

Blockchain-Based KYC Model for Credit Allocation in Banking

Ms CH. SIRISHA

Assistant Professor, Dept of CSE
CMR Technical Campus
Hyderabad, Telangana, India
chsirisha326@gmail.com

Ms B. POOJA REDDY

Assistant Professor, Dept of CSE
CMR Technical Campus
Hyderabad, Telangana, India
poojareddybejjanki@gmail.com

Mr B. ADITYA

Assistant Professor, Dept of CSE
CMR Technical Campus
Hyderabad, Telangana, India.
Adi.sacs@gmail.com

NATHA CHARMI

UG Student, Dept of CSE
CMR Technical Campus
Hyderabad, Telangana, India
nethacharmi@gmail.com

SINGIREDDY SATHVIK REDDY

UG Student, Dept of CSE
CMR Technical Campus
Hyderabad, Telangana, India
sathvikreddysingireddy0@gmail.com

YASARAPU KEERTHHI

UG Student, Dept of CSE
CMR Technical Campus
Hyderabad, Telangana, India
Keerthi80088008@gmail.com

Abstract:

The banking industry has witnessed a rise in the demand for robust and efficient security systems that provide a tamper-proof mechanism in the management of Know Your Customer (KYC) and credit risk assessment processes. This project aims at designing a Blockchain-based KYC system using Machine Learning (ML) for the automation of the Credit Risk Assessment process. The proposed system uses smart contracts on the Ethereum blockchain platform and a private blockchain environment called Ganache. The proposed system has a Django-based web application that connects with the blockchain environment and allows banking institutions to manage the KYC process efficiently. Moreover, the proposed system uses a machine learning model that predicts the loan approval process using key attributes of applicants. The proposed system integrates the benefits of blockchain and machine learning technologies in the management of the banking industry.

Keywords: Blockchain, Know Your Customer (KYC), Machine Learning, Smart Contracts, Credit Risk Assessment, Ethereum Blockchain, Loan Approval Prediction, and Financial Technology (FinTech).

INTRODUCTION

The banking industry relies heavily on the know-your-customer (KYC) process, in which banking institutions verify customers' identities and ensure regulatory compliance. However, the conventional KYC process is time-consuming, redundant, and costly, as customers must submit the same documents to various financial institutions. In addition, the centralized nature of conventional systems is vulnerable to security breaches, fraud, and a lack of transparency.

Fortunately, recent advances in digital technology have led to the introduction of blockchain technology, which has the potential to solve the problems associated with conventional systems.

This is because blockchain technology offers a robust, transparent, and tamper-proof environment that is suitable for storing sensitive financial information, such as customers' verified details. This would allow banks to access reliable customers verified information without undergoing the tedious and redundant verification process.

In addition, the application of the machine learning algorithm would help the banking system in the following ways:

- Improves accuracy of loan approval prediction
- Reduces human bias in decision-making
- Speeds up credit evaluation process
- Enhances risk assessment efficiency

This project aims to create a blockchain-based KYC model using machine learning to enhance the overall banking system. This project will use Ethereum-based smart contracts that will run on a private blockchain environment using the Ganache platform. This project will also use a Django-based web application to interact with the blockchain environment, and a machine learning algorithm will be integrated into the project to make loan approval prediction more efficient and intelligent.

I. PROBLEM DEFINITION

A. Repetitive and Inefficient Process

The existing Know Your Customer (KYC) process followed by most banking institutions is highly repetitive and inefficient. In the existing system, customers are required to submit the same identity proofs and other personal documents to different banks and financial institutions whenever they wish to obtain new financial services. Thus, the existing process makes the entire process highly inefficient and repetitive. Moreover, the lack of coordination between existing banks and financial institutions makes the entire process inefficient and repetitive.

B. Security and Data Privacy Issues

The existing Know Your Customer process followed by most banks and financial institutions makes use of a centralized database to maintain the confidential information of their customers. In the existing system, the centralized database makes the entire process highly vulnerable to cyber-attacks and other security threats. If the central server of the existing system is hacked, it can reveal the confidential information of a large number of customers. Thus, ensuring the security and privacy of customer data has become a major challenge for existing banks and financial institutions.

C. Lack of Transparency and Trust

The existing Know Your Customer process followed by most banks and financial institutions lacks transparency. In the existing system, banks and financial institutions maintain the entire credit approval process, and customers and other

financial organizations lack transparency regarding the process. Thus, the lack of transparency makes the entire process highly inefficient.

RELATED WORK

Various studies have been conducted to explore the implementation of blockchain technology to enhance the Know Your Customer (KYC) process in the banking industry. The conventional KYC process has often proven to be costly and time-consuming, as banks perform the verification process individually, resulting in a duplication of tasks. The conventional KYC process has often proven to be costly and time-consuming, as banks perform the verification process individually, resulting in a duplication of tasks. The implementation of blockchain technology has been suggested to enhance the process of carrying out the KYC process, as the technology has the capability to provide a highly secure environment for storing customer data.

A research study has suggested a decentralized verification of the Know Your Customer (KYC) process based on Ethereum blockchain technology, in which banks participating in the network have the capability to collectively verify customer data. The conventional process of carrying out the KYC process has often proven to be costly and time-consuming, as banks perform the verification process individually, resulting in a duplication of tasks. The implementation of blockchain technology has been suggested to enhance the process of carrying out the KYC process, as the technology has the capability to provide a highly secure environment for storing customer data. The implementation of blockchain technology has been suggested to enhance the process of carrying out the KYC process, as the technology has the capability to provide a highly secure environment for storing customer data. A research study has suggested a blockchain-based digital identity to enhance the process of carrying out the KYC process, in which a single trusted identity can be used to access multiple financial institutions, thus avoiding redundancy.

II. METHODOLOGY

The proposed system aims to incorporate a combination of blockchain and machine learning technologies to facilitate a secure and efficient KYC verification and credit allocation process for banking institutions. The methodology of the proposed system is as follows.

1. Data collection and user registration:

Customers must register through a web-based application and provide the required information for KYC verification. The system collects the required information from customers, including identity documents and other essential attributes required for verification and credit allocation.

2. KYC Verification using Blockchain

After the customers have provided the required information, the KYC data is verified and stored on a blockchain network. For this purpose, smart contracts of the Ethereum network have been used to securely record the verified information. As the blockchain network provides immutability and decentralization, the stored KYC data cannot be modified or altered.

3. Blockchain network implementation:

A private blockchain environment was implemented using Ganache to mimic the Ethereum network. On this network, smart contracts were used to manage the customer's KYC data and provide access control to authorized banking institutions.

4. Web application integration:

A web application has been designed based on the Django framework to provide a platform for users and banking institutions. Banks will be able to add, retrieve, and verify customer-related information through the provided interface, which will be obtained from the blockchain network. This phase of the design acts as a communication medium between the blockchain network and users.

5. Machine Learning-based Credit Prediction:

A machine learning model is trained based on various financial and personal factors of customers. The model predicts the outcome of customer-related data, which helps banks make decisions regarding loan approvals. This phase of the design helps banks make efficient decisions regarding credit allocation to customers.

6. Credit Allocation and Decision Support

After the verification process, the system predicts loan approval or rejection based on the output of the machine learning model. Banking institutions can utilize this information to make efficient decisions with respect to credit allocation to customers. The overall design methodology enables a secure identity verification process, eliminates redundancy in customer-related verification, and enhances the overall efficiency of the credit allocation process in the banking industry.

III. PROPOSED SYSTEM

The proposed system aims to address the weaknesses and flaws in conventional systems used in KYC and credit assessment processes. This is accomplished by utilizing blockchain technology and a machine learning-based loan prediction model. In the proposed system, the Ethereum blockchain, using the Ganache local implementation, is used to create a decentralized and unalterable ledger for customers' KYC information. This ensures that customer information is tamper-proof and can be accessed by all institutions without duplication.

Additionally, customer blockchain records contain the customer's full name, national ID, status, and the name of the originating bank. This ensures that customer records are updated in accordance with the set rules, thus preventing fraud and unauthorized access to customer records. Moreover, a transaction hash is created each time a customer's KYC is made, and it is used as a unique identifier in the customer database.

To enhance the secure identity verification process, a trained machine learning model has also been integrated to make predictions based on loan

approval decisions. The parameters considered during the process are various financial parameters, including age, income, work experience, credit score, and history of defaults. The data used to make predictions are preprocessed with label encoding before being fed into the model, which finally makes a binary prediction, that is, an approval or denial of the loan.

The proposed system overcomes the limitations of the conventional KYC and credit assessment processes with the integration of blockchain technology along with a machine learning model for loan prediction. The proposed system uses the Ethereum blockchain, which has been implemented locally with the help of Ganache. Blockchain technology was used to create a decentralized ledger for KYC. Once verified, customer information is stored in the blockchain ledger with the help of smart contracts.

The details included in the blockchain for every customer are the customer's name, national ID, KYC, and the bank's name from which the customer originated. Smart contracts help in updating the data according to specific rules, thus avoiding unauthorized access to the data. The hash of the transaction is created for every KYC transaction, which becomes a unique identifier for the customer's data stored in the blockchain and the database of the institution.

In addition to the secure identity verification process, a machine learning model has been integrated into the system to predict the outcome of the loan approval process. The model considers various parameters, including financial parameters, age, income, work experience, credit score, and previous defaults. The data is preprocessed using label encoding before being fed into the model, which predicts the outcome as approved or denied loan.

The system implementation utilizes blockchain technology, machine learning, and a web-based application framework to ensure secure verification of customer KYC and decision-making for credit assessment.

IV. IMPLEMENTATION DETAILS

A. BLOCKCHAIN IMPLEMENTATION

A private blockchain network will be created using Ganache software, which provides a simulated environment for Ethereum. Smart contracts will be developed and deployed on the blockchain network to ensure the secure storage and management of customer KYC information. Blockchain technology provides immutability, transparency, and tamper-proof storage of customer identity information. Each verified customer KYC information will be stored as a transaction on the blockchain network.

B. SMART CONTRACT DEVELOPMENT

Smart contracts will be developed to facilitate the execution of certain operations, such as adding new customer KYC information, retrieving existing information, and verifying identity records. These smart contracts will ensure the implementation of the rules and permissions for accessing and managing the stored information, allowing only authorized institutions to interact with the stored information.

C. WEB APPLICATION DEVELOPMENT

A web application will be developed using the Django framework, which provides a framework for the implementation of the system and allows the creation of a user-friendly interface for users and banking institutions.

D. BLOCKCHAIN INTEGRATION

The integration of the Django application with the Ethereum blockchain is performed using Web3 libraries. This allows for communication between the web application and smart contracts on the blockchain. The application can send transaction requests and retrieve stored KYC information.

E. MACHINE LEARNING MODEL

A machine learning model was created to predict loan approval decisions for applicants. The prediction is based on various applicant

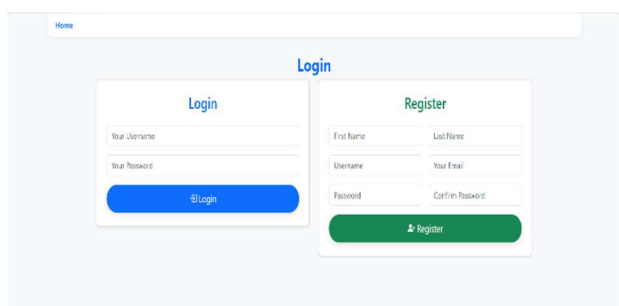
characteristics, including income, employment, credit history, and other financial information. The machine learning model was created using Python.

F. SYSTEM WORKFLOW

The system workflow involves a user registering on a web application and submitting their KYC information. The information is verified and securely stored on the blockchain through smart contracts. When a user requests a loan, their information is used by an ML model to predict their loan approval status. The banking institution can then make decisions regarding credit allocation.

The proposed system was implemented using a private Ethereum blockchain with Ganache and a Django-based web application. The results reveal that the proposed system can securely store and verify customers' KYC details on the blockchain, which is transparent and tamper-proof. In addition, the inclusion of a machine learning model in the proposed system was successful in providing automated predictions on loan approvals based on the details provided by the applicant. Overall, the proposed system successfully reduced the time taken to verify customers' KYC details, improved data security, and provided faster and more accurate results in credit operations.

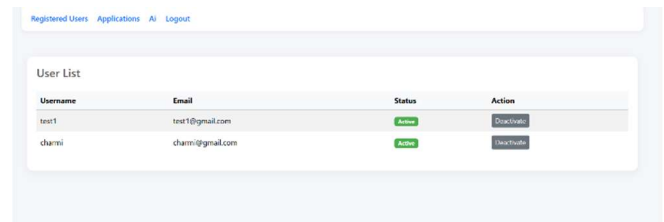
System Interface – Home Page:



The system has a user-friendly home page that acts as the interface for the users as well as the banking institutions. The home page has options such as user registration and login. Users who want to register with the system can enter the required personal information in the registration section of the home page. Users who want to use the system can enter the required information in the login section of the home page. This interface allows the users to

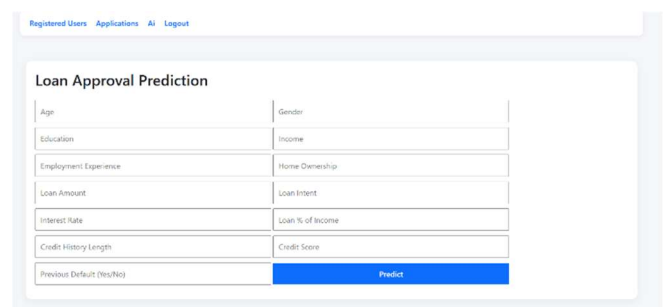
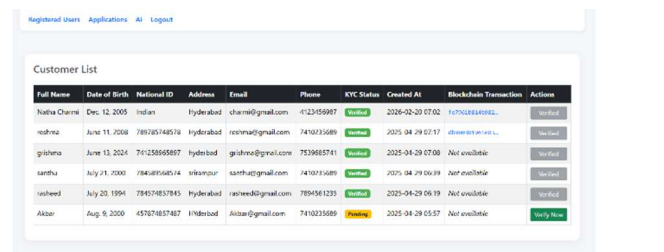
interact with the blockchain-based KYC verification and credit allocation system.

Fig. 1: Registered Users Page showing list of users in the system.



The page shows a list of users who have successfully registered in the system. It is used by authorized institutions to view and manage users for further verification.

Fig. 2: Customers List interface used for loan evaluation input.



This page enables users to fill in the necessary financial and personal information to be used in evaluating the loan. The process of evaluating the loan approval or rejection is done through a machine learning model.

VII. CONCLUSION

With the integration of blockchain technology and machine learning, an efficient solution is proposed to enhance the management of KYC in banking operations, along with the approval of loans. In this regard, the blockchain technology ensures the secure, transparent, and tamper-proof storage of customers' KYC details, which eliminates redundancy in the verification process. Additionally, the proposed machine learning model

facilitates an automated decision-making approach in the approval of loans, which provides accurate results. Overall, the proposed system ensures security, efficiency, and reliability in banking operations.

VIII. FUTURE SCOPE

The system can eventually be implemented on consortium or public blockchain networks to facilitate actual banking settings. Integration with national digital identity systems can lower fraud and enhance KYC verification. Additionally, the system can be expanded to handle the entire loan lifecycle, including tracking repayments and disbursements. Accessibility, security, and system management can be further improved with additional improvements like mobile applications, privacy-preserving methods, and monitoring dashboards.

IX. ACKNOWLEDGMENT

We would like to sincerely thank the CMR Technical Campus Department of Computer Science and Engineering for providing the lab space, computational infrastructure, and academic support needed to successfully complete this project. The development of our research work was greatly aided by the supportive learning environment and ongoing direction from the faculty.

We are particularly appreciative of Ch. Sirisha, our project guide, for her invaluable guidance, technical know-how, and unwavering support during the project. Her helpful criticism and perceptive recommendations enabled us to improve our approach and successfully apply the suggested blockchain-based KYC system and machine learning model for credit prediction.

X. REFERENCES

- [1] V. L. Lemieux, "Trusting records: Is blockchain technology the answer?" *Records Management Journal*, vol. 26, no. 2, pp. 110–139, 2016.
- [2] W. Viriyasitavat and D. Hoonsopon, "Blockchain characteristics and consensus in modern business processes," *Journal of Industrial Information Integration*, vol. 13, pp. 32–39, 2019.
- [3] S. Nakamoto, *Bitcoin: A Peer-to-Peer Electronic Cash System*, 2008.
- [4] S. Perera et al., "Blockchain technology: Is it hype or real in the construction industry?" *Journal of Industrial Information Integration*, vol. 17, pp. 1–20, 2020.
- [5] B. Karadag, A. Akbulut, and A. H. Zaim, "A review on blockchain applications in fintech ecosystem," in *Proc. Int. Conf. Advanced Creative Networking and Intelligent Systems (ICACNIS)*, 2022.
- [6] Ethereum, *Ethereum Whitepaper*, 2023.
- [7] N. Mansoor et al., "A review of blockchain approaches for KYC," *IEEE Access*, vol. 11, pp. 121013–121042, 2023.
- [8] D. George, A. Wani, and A. Bhatia, "A blockchain-based solution to Know Your Customer (KYC) dilemma," in *Proc. IEEE Int. Conf. Advanced Networks and Telecommunications Systems (ANTS)*, 2019.
- [9] D. Roman and G. Stefano, "Towards a reference architecture for trusted data marketplaces: The credit scoring perspective," in *Proc. Int. Conf. Open Big Data*, 2016.
- [10] H. Karaylan, "Blockchain and its applications for financial technology solutions," *Yüksek Öğretim Dergisi*, 2019.
- [11] H. Byström, "Blockchains, real-time accounting, and the future of credit risk modeling," *Ledger*, vol. 4, pp. 40–47, 2019.
- [12] S. Chakraborty et al., "A blockchain-based credit analysis framework for efficient financial systems," in *Proc. Int. Conf. Advanced Communication Technology*, 2019.
- [13] S. Patel et al., "KiRTi: A blockchain-based credit recommender system for financial institutions," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1044–1054, 2021.
- [14] F. Yang et al., "BACS: Blockchain and Auto ML-based technology for efficient credit scoring classification," *Annals of Operations Research*, 2022.
- [15] S. Wang et al., "Blockchain-enabled smart contracts: Architecture, applications, and future trends," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 49, no. 11, pp. 2266–2277, 2019.