Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 15, Issue. 1 : 2024
ISSN : **1906-9685**

PREDICTION OF FRAUD TRANSACTIONS IN CRYPTO SYSTEMS USING VOTING CLASSIFIER ENSEMBLE MODEL

#1 Mr.SK.UDDANDU SAHEB, #2 G.VIVEK, #3 T.RUPAS, #4 K.SIDDARDHA PRASAD, #5 A.PRANATHI

#1Assistant professor in Department of IT, DVR & Dr.HS MIC College of Technology,Kanchikacherla-521180

#2#3#4#5 B.Tech with Specialization of Information Technology , DVR & Dr.HS MIC College of Technology,Kanchikacherla-521180

Abstract_ In order to solve the issues with the centralised transaction system, cryptocurrency has arisen as a decentralised transaction. Instead of sending money physically, this strategy has increased the number of fraudulent transactions, despite becoming a popular trend in online crypto transactions and mobile wallets. due to the possibility of fraudulent transactions resulting from shared data and online transaction histories. Preprocess identification of fraudulent crypto transactions is turning into a pressing area of inquiry. With artificial intelligence growing at an exponential rate, many disciplines have succeeded in using machine learning to predict social challenges. From this perspective, this paper proposes an ensemble learning approach for fraudulent cryptocurrency transactions with the combination of machine learning algorithms: We use all machine learning algorithms (DT, RFC, and NB) for training, combining them to create an ensemble voting classifier model. The accuracy and losses from training and test datasets are then compared between the ensemble using the bagged and boosted methodology..

1.INTRODUCTION

Digital currency has arisen as a thrilling stage with the possibility to defeat issues related with the current methods of installments and exchanges . The gigantic expansion in the utilization of digital currency in the installment region has opened more open doors and difficulties as well as elaborate crimes. As per one gauge, 1,000 digital forms of money enter the market every month with various convenience . Besides, more than 2.4M digital currencies are accessible in 2024, and 70 of these cryptographic forms of money have a market cap of more than one billion bucks . Blockchains are utilized for the improvement of digital currencies that keep up with public records for overseeing cryptographic money exchanges . Digital money exchanges are decentralized and kept in a shared organization called a blockchain dispensing with the requirement for a focal power. Bitcoin is a spearheading digital currency, yet there are likewise numerous different coins with huge potential. For instance, Ethereum is the second-biggest cryptographic money by market capitalization that permits shrewd agreements . An outline of Ethereum's worth is introduced in Fig.

1. The Bitcoin network has a restricted ability to rapidly deal with numerous exchanges; accordingly, it isn't adaptable . Ethereum diminishes the issue of the versatility of Bitcoin . Vitalik Buterin creates Ethereum to assign capacity to the client . The principal benefit of Ethereum is its low exchange charge (gas) expected to execute an exchange on Ethereum, paying little heed to exchange achievement or disappointment.

Each gwei (Ethereum gas unit) is equivalent to 0.000000001 ETH (10−9 ETH). Ethereum is more versatile to shrewd agreements and exchanges. Shrewd agreements are a sort of Ethereum account. This implies they have an

equilibrium and can be the objective of exchanges. In this manner, fake exchanges might happen through savvy contracts. The decentralized blockchain approach permits activities without focal control and middle people with a few ben efits connected with protection and security , e.g., exchange namelessness . Such advantages make fake conduct extremely normal i.e., the decentralized control of blockchain and exchanges' namelessness prompts continuous deceitful exchange conduct in digital money . As per CipherTrace, a digital currency legal sciences organization's trick prompted a deficiency of 4.5 billion bucks in 2019. In addition, the digital money observing organizations pronounced that Ethereum is the principal decision for fake exchanges .

Albeit a client's secrecy is entirely reasonable for false exchanges with any Digital currency network the absence of control by a power and namelessness is exceptionally alluring for deceitful movement. The Ponzi plot is one of the most energetic tricks related with digital money. The disguising plans are normal on the Ethereum organization . In light of changelessness and client namelessness, deceitful exchanges are 95214 hard to switch, subsequently making them extremely alluring for fake exchanges. It is additionally extremely challenging and tedious to sort for fake exchanges physically. Such a colossal arrangement of exchanges makes recognizing deceitful exchanges almost incomprehensible. The issue is likewise hard in regards to time and different assets expected to distinguish strange exercises . AI is an optimal possibility for this reason. Numerous endeavors were placed into impact utilizing AI to identify irregular movement according to an alternate point of view . The reason for the proposed arrangements is to distinguish deceitful exchanges utilizing an AI model. The review means to distinguish deceitful exchanges on the Ethereum plat structure with restricted highlights. Note that we select the Ethereum stage since it is generally satisfactory and more versatile to brilliant agreements. AI approaches are generally applied to work on the exactness of distinguishing deceitful exchanges. The review needs to work on the exactness of recognizing fake exchanges in Ethereum utilizing casting a ballot classifier group model. The concentrate likewise intends to investigate the exhibition of various AI calculations to recognize the event of false exchanges utilizing casting a ballot classifier troupe model. The paper features are as per the following: • Distinguishing proof of the fake exchanges on the Ethereum network with high exactness. • Presentation of a troupe AI way to deal with work on the exactness of recognizable proof of deceitful exchanges on the Ethereum organization. • Top to bottom correlation of various AI models against the proposed troupe approach of distinguishing proof of fake exchanges

2.LITERATURE SURVEY

A machine learning based method for automated blockchain transaction signing including per sonalized anomaly detection

B. Podgorelec, M. Turkanović, and S. Karakatič,

The basis of blockchain-related data, stored in distributed ledgers, are digitally signed transactions. Data can be stored on the blockchain ledger only after a digital signing process is performed by a user with a blockchain-based digital identity. However, this process is time-consuming and not user-friendly, which is one of the reasons blockchain technology is not fully accepted. In this paper, they proposed a machine learning-based method, which introduces automated signing of blockchain transactions, while including also a personalized identification of anomalous transactions. In order to evaluate the proposed method, an experiment and analysis were performed on data from the Ethereum public main network. The analysis shows promising results and paves the road for a possible future integration of such a method in dedicated digital signing software for blockchain transactions

Fraudulent account recognition using supervised learn ing in Ethereum,

P. K. Choudhary,

Blockchain has gained significant popularity in the modern era. Almost all kinds of financial

transactions are supported by the Blockchain platform. Ethereum is one of the most used Blockchain platforms. After Bitcoin, Ethereum is the contributor to the second-largest cryptocurrency. The number of transactions performed on Ethereum in a day exceeds 1 million. The security and ease of transactions make it ideal for all kinds of transactions. But despite all the security features provided by Ethereum, there is a significant quantity of illegal activities that are conducted on Ethereum. These illegal activities significantly harm the spread and usage of Ethreum by people and organizations. Thus, there is a need for a mechanism to detect the illegal activities on Ethereum Blockchain. The category of illegal activities is huge and the scope of this work is limited to the detection of illicit accounts on Ethereum using machine learning techniques. A novel convolution neural network architecture followed by an XGBoost classifier is proposed to segregate the accounts as illicit or normal based on transaction history. The XGBoost model is a tree-based ensemble classifier. XGBoost classifier has been used to improve the accuracy of the proposed model without overfitting the model too much. The implicit regularization supported by XGBoost helps the model to generalize well for the dataset. XGBoost provides another benefit in that the individual trees are parallelly created while training. This decreases the time required for training the model and makes it more scalable. The historical transactions of over 4000 Ethereum accounts are used as a dataset to train the model and perform prediction. The dataset is a balanced dataset as the normal accounts and fraudulent accounts both are in nearly equal proportion. The accuracy achieved is 98.39 % and an average AUC which is better than standard machine learning models.

Temporal debiasing using adversarial loss based GNN architecture for crypto fraud detection

Singh, A. Gupta, H. Wadhwa, S. Asthana, and A. Arora

The tremendous rise of cryptocurrency in the payment domain has unlocked huge opportunities but also raised numerous challenges in parallel involving cybercriminal activities like money laundering, terrorist financing, illegal and risky services, etc, owing to its anonymous and decentralized setup. The demand for building a more transparent cryptocurrency network, resilient to such activities, has risen extensively as more financial institutions look to incorporate it into their network. While a plethora of traditional machine learning and graph based deep learning techniques have been developed to detect illicit activities in a cryptocurrency transaction network, the challenge of generalization and robust model performance on future timesteps still exists. In this paper, they showed that the model learned on transactional feature set provided in dataset (Elliptic Dataset) carry a temporal bias, i.e. they are highly dependent on the timesteps they occur. Deploying temporally biased models limits their performance on future timesteps. To address this, they proposed a temporal debiasing technique using GNN based architecture that ensures generalization by adversarially learning between fraud 1 classification and temporal classification. The adversarial loss constructed optimizes the embeddings to ensure they 1.) perform well on fraud classification task 2.) does not contain temporal bias. The proposed architecture capture the underlying fraud patterns that remain consistent over time. They evaluated the performance of the proposed architecture on the Elliptic dataset and compare the performance with existing machine learning and graph-based architectures. 1 Fraud and illicit are used interchangeably in this paper. Firstly, the data is per-processed and then CNN extracts the features from the image.

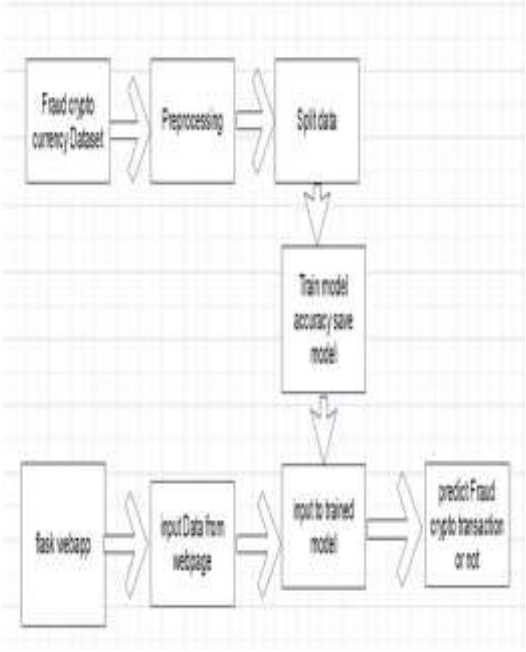

Fig 1:Architecture

3.PROPOSED SYSTEM

• The suggested solutions use a machine learning algorithm to identify fraudulent transactions. The study attempts to identify fraudulent transactions on the feature-limited Ethereum platform. Keep in mind that the Ethereum platform was chosen because it is more generally accepted and better suited for smart contracts. Machine learning techniques are frequently used to increase the precision of fraudulent transaction identification. We employ the ensemble model of voting classifier (RFC, NB, DT) among these methods. The goal of the project is to use ensemble machine learning techniques to increase the accuracy of fraudulent transaction detection in Ethereum. The project also intends to investigate how well various ensemble models perform when employing ensemble voting classifier learning algorithms to identify fraudulent transactions. The following are highlights of the paper:

• Accurate identification of fraudulent transactions on the Ethereum network

• Introducing an ensemble machine learning technique to increase the precision with which fraudulent transactions on the Ethereum network are identified.

• A thorough analysis contrasting various machine learning models with the suggested ensemble method for identifying fraudulent transactions.

3.1 IMPLEMENTATION

DATASET:

• In this project se are using Ethereum fraud detection dataset which is collected from Kaggle website. 7 features of transactions are considered in this dataset and label is used as fraud or not.

PRE-PROCESSING:

• • Features are extracted from data set and stored in variable as xtrain variable and labels are stored in y train variable. Data is preprocessing by standard scalar function and new features and labels are generated.

METHODOLOGY:

As seen from the above figure, we can see how the data is divided into different sets and then trained for different models. • The dataset was first divided into training set (80%) and pre-training set(20%). • The pre-training set was divided into pre-train(80%) and pre- test(20%) • Now, the training set is further is divided into train(80%) and validation set(20%). This train set is again divided into train(80%) and test set(20%). So, now I have train validation and test sets separate which are nonoverlapping. • The pretrain set was used to find the best models for the given dataset. I took best 4 models using pretest set. Their performance was compared based on their mean absolute errors. • Once the best 4 models were obtained, hyperparameters for these models were tuned and the best parameter was selected.

Testing training:

• In this stage data is sent to testing and training function and divided in to four parts x test train, and y test train. Train variables are used for passing to algorithm where as test are used for calculating accuracy of the algorithm.

Initializing Multiple Algorithms and training with Voting classifier Ensemble Model:

• In this stage machine learning algorithms are initialized and train values are given to algorithm by this information algorithm will know what are features and what are labels. Then data is modeled and stored as pickle file in the system which can be used for prediction.

• Data set is trained with multiple algorithms and accuracy of each model is calculated and best model is used for prediction

Predict data:

• In this stage new data is taken as input and trained models are loaded using pickle and then values are preprocessed and passed to predict function to find out result which is showed on web application.

4.RESULTS AND DISCUSSION

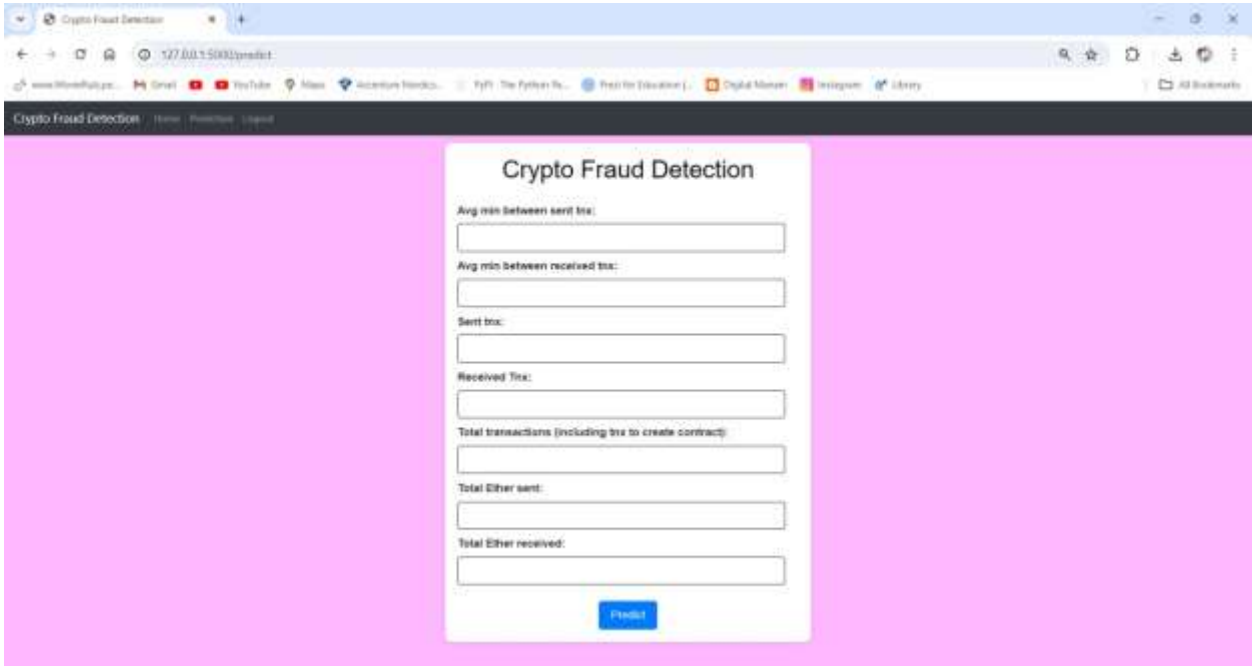

Fig 2:FEATURES

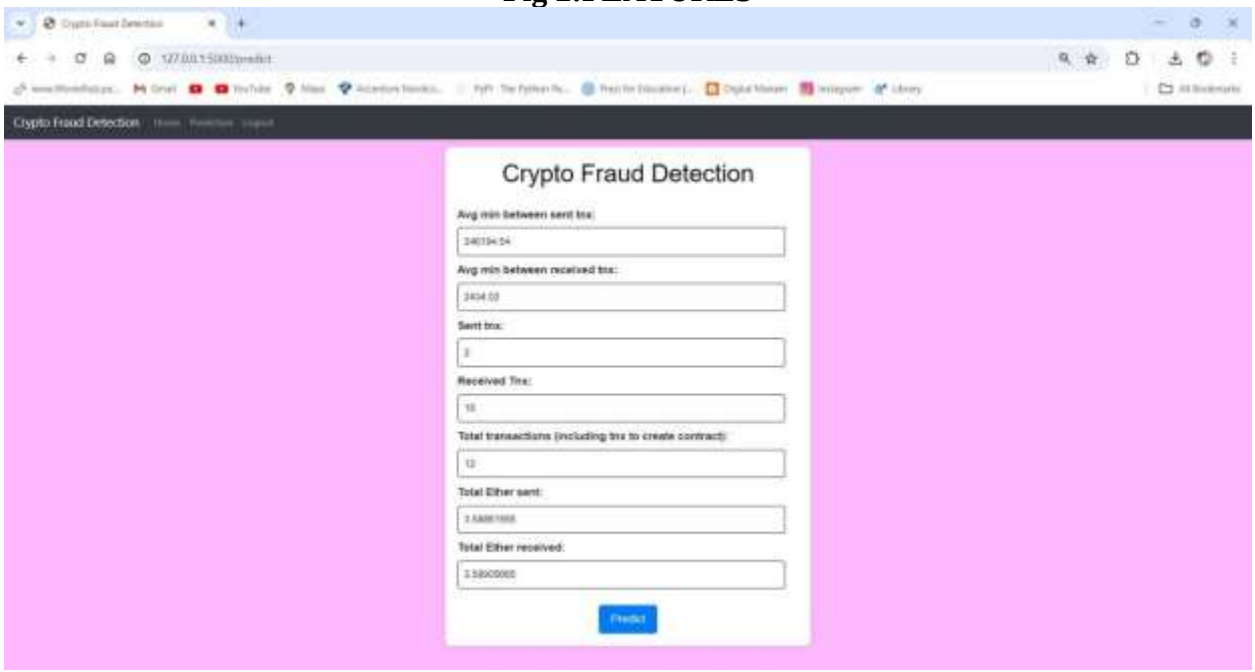

FIG 3:INPUT-1

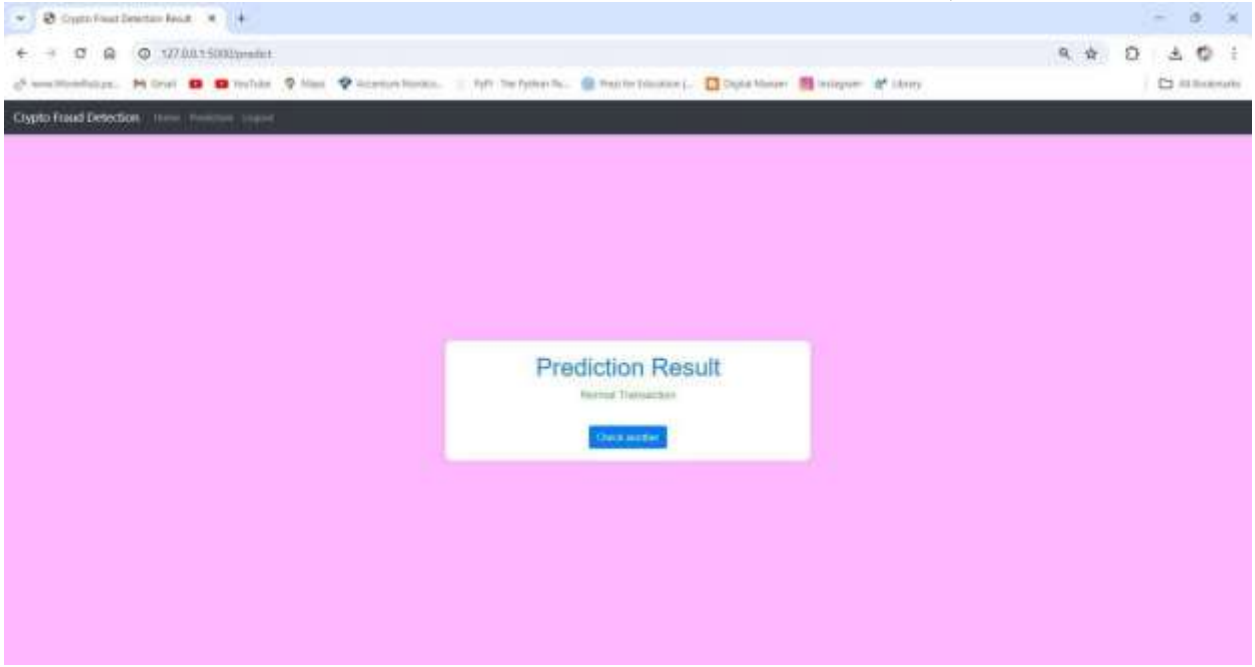

FIG 4:OUTPUT-1

FIG 5:INPUT-2

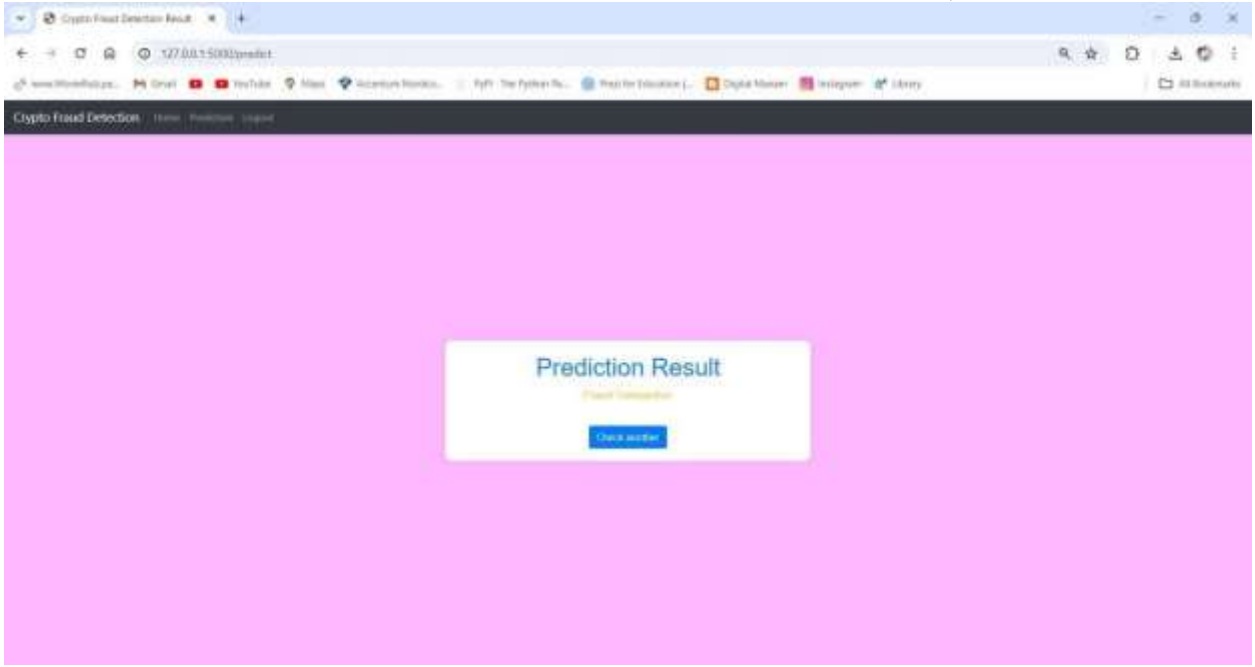

FIG 6:0UTPUT-2

5.CONCLUSION

Accomplishing a 92 percent precision in extortion identification utilizing AI (ML) models in digital currency exchanges is a remarkable achievement. Such a high precision rate shows that the ML model is capable at recognizing genuine and deceitful exchanges inside the dataset it was prepared and tried on. Notwithstanding, it's fundamental to consider a few variables while reaching determinations from this outcome:

Dataset Quality: The exactness of ML models intensely depends on the quality and representativeness of the dataset utilized for preparing and testing. On the off chance that the dataset is one-sided or doesn't sufficiently catch the variety of deceitful exercises, the model's presentation may not sum up well to true situations.

Assessment Measurements: Precision alone probably won't give a thorough image of the model's exhibition, particularly in imbalanced datasets where fake exchanges are uncommon contrasted with genuine ones. It's vital to consider different measurements, for example, accuracy, review, F1 score, and region under the beneficiary working trademark (ROC) bend to evaluate the model's viability in distinguishing extortion while limiting misleading up-sides.

Generalization: The revealed exactness relates to the presentation of the model on the dataset it was assessed on. It's fundamental to approve the model's exhibition on inconspicuous information to guarantee its capacity to sum up to new and concealed fake examples.

Bogus Up-sides and Misleading Negatives: While a 92 percent exactness rate sounds great, examining the model's bogus positive and misleading negative rates is fundamental. Misleading up-sides (genuine exchanges delegated fake) and bogus negatives (deceitful exchanges named real) can have various ramifications relying upon the setting of misrepresentation recognition.

Heartiness and Flexibility: ML models ought to be powerful to changes in the fundamental information conveyance and versatile to arising misrepresentation designs. Normal observing and retraining of the model are important to keep up with its adequacy over the long haul.

Reconciliation and Arrangement: Contemplations for coordinating the ML model into the functional work process of cryptographic money stages or monetary foundations, including inertness prerequisites, adaptability, and consistence with administrative principles.

All in all, accomplishing a 92 percent precision rate in extortion discovery utilizing ML models is excellent, yet it's fundamental to decipher this outcome inside the setting of the elements referenced previously. Consistent assessment, refinement, and approval are pivotal for

guaranteeing the model's viability in true situations.

REFERENCES

[1] M. J. Shayegan and H. R. Sabor, ‘‘A collective anomaly detection method over Bitcoin network,’’ 2021, *arXiv:2107.00925*.

[2] B. Podgorelec, M. Turkanović, and S. Karakatič, ‘‘A machine learning- based method for automated blockchain transaction signing including per- sonalized anomaly detection,’’ *Sensors*, vol. 20, no. 1, p. 147, Dec. 2019.

[3] P. K. Choudhary, ‘‘Fraudulent account recognition using supervised learn- ing in Ethereum,’’ Ph.D. dissertation, Indian Institute Technol. Jodhpur, India, 2021.

[4] A. Singh, A. Gupta, H. Wadhwa, S. Asthana, and A. Arora, ‘‘Temporal debiasing using adversarial loss based GNN architecture for crypto fraud detection,’’ in *Proc. 20th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2021, pp. 391–396.

[5] C. Prices, ‘‘Charts and market capitalizations| coinmarketcap. (SEM data),’’ CoinMarketCap, Online, Tech. Rep., 2022.

[6] K. Lašas, G. Kasputyté, R. Užupyté, and T. Krilavičius, ‘‘Fraudulent behaviour identification in Ethereum blockchain,’’ in *Proc. CEUR Work- shop, Inf. Soc. Univ. Stud.*, Kaunas, Lithuania, 23, Apr. 2020, pp. 1–8.

[7] J. T. Hamrick, F. Rouhi, A. Mukherjee, A. Feder, N. Gandal, T. Moore, and M. Vasek, ‘‘An examination of the cryptocurrency pump-and- dump ecosystem,’’ *Inf. Process. Manag.*, vol. 58, no. 4, Jul. 2021, Art. no. 102506.

[8] U. W. Chohan, ‘‘Are cryptocurrencies truly trustless?’’ in *Cryptofi- nance and Mechanisms of Exchange*. Berlin, Germany: Springer, 2019, pp. 77–89.

[9] H. Nghiem, G. Muric, F. Morstatter, and E. Ferrara, ‘‘Detecting cryptocur- rency pumpand-dump frauds using market and social signals,’’ *Exp. Syst. Appl.*, vol. 182, Nov. 2021, Art. no. 115284.

[10] A. Chauhan, O. P. Malviya, M. Verma, and T. S. Mor, ‘‘Blockchain and scalability,’’ in *Proc. IEEE Int. Conf. Softw. Qual., Rel. Secur. Companion (QRS-C)*, Jul. 2018, pp. 122– 128.

[11] A. Trozze, J. Kamps, E. A. Akartuna, F. J. Hetzel, B. Kleinberg, T. Davies, and S. D. Johnson, ‘‘Cryptocurrencies and future financial crime,’’ *Crime Sci.*, vol. 11, no. 1, pp. 1–35, Dec. 2022.

[12] F. Leal, A. E. Chis, and H. González–Vélez, ‘‘Multi-service model for blockchain networks,’’ *Inf. Process. Manag.*, vol. 58, no. 3, May 2021, Art. no. 102525.

[13] *Top 10 Cryptocurrencies Price Analysis | Cointelegraph*. Accessed: Jul. 22, 2022. [Online]. Available: https://cointelegraph.com/category/top- 10-cryptocurrencies

[14] J. Barna, ‘‘Blockchain and cryptocurrencies,’’ Harvard Model Congr., Boston, MA, USA, Tech. Rep., 2022.

.

Author’s Profiles

#1:-Mr.SK.UDDANDU SAHEB working as Assistant Professor in Department of IT in DVR & Dr.HS MIC College of Technology,Kanchikacherla-521180

#2:-G.VIVEK(20H71A1262) B.Tech with Specialization of Information Technology in DVR & Dr.HS MIC College of Technology,Kanchikacherla-521180

#3:- T.RUPAS(20H71A1231) B.Tech with Specialization of Information Technology in DVR & Dr.HS MIC College of Technology,Kanchikacherla-521180

#4:-K.SIDDARDHA PRASAD(20H71A1240) B.Tech with Specialization of Information Technology in DVR & Dr.HS MIC College of Technology,Kanchikacherla-521180

#5:-A.PRANATHI(20H71A1225) B.Tech with Specialization of Information Technology in DVR & Dr.HS MIC College of Technology,Kanchikacherla-521180