lending rate data in this testing dataset are new to the model and were not present during training. We can evaluate the predictive accuracy of the model by contrasting its predictions with the actual lending rates in this testing dataset.

3 Results

3.1 Predective Accuracy

Our LSTM model's main goal is to predict High Value (highest price at which the asset is traded during the period) with accuracy. We divided the data into training and testing segments. The Percentage of division for training and testing data is 75% and 25%. With its multiple LSTM layers and recurrent architecture, it is capable of capturing complex patterns and temporal dependencies, which are crucial for comprehending and forecasting DEFI High Values. To implement LSTM we used TensorFlow and Keras libraries [9-10]. While implementing LSTM we used 180 epochs to get optimal result. Several important performance metrics, such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Standard Deviation, and Symmetric Mean Absolute Percentage Error (SMAPE), are used to measure the precision of our forecasts [11]. These matrices are represented in Table IV which are crucial in enabling users and investors to make informed decisions.

3.2 Comparative Analysis

Our comparison of the DEFI lending protocols reveals the following distinct trends and patterns:

- Comparable mean rates between AAVE and COMPOUND show a similar range of interest rates for lenders. AAVE, on the other hand, typically provides predictions that are more precise and stable with lower RMSE, MAE, and standard deviation values.

- With a noticeably higher mean rate than the average, MAKER stands out and may offer lenders greater returns. Its higher RMSE, MAE, and standard deviation values, on the other hand, highlight its greater variability and unpredictability.

- While VENUS offers the most precise and stable predictions of rates, it also has the lowest mean lending rate, which means a relatively lower return for lenders.

Table 4: Error values on the Testing Dataset

METRIC	AAVE	COMPOUND	MAKER DAO	VENUS
ROOT MEAN SQUARED ERROR	197.96	283.66	1908.42	16.26
MEAN ABSOLUTE ERROR	166.12	212.65	1669.63	12.42
STANDARD DEVIATION	107.84	187.93	924.50	10.55
MEAN ABSOLUTE PERCENTAGE ERROR	199.32	199.16	199.92	196.22

These findings highlight the distinctive qualities and performance nuances of each DEFI lending protocol, giving users and investors useful information to help them make decisions about the DEFI lending market.

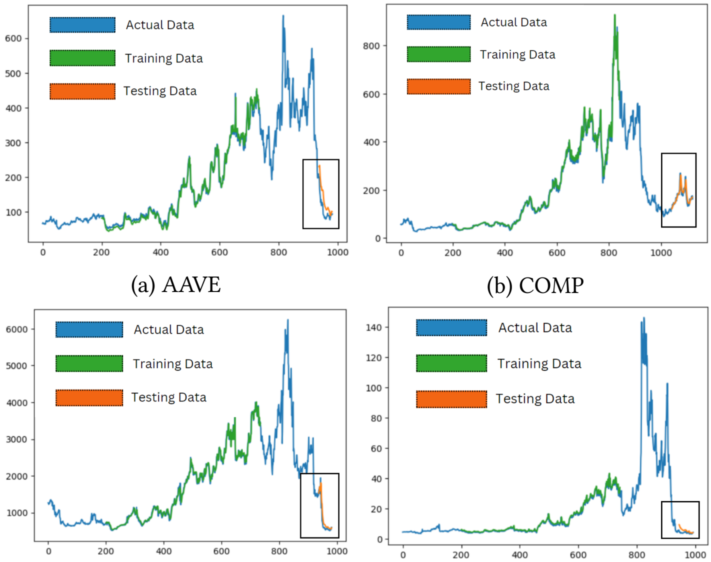
We can improve our comprehension of DEFI lending protocols and predictive modeling by addressing these constraints and investigating these new lines of inquiry, which will ultimately help users, investors, and other stakeholders in the dynamic world of decentralized finance.

4 Discussion

This paper presents a thorough comparison and analysis of four major DeFi lending protocols: AAVE, COMPOUND, VENUS Finance, and MAKERDAO. We used LSTM neural networks to measure and forecast their key performance indicators. The main goal was to give useful information about their performance and risk levels, as well as their long-term sustainability potential.

We discovered several important insights about these DeFi lending protocols. AAVE and COMPOUND had similar average interest rates for lenders, with AAVE having more accurate and stable forecasts. MAKERDAO had higher average rates, which could mean higher returns, but it also had more variation and uncertainty. VENUS Finance had the most accurate and stable forecasts but the lowest average lending rate, which could mean lower returns for lenders. Statistical tests confirmed the differences in average lending rates and predictive accuracy, highlighting the influence of platform policies, user demand, and general market trends. Our research followed a systematic approach for evaluating and comparing the performance indicators of these DeFi lending protocols, using carefully selected datasets and LSTM models designed for time series data.

While this paper provides valuable insights into the performance and risk factors of these four leading DeFi lending protocols, there are ways to improve and extend the research. Including more DeFi lending protocols and more variables in the dataset could provide a wider view of the DeFi ecosystem's behavior and better support decisionmaking. Improving deep learning models and exploring other machine learning techniques could lead to more precise and reliable forecasts. Developing real-time monitoring and prediction systems that can adjust to the fast-changing DeFi environment would be very useful for investors and users. Conducting deeper risk analysis, including factors like smart contract vulnerabilities and liquidity risks, can help to understand the risks associated with DeFi lending protocols better. Investigating user experience factors and the governance models of these protocols could offer additional insights into their appeal to participants and long-term viability.



(a) AAVE (b) COMP

(c) MAKER (d) VENUS

Figure 1: Graphical representation of Actual, Training and Testing Data

In Figure 1, the graphs exhibit the model performance on the training and testing data. The X-axis represents the time stamps in days whereas the Y-axis represents the price of a given DeFi protocol. The blue line indicates the actual data, the green line indicates the training data and the orange line indicates the testing data [8].

References

[1] D. A. Zetzsche, D. W. Arner, and R. P. Buckley, “Decentralized finance (DeFi),” J. Financ. Regul., vol. 6, no. 2, pp. 172–203, Sep. 2020.

[2] J. Xu and N. Vadgama, “From banks to DeFi: the evolution of the lending market,” in Future of Business and Finance, S. K. Das and A. K. Saha, Eds. Cham, Switzerland: Springer International Publishing, pp. 53–66, 2022

[3] A. Paliwal, “Analysis between different Decentralized Lending and Borrowing Protocols,” J. Bus. Anal. Data Vis., vol. 3, no. 1, pp. 15-23, 2022.

[4] Van Houdt, G., Mosquera, C. and Nápoles, G. ‘A review on the long short-term memory model’, Artificial Intelligence Review, 53(8), pp. 5929–5955, May 2020.

[5] DefiLlama: https://defillama.com/, accessed on August 11, 2023.

[6] CoinGecko: https://www.coingecko.com/, accessed on August 11, 2023.

[7] CoinMarketCap: https://coinmarketcap.com/, accessed on August 11, 2023.

[8] Shah, K. et al., ”A systematic review of Decentralized Finance Protocols”, International Journal of Intelligent Networks, 4, pp. 171–181, 2023.

[9] Metelski, D. and Sobieraj, J. ”Decentralized finance (DEFI) projects: A Study of Key Performance Indicators in terms of Defi Protocols’ valuations”, International Journal of Financial Studies, 10(4), p. 108, November 2022.

[10] TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, Available at: https://www.tensorflow.org/about/bib ,2015

[11] Chollet,Francois and others. Keras. Available at: https://github.com/fchollet/keras., 2015