JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

The Credit Recommender System using Block Chain

[1]Mandava Sandeep, [2]Mallu Yashwanth Reddy, [3]MrsV.Subapriya, [4]N. Senthamilarasi

[1]Student/Research Scholar, [2]Student/Research Scholar, [3]Student/Research Scholar, [4]Student/Research Scholar

Department of Computer Science and Engineering,

Sathyabama Institute of Science and Technology, Chennai, Tamilnadu , India

Abstract : Rather than relying on third-party rating agencies for credit scores, this paper proposes KIRTi, a deep-learning credit-recommender system based on public blockchain. PBs and PLs will be able to make smart loans to each other under the scheme. By securing, approving, and automating the loan grant process, PL was able to speed up the disbursement process to PB. As PB's assets, liabilities, and transactions are recorded on a public blockchain, KiRTi keeps track of their past operations. Based on the suggested lending algorithms for both PB and PL, a long-short memory (LSTM) model retrieves and processes blockchain sequence data.We update CS's edge weights in real-time based on PB and PL's boolean indications of repayment success and loan default. Iterating the procedure gets better edge-weights and CS, ensuring that PB is credible enough for future loans. In order to automate loan repayments, it is suggested that PB and PL use smart contracts (SC). LSTM recommender systems are trained by using German credit datasets from the UCI repository. PB's credit histories are contained in this dataset. 700 of these have been repaid successfully, and 300 have defaulted. Using KiRTi, 97.5% accuracy can be achieved with Fmeasure 0.98304. Based on the security assessment, KiRTi's compute and communication costs are 20.96 ms and 121 bytes respectively.

***IndexTerms* - Drug-related, Attack-related content, Positive online environment, Support Vector Machines (SVMs), Random Forests (RFs).**

I. Introduction

KiRTi's blockchain-based deep-learning credit- recommender scheme eliminates the need for credit rating agencies to generate credit scores by facilitating smart lending operations between prospective lenders and prospective borrowers. PL is therefore required to secure, approve, and automate loan grants to PB in order to speed up the disbursement procedure. KiRTi records PB's assets, liabilities, and past transactions on a public blockchain. As a result of the PB and PL lending algorithms, a long-short term memory (LSTM) model is used to retrieve sequencing data from the blockchain for the purpose of providing CS. As PB and PL provide boolean indications of successful repayments and loan defaults, edge weights are updated in real-time for CS.To make sure that PB is legitimate for future loans, the procedure is repeated until the edge-weights and produced CS are as accurate as possible. It is suggested that PB and PL use smart contracts (SC) to automatically set up loan repayments. We use a German credit dataset from the UCI repository to train our LSTM recommender system. This dataset contains 1000 credit-histories of PB, 700 of which have had successful repayments and 300 of which have defaulted.

Existing system

The current method of disbursing loans is cumbersome and time-consuming. This is because of the intricate web of relationships that exists among several factors, including as the borrower's credit history, the frequency of defaults, and the duration of payback cycles. Additionally, the degree of uncertainty around loan approvals varies according to the PL's expected profit margins. Some PL place less emphasis on gains (repayments) than loss (defaults). PL's risk assessment is proportional as a result of the recommended models providing different ratings for the same PB.Even a missing payments may impact CS, leading to the computation of high risk; hence, this technique is faulty. Since recommender models do not include new PB's credit histories, they also do not account for new PB with a strong reputation and significant assets.

Proposed system

The suggested system has a lengthy loan disbursement period. Prospective borrowers and lenders talk to each other face-to-face. In order for them to understand all of the intricacies. Here, we read borrower evaluations using a sentiment analysis. With the help of sentiment analysis, we can calculate the overall review score. Reviews and the ratio of positive to negative ratings are taken into consideration by PL when deciding whether to provide the loan. In order to make loans accessible to as many people as possible. The information on the prospective lender's loan and the prospective borrower's payment history will be recorded in the blockchain.

II. LITERATURE SURVEY

Seemann et al. (2023) conducted a comprehensive study on abusive language detection, addressing the challenge of model generalization. Despite the observed proficiency of models on training datasets, the authors noted difficulties in extending performance to new datasets. In contrast to prior research, they emphasized the use of homogeneous datasets, assuming enhanced generalizability. This study investigated the degree to which datasets must be similar for effective model generalization and found that the training method had a significant impact on generalizability. Four German datasets, including successive GermEval challenges, were analyzed using two deep learning models and three traditional machine learning models.The findings revealed partial generalization, even with nearly identical annotation schemes, highlighting that generalizability is contingent on specific training set combinations, irrespective of the underlying method.

Using online harassment to detect the severity of abusive comments, Marshan et al. (2023) address the issue of online harassment. An examination of the effects of text pre- processing and a comparison of machine learning models is presented in the study, which also explores the impact of neural networks on the result. A Random Forest model with bigrams performs best, achieving an accuracy rate of 94 percent, precision rate of 91 percent, recall rate of 94 percent, and F1 score of 92 percent.The research provides insights for theory and practical applications in combating abusive language on online platforms.

By introducing a Hybrid Deep Learning Architecture (HDLA), Kazbekova et al. (2023) address the prevalence of offensive language on online social networks (OSNs). To distinguish nuanced communication on OSNs, we propose a model that combines Convolutional Neural Networks (CNNs) and Long Short Term Memory (LSTMs). CNN reveals concealed offensive patterns by extracting spatial features, while LSTM captures context-dependent dependencies. Various measures of precision, recall, and F1- score demonstrate improved performance in detecting offensive language across a variety of subtleties and complexity levels.The HDLA offers insights into how offensive content propagates and can be used to foster healthier online communities through effective curation of digital content.

Khan et al. (2023) address the challenge of offensive language detection in low-resource languages like Pashto. They propose a model trained on a Roman Pashto dataset created from 60,000 manually labeled social media comments. The study evaluates three feature extraction approaches and employs traditional classifiers alongside a deep sequence model. Combining random forest classification with n-grams yields the highest accuracy(94.07%), whereas bidirectional long short-term memory gives the best overall accuracy (97.21%). Researchers working in this field can benefit from the Pashto corpus provided by the authors.

Panchala et al. (2022) address the pressing issue of hate speech and offensive language in the cross-cultural Internet era.With the surge in online condemnations, judgments, and trolling, the authors emphasize the need for restoring peace and harmony. They highlight the challenges of automatically detecting such content, especially in the context of 5G evolution and increased online engagement. Existing research is noted to be language-specific, limiting its applicability. Social media platforms are increasingly being used for offensive language, highlighting the importance of automatic methods to address the issue. Methods for detecting hate speech on Twitter include machine learning and natural language processing. Social media platforms are widely used to use offensive language, Hajibabaee et al. (2022) highlight its potential to harm individuals or communities. In their study, modular cleaning, tokenization, embedding methods, and eight classifiers are integrated into a robust model for text classification. TF-IDF embedding yields the best average F1- score when used on a Twitter dataset, with AdaBoost, SVM, and MLP showing the best results. To detect and prevent offensive language online, the research contributes valuable insight to ongoing research efforts.

In the context of natural language processing, Sokolová et al. (2022) present a study of hate speech and offensive language detection in Slovak.A database of short texts derived from Slovak social media is created and processed in the paper. In order to detect hate speech using sentiment analysis, recursive convolutional neural networks are combined with components from convolutional neural networks.Notably, the study reports a detection accuracy of 61.32% on a small set of training data, which is balanced in terms of positive, neutral, and negative sentiments. In particular, it discusses Slovak language and social media content in the context of hate speech detection.

In their paper (2022), Li et al. propose a deep semantic feature fusion approach for identifying offensive languages. By combining BERT word-level embeddings and RCNN with an attention mechanism, the model identifies languages that convey offensive meanings without using common offensive words. In this model, a label encoder and an offensive predictor reduce the need for offensive language lexicons. By understanding context, this model outperforms existing methods for detecting potential offensive meanings in Wikipedia and Twitter comments.

Kumar et al. (2021) address the surge in offensive language on social media platforms by proposing an automatic detection system. In contrast to other machine learning classifiers, they implement and compare LSTM and BERT models based on the Davidson dataset. It was found that their deep learning methods, especially LSTM and BERT, outperformed other classifiers in identifying offensive language in social media posts.

The methods of deep learning were used by Yadav et al. (2021) in order to detect hate speech and offensive language. This study examines recurrent neural networks (RNNs), convolutional neural networks (CNNs), long short-term memories (LSTMs), and transformer-encoder representations from bidirectional encoders (BERTs).The study investigates the impact of class weighting techniques on model performance. Findings indicate that the pre-trained BERT model excels in hate speech classification, while RNN and CNN models outperform others in offensive language classification. The study emphasizes the substantial improvement achieved by employing class weighting techniques across all models for hate speech detection.

III. METHODOLOGY

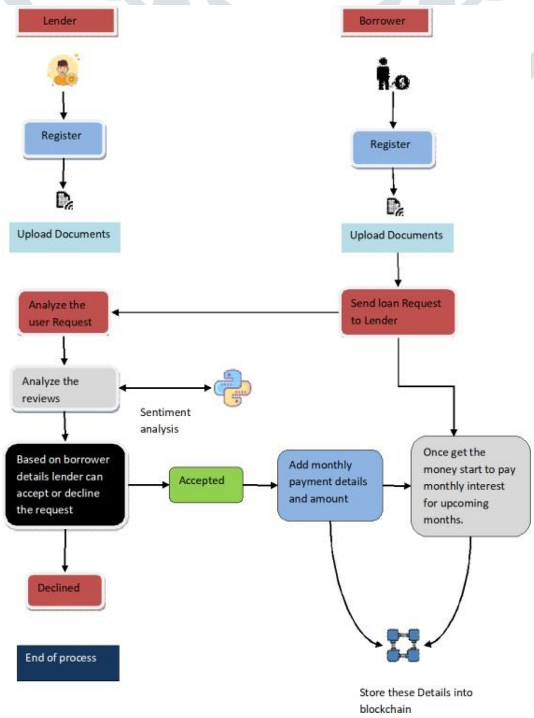


Fig 1: System Architecture

SYSTEM DESIGN

The suggested system has a lengthy loan disbursement period. Prospective borrowers and lenders talk to each other face-to-face. In order for them to understand all of the intricacies. Here, we read borrower evaluations using a sentiment analysis. With the help of sentiment analysis, we can calculate the overall review score. Reviews and the ratio of positive to negative ratings are taken into consideration by PL when deciding whether to provide the loan. In order to make loans accessible to as many people as possible. The information on the prospective lender's loan and the prospective borrower's payment history will be recorded in the blockchain.

MODULES

- PB and PL Registration
- Loan Request by PB
- Sentiment Analysis
- Payment Tracking

PB AND PL REGISTRATION

Prospective borrowers and lenders will fill out their profiles and income data in the first module. In addition, official papers provided by the government will be uploaded throughout the registration procedure.

LOAN REQUEST BY PB

The information of registered potential lenders will be shown as a list in the second module when the perspective borrower logs in. The borrower has the freedom to choose the lender and verify their identity by seeing official government documents. After that, he might approach the lenders with the request. All requests should be sent by a single PB to a single PL.

SENTIMENT ANALYSIS

The third module is where the lender's side lists the borrower's data. The lender has the right to verify the borrower's identity and verify their personal information. He will proceed to sentiment analysis when he has verified the fundamental procedures. By using sentiment analysis, the lender may determine the proportion of favorable and negative ratings.

PAYMENT TRACKING

The last module involves storing the borrower's monthly payment on blockchain. In order for creditors to be aware of when payments are due and if they are late or on time.

IV. RESULTS AND DISCUSSION

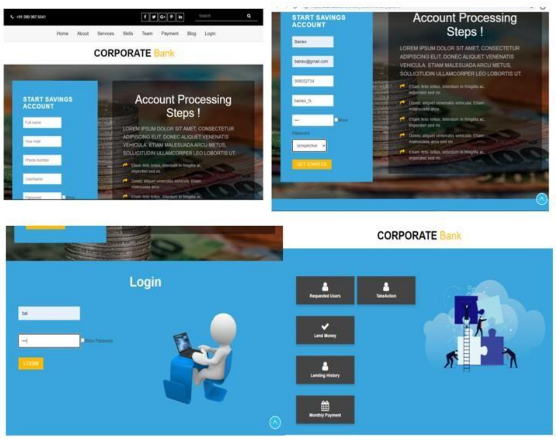
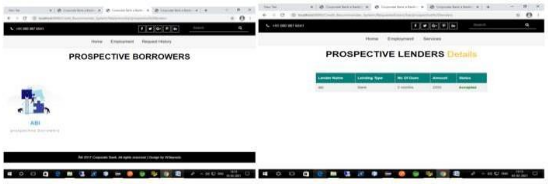


Fig 2: Log in and account processing



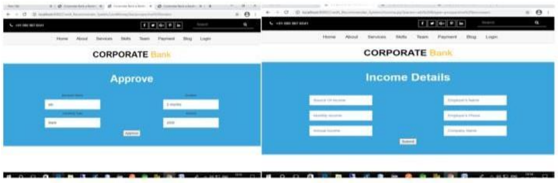
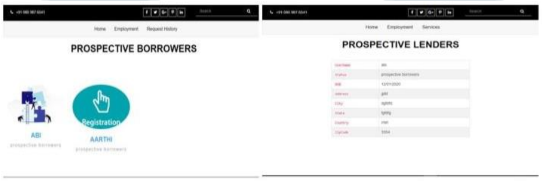


Fig 3: Prospective Borrowers and Lenders



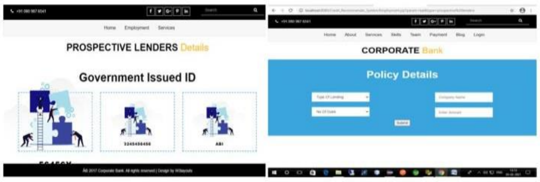


Fig 4: Approve and income details

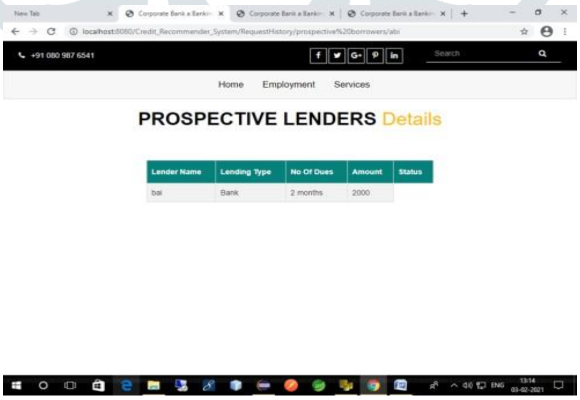


Fig 5: Prospective lenders

V. CONCLUSION

PBs are granted loans using recommender models by modern smart lending operations in financial institutions. CS production and loan award reviews will be transparent through the proposed approach.To enable a recommender ecosystem built on trust, KiRTi combines deep learning with blockchain technology. To guarantee precise predictions on PB historical repositories, KiRTi applies the LSTM model, and blockchain provides immutability in recorded data. In order to enhance decision analytics over time and decrease bias, we created a feedback method that changes the edge-weights of LSTM in real-time. SC need PB to PL to automate re-payments. In comparison to current credit systems, the anticipated outcomes show that the program is effective. But in order to update CS, the LSTM model uses feedbacks, which include latency overheads.

FUTURE WORK

To enhance the learning rate and optimize the loss functions, the authors want to enhance the regularization parameters in the future by ending training samples early. This will increase the total reaction time of CS generation, which will lead to speedier repayments while maintaining the appropriate degree of security and accuracy. It will also induce responsive feedbacks.

VI. REFERENCES

[1] N. Kabra, P. Bhattacharya, S. Tanwar, and S. Tyagi, “Mudrachain Blockchain-based framework for automatedcheque clearance in financial institutions,” Future Generation Computer Systems, vol. 102, pp. 574–587, 2020.

[2] “Equifax data breach settlement.” https://www.ftc.gov/enforcement/ casesproceedings/refunds/equifax-data-breach- settlement. Accessed: 2019-12-27.

[3] “Experian lapse allowed id theft service access to 200 million consumer records.” https://krebsonsecurity.com/2014/03/experian-lapseallowed- id-theft-service-to-access-200-million-consumer-records/. Accessed: 2019-12-27.

[4] P. Bhattacharya, S. Tanwar, U. Bodke, S. Tyagi, and N. Kumar, “Bindaas: Blockchain-based deep-learning as-a- service in healthcare 4.0 applications,” IEEE Transactions on Network Science and Engineering, pp. 1–1, 2019.

[5] H. Wang, C. Guo, and S. Cheng, “Loc — a new financial loan management system based on smart contracts,” Future Generation Computer Systems, vol. 100, pp. 648 – 655, 2019.

[6] S. Chakraborty, S. Aich, S. J. Seong, and H. Kim, “A blockchai based credit analysis framework for efficient financial systems,” in 201 21st International Conference on Advanced Communication Technology (ICACT), PyeongChang, Korea (s), pp. 56–60, Feb 2019.

[7] V. Thakur, M. Doja, Y. K. Dwivedi, T. Ahmad, and G. Khadanga, “Land records on blockchain for implementation of land titling in india,” International Journal of Information Management, p. 101940, 2019.

[8] V. Hassija, G. Bansal, V. Chamola, N. Kumar, and M. Guizani, “Secure lending: Blockchain and prospect theory-based decentralized credit scoring model,” IEEE Transactions on Network Science and Engineering, vol. PP, 03 2020.

[9] T. Yuan, H. Wu, J. Zhu, L. Shen, and G. Qian, “Ms-ucf: A reliable recommendation method based on mood- sensitivity identification and user credit,” in 2018 International Conference on Information Management and Processing (ICIMP), London, UK, pp. 16–20, Jan 2018.

[10] K. Ren and A. Malik, “Recommendation engine for lower interest borrowing on peer to peer lending (p2pl) platform,” in 2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI), Thessaloniki Greece,pp. 265–269, Oct 2019.

[11] N. Morar, C. Baber, F. McCabe, S. D. Starke, I. Skarbovsky, A. Artikis, and I. Correai, “Drilling into dashboards: Responding to computer recommendation in fraud analysis,” IEEE Transactions on Human-Machine Systems, vol. 49, pp. 633–641, Dec 2019.

[12] S. Cai and J. Zhang, “Exploration of credit risk of p2p platform based on data mining technology,” Journal of Computational and Applied Mathematics, vol. 372, p. 112718, 2020.

[13] P. Pławiak, M. Abdar, J. Pławiak, V. Makarenkov, and U. R. Acharya, “Dghnl: A new deep genetic hierarchical network of learners for prediction of credit scoring,” Information Sciences, vol. 516, pp. 401 – 418, 2020.

[14] U. Bodkhe, P. Bhattacharya, S. Tanwar, S. Tyagi, N. Kumar, and M. S. Obaidat, “Blohost: Blockchain enabled smart tourism and hospitality management,” in 2019 International Conference on Computer, Information and Telecommunication Systems (CITS), China, pp. 1–5, Aug 2019.

[15] A. Srivastava, Bhattacharya, A. Singh, A. Mathur, O. Prakash, and R. Pradhan, “A distributed credit transfer educational framework based on blockchain,” in 2018 Second International Conference on Advances in Computing, Control and Communication Technology (IAC3T), Allahabad, India, pp. 54–59, IEEE, 2018.

[16] Y. Liu, A. Ghandar, and G. Theodoropoulos, “A metaheuristic strategy for feature selection problems: Application to credit risk evaluation in emerging markets,” in 2019 IEEE Conference on Computational Intelligence for Financial Engineering Economics (CIFEr), Shenzhen, China, pp. 1–7, May 2019.

[17] S. Guo, H. He, and X. Huang, “A multi-stage self- adaptive classifier ensemble model with application in credit scoring,” IEEE Access, vol. 7, pp. 78549–78559, 2019.

[18] B. Zhu, W. Yang, H. Wang, and Y. Yuan, “A hybrid deep learning model for consumer credit

scoring," in 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), Las Vegas, USA, pp. 205– 208, May 2018.

[19] J. Srinivas, A. K. Das, N. Kumar, and J. Rodrigues, "Cloud centric authentication for wearable healthcare monitoring system," IEEE Transactions on Dependable and Secure Computing, 2018.

[20] M. D. Angelo and G. Salzer, "A survey of tools for analyzing ethereum smart contracts," 2019 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPCON), pp. 69–78, 2019.

[21] "Mythril security analysis framework on evm bytecodes to detect security vunerabilities of smart contracts." https://github.com/ConsenSys/ mythril. Accessed: 2020-01-26.

[22] D. Dua and C. Graff, "UCI machine learning repository," 2017.

[23] V. Odelu, A. K. Das, M. Wazid, and M. Conti, "Provably secure authenticated key agreement scheme for smart grid," IEEE Transactions on Smart Grid, vol. 9, no. 3, pp. 1900–1910, 2018.