IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, June 2022
Impact Factor: 6.252

Comparative Study of Machine Learning Algorithms for Fraud Detection in Blockchain

Rohan Kumar C L[1], Ali Mohammed Zain[2], Sanjay Kumar H P[3], Prajwal A V[4], Dr. Sudarshan R[5]
Students, Department of Information Science and Engineering[1,2,3,4]
Professor & Head, Department of Information Science and Engineering[5]
Vidya Vikas Institute of Engineering and Technology, Mysuru, Karnataka, India

Abstract: *Fraudulent transactions have a huge impact on the economy and trust of a block chain network. Consensus algorithms like proof of work or proof of stake can verify the validity of the transaction but not the nature of the users involved in the transactions or those who verify the transactions. This makes a block chain network still vulnerable to fraudulent activities. One of the ways to eliminate fraud is by using machine learning techniques. Machine learning can be of supervised or unsupervised nature. In this paper, we use various supervised machine learning techniques to check for fraudulent and legitimate transactions. We also provide an extensive comparative study of various supervised machine learning techniques.*

Keywords: Fraud, Comparative Study, Blockchain, Machine learning.

I. INTRODUCTION

The problem of detecting fraudulent transactions is being studied for a long time. Fraudulent transactions are harmful to the economy and discourage people from investing in bit coins or even trusting other block chain- based solutions. Fraudulent transactions are usually suspicious either in terms of participants involved in the transaction or the nature of the transaction. Members of a block chain network want to detect fraudulent transactions as soon as possible to prevent them from harming the block chain network's community and integrity. Many Machine Learning techniques have been proposed to deal with this problem, some results appear to be quite promising, but there is no obvious superior method. This paper compares the performance of various supervised machine learning models and few deep learning models in detecting fraudulent transactions in a block chain network. Such comparative study will help decide the best algorithm based on accuracy and computational speed trade-off. Our goal is to see which users and transactions have the highest probability of being involved in fraudulent transactions.

II. EXISTING SYSTEM

The problem of detecting fraudulent transactions is being studied for a long time. Fraudulent transactions are harmful to the economy and discourage people from investing in bitcoins or even trusting other blockchain- based solutions. Fraudulent transactions are usually suspicious either in terms of participants involved in the transaction or the nature of the transaction. Members of a blockchain network want to detect Fraudulent transactions as soon as possible to prevent them from harming the blockchain network's community and integrity. Many Machine Learning techniques have been proposed to deal with this problem, some results appear to be quite promising, but there is no obvious superior method. This paper compares the performance of various supervised machine learning models like SVM, Decision Tree, Naive Bayes, Logistic Regression, and few deep learning models in detecting fraudulent transactions in a blockchain network. Such comparative study will help decide the best algorithm based on accuracy and computational speed trade-off. Our goal is to see which users and transactions have the highest probability of being involved in fraudulent transactions.

III. PROPOSED SYSTEM

As we discussed the objective and subjective frauds. They conclude that blockchain effectively detects objective fraud but not subjective fraud and thus uses Machine Learning to mitigate the weakness. worked on illegal applications for tax returns and caused huge loss to the government. They mainly focus on international transactions, including double tax at the country where the transaction is generated and the country where the transaction is received, insight into a different type of fraudulent activity related to tax returns. We discussed the types of fraudulent activities that blockchain

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, June 2022
Impact Factor: 6.252

can detect and the ones that blockchain is still vulnerable to. This paved a path towards ideas about what problems a Machine learning part needs to consider. She specifies that attacks like Identity theft and system hacking are still possible and challenging to detect using blockchain as it just uses some predetermined rules. We used Supervised Machine Learning methods to detect fraudulent activities. They focused on the fact that malicious actors can steal money by applying well-known malware software or fake emails. Therefore, they used the capabilities of Random Forests, Support Vector Machines, and XGBoost classifiers to identify such accounts based on a dataset of more than 300 thousand accounts.

3.1 Objectives

This paper compares the performance of various supervised machine learning models like SVM, Decision Tree, Naive Bayes, Logistic Regression, and few deep learning models in detecting fraudulent transactions in a blockchain network. Such comparative study will help decide the best algorithm based on accuracy and computational speed trade-off. Our goal is to see which users and transactions have the highest probability of being involved in fraudulent transactions.

IV. LITERATURE SURVEY

Cai, Y., Zhu, D. Fraud detections for online businesses: a perspective from blockchain technology.
Finance Innov 2, 20 (2016).
Yuanfeng Cai et al. discussed the objective and subjective frauds. They conclude that blockchain effectively detects objective fraud but not subjective fraud and thus uses Machine Learning to mitigate the weakness.
Method: Blockchain

Xu, J.J. Are blockchains immune to all malicious attacks? Finance Innov 2, 25 (2016).
Jennifer J. Xu discussed the types of fraudulent activities that blockchain can detect and the ones that blockchain is still vulnerable to. This paved a path towards ideas about what problems a Machine learning part needs to consider. She specifies that attacks like Identity theft and system hacking are still possible and challenging to detect using blockchain as it just uses some predetermined rules.
Method: Blockchain

Ostapowicz M., ˙ Zbikowski K. (2019) Detecting Fraudulent Accounts on Blockchain: A Supervised Approach. In: Cheng R., Mamoulis N., Sun Y., Huang X. (eds) Web Information Systems Engineering
Michał Ostapowicz et al. used Supervised Machine Learning methods to detect fraudulent activities. They focused on the fact that malicious actors can steal money by applying well-known malware software or fake emails. Therefore, they used the capabilities of Random Forests, Support Vector Machines, and XGBoost classifiers to identify such accounts based on a dataset of more than 300 thousand accounts.
Method: Supervised Machine Learning

Podgorelec, B., Turkanovi´c, M. and Karakati˘c, S., 2020. A Machine Learning-Based Method for Automated Blockchain Transaction Signing Including Personalized Anomaly Detection.
Bla˘z Podgorelec et al. devised a method using Machine Learning for the automated signing of transactions in the blockchain. Hence, it also uses a personalized identification of anomalous transactions.
Method: Machine Learning

Farrugia S, Ellul J, Azzopardi G. Detection of illicit accounts over the Ethereum blockchain. Expert Systems with Applications
Steven Farrugia et al. detected illicit accounts in the Ethereum Blockchain based on their transaction history. They found out that ‘Time difference between first and last (Mins)’, ‘Total Ether balance’ and ‘Min value received’ are the three major contributing factors for detecting illicit accounts.
Method: Blockchain.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, June 2022
Impact Factor: 6.252

V. METHODOLOGY

5.1 Machine Learning

Machine learning is a process of studying a system based on data. Machine learning is a part of data science where we use machine learning algorithms to process data.

5.2 Supervised Learning

This involves learning from a training dataset with labelled data using classification and regression models. This learning process continues until the required level of performance is achieved.

This model compares eight different Machine Learning Algorithms:

1. **Logistic Regression:** This is a simple linear classifier. Logistic regression works well for binary classification problems
2. **Multilayer Perceptron**: Multilayer perceptron helps in separation data that cannot be classified using a linear classifier by introducing non linearity.
3. **Naive Bayes:** This model uses the Bayes theorem to calculate the probability of a transaction being fraudulent.
4. **Adaboost:** This is an ensemble learning method to boost the performance of binary classifiers.
5. Decision Tree: This classifier has a sequence of conditions and questions on data based on various features.
6. **SVM:** It uses a kernel method to transform the data in the dataset, and based on these transitions, it finds a boundary between all possible outputs.
7. **Random Forest Classifier:** This classifier fits a number of decision trees on small batches of the dataset.
8. **Neural Network:** This model consists of six dense layers and four hidden layers. Relu and sigmoid were used as activation functions.

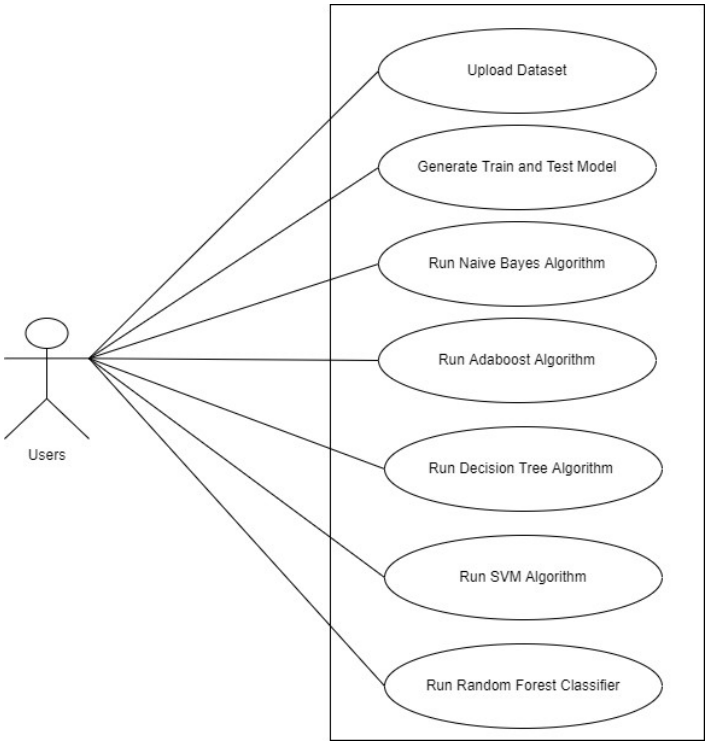

Fig. 1 Use Case Diagram

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, June 2022
Impact Factor: 6.252

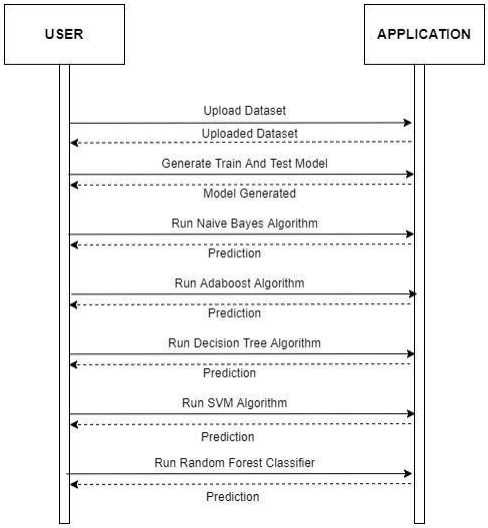

Fig. 2 SEQUENCE DIAGRAM

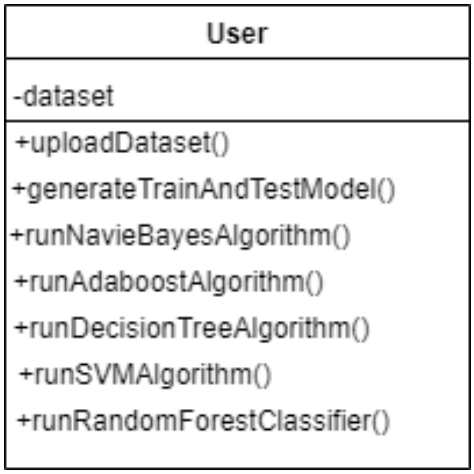

Fig. 3 CLASS DIAGRAM

VI. IMPLEMENTATION

After the Block chain network has approved a transaction after all basic checks, our proposed system kicks in and does additional checks to detect if the transaction can be fraudulent.

The work done can be divided mainly into three phases:

1. Pre-processing phase
2. Building and training various models

Performance evaluation of all the models

VII. CONCLUSION AND FUTURE ENHANCEMENT

A method has been proposed for the detection of fraudulent transactions in a blockchain network using machine learning. In this method, various supervised learning approaches like support vector machines, decision trees, logistic regression, were analyzed. A thorough comparative analysis of all the approaches is performed through accuracy. This work can be extended for the comparative study of unsupervised algorithms like clustering. In the future, we also plan to do an exhaustive study on fraudulent activities in a private blockchain.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, June 2022
Impact Factor: 6.252

REFERENCES

[1]. Cai, Y., Zhu, D. Fraud detections for online businesses: a perspective from blockchain technology. Financ Innov 2, 20 (2016).

[2]. Hyv¨arinen, H., Risius, M. & Friis, G. A Blockchain-Based Approach Towards Overcoming Financial Fraud in Public Sector Services. Bus Inf Syst Eng 59, 441–456 (2017). https://doi.org/10.1007/s12599-017- 0502- 4

[3]. Xu, J.J. Are blockchains immune to all malicious attacks? Finance Innov 2, 25 (2016). https://doi.org/10.1186/s40854-016-0046-5

[4]. Ostapowicz M., ˙ Zbikowski K. (2019) Detecting Fraudulent Accounts on Blockchain: A Supervised Approach. In: Cheng R., Mamoulis N., Sun Y., Huang X. (eds) Web Information Systems Engineering – WISE 2019. WISE 2020. Lecture Notes in Computer Science, vol 11881. Springer, Cham. https://doi.org/10.1007/978-3-030-34223-4 2

[5]. Podgorelec, B., Turkanovi´c, M. and Karakati˘c, S., 2020. A Machine Learning-Based Method for Automated Blockchain Transaction Signing Including Personalized Anomaly Detection. Sensors, 20(1), p.147.

[6]. Farrugia S, Ellul J, Azzopardi G. Detection of illicit accounts over the Ethereum blockchain. Expert Systems with Applications. 2020 Jul 15;150:113318.

[7]. Pham, Thai, and Steven Lee. "Anomaly detection in bitcoin network using unsupervised learning methods." arXiv preprint arXiv:1611.03941 (2016).

[8]. Monamo, Patrick, Vukosi Marivate, and Bheki Twala. "Unsupervised learning for robust Bitcoin fraud detection." 2016 Information Security for South Africa (ISSA). IEEE, 2016.