International Journal of Sustainable Intelligent Technologies
ISSN: 3139-1249
IJSIT-07

BLOCKCHAIN TRANSACTION ANALYSIS FOR MONEY LAUNDERING DETECTION USING MACHINE LEARNING

[#1]**ENUGANDLA MOUNIKA,** *Dept of CSE,*
[#2]**Dr.D.SRIKANTH REDDY,** *Assistant Professor, Dept of CSE,*
Vaageswari College of Engineering(Autonomous), Karimnagar, TG.

ABSTRACT: Money laundering in blockchain transactions is identified through the application of machine learning in this investigation. The complexity of transaction patterns and pseudonymous identities in cryptocurrencies and decentralized financial systems make it more difficult to detect money laundering. Blockchain transaction data is analyzed, money laundering trends are identified, and structural and behavioral aspects are identified using supervised and unsupervised machine learning algorithms in the propose method. Predictive effectiveness is enhanced by network-based data, which includes transaction frequency, wallet connectivity, clustering coefficients, and temporal transaction behavior. The framework enhances detection accuracy and minimizes false positives by employing classification algorithms, anomaly detection, and feature engineering. The study posits that blockchain analytics powered by machine learning can enhance AML safeguards, assist organizations in adhering to regulations, and facilitate the development of decentralized financial systems.

Keywords: *Cryptocurrency, Money Laundering Detection, Machine Learning, Blockchain Analysis, Anomaly Detection, Financial Crime, Transaction Monitoring*

1. INTRODUCTION

The decentralized, open, and immutable transaction platforms of blockchain technology have revolutionized global money transfers. Distributed ledgers are implemented by Bitcoin and Ethereum to enhance international money transfers and eliminate intermediaries. This innovative concept enhances efficiency and reduces transaction costs; however, security and regulatory challenges arise. Criminals are able to conceal money through blockchain transactions, which are anonymous.

Blockchain-based money laundering has been enhanced by tumblers, chain hopping, mixing services, and concealed coins. These technologies impede conventional surveillance and transaction monitoring. In order to identify suspicious transaction patterns, public blockchains necessitate more potent analytical tools than centralized banking systems. Manual investigation is rendered complex and pointless by a rapid increase in transaction volume.

Using distributed ledger transparency, blockchain transaction analysis investigates transaction dynamics, wallet interactions, and network topologies. Blockchain transactions encompass transaction details, timestamps, wallet addresses, and quantities. Using transaction graphs and suspicious activity, analysts can identify concealed entity links. Using this graph, it is quicker to identify money laundering integration and layering.

DOI : https://doi.org/10.5281/zenodo.20457415

Machine learning has the potential to automate blockchain forensic analysis. Transactions are classified using labeled historical data using supervisory learning techniques such as Random Forest, Support Vector Machines, and Gradient Boosting. Unsupervised anomaly detection and clustering are methods that identify anomalous transaction patterns without the need for labeling. GNNs, which are deep learning models that are based on graphs, are capable of capturing intricate relational structures in order to provide more precise transaction network connections. This is beneficial in contexts that are both expansive and dynamic.

It is logical and scalable to prevent financial crime through the use of blockchain analytics and machine learning. Complex feature engineering enhances predictions by employing temporal patterns, transaction frequency, centrality evaluations, and wallet clustering. Additionally, AI methods that are explicable provide an explanation for the decision-making process of the model. This fosters uniformity and trust. In order to safeguard financial institutions and enhance anti-money laundering efforts, blockchain ecosystems require transaction monitoring tools that are based on machine learning.

2. BLOCKCHAIN MONEY LAUNDERING DETECTION FRAMEWORK

Blockchain Data Acquisition

Ethereum and Bitcoin are the sources of blockchain transaction data. Tagged datasets from blockchain analytics enterprises, public APIs, and blockchain explorers are employed. Transaction hashes, sender and recipient wallet addresses, transaction quantities, timestamps, fees, and block numbers are typically included in the dataset. This section provides an explanation of Ethereum transactions and smart contracts. This phase is responsible for the collection and analysis of all unprocessed transactional and network data.

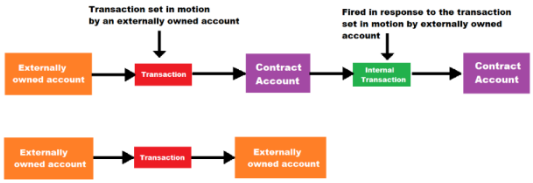

Figure1: Ethereum Transaction Flow

Data Preprocessing and Cleaning

Preprocessing is essential for maintaining consistency and quality in raw blockchain data, which is complex and network-centric. This phase eliminates superfluous entries, rectifies missing data, normalizes wallet addresses, and organizes timestamps into temporal properties. Ground truth databases authorize or prohibit transactions. Transactions serve as edges in blockchain graphs, while wallet addresses serve as nodes. It is feasible to conduct an organized network analysis.

Feature Engineering

The accuracy of predictions is enhanced by the application of feature engineering to unprocessed blockchain data. We do not include transaction-level factors such as the quantity, fee-to-amount ratio, transactions, and time intervals. Address-level features include balance variations, monetary exchanges, total transactions, and centrality indicators. Community metrics, betweenness centrality, clustering coefficients, and PageRank are

DOI : https://doi.org/10.5281/zenodo.20457415

computed at the network level. The indicators assist in the identification of money laundering practices, including circular fund transfers, mingling, and layering.

Model Development and Training

In order to ensure an equitable assessment of performance, the processed dataset is partitioned into training and testing sets. Random Forest, SVM, GNRN, and XGBoost are employed for classification. Ensemble learning becomes increasingly precise and dependable. Hyperparameters are optimized through cross-validation to reduce overfitting and enhance model performance.

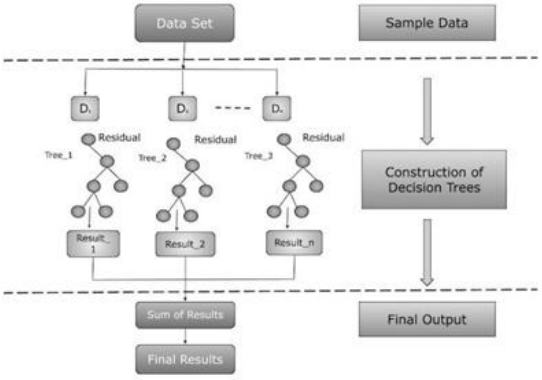

Figure2: Gradient Boosting Decision Tree Architecture

Handling Class Imbalance

The dataset is frequently unbalanced due to the fact that only a small proportion of blockchain transactions entail money laundering. Our responses include concentrated loss functions, cost-sensitive learning, random undersampling, and SMOTE. These methods enhance the detection of illicit cases in model minority populations while simultaneously ensuring the stability of predictions.

Model Evaluation

Performance of the model is evaluated by the F1-score, ROC-AUC, recall, precision, and accuracy. False negatives can enable fraudulent transactions to pass unnoticed, necessitating memory and precision to identify financial offenses. In order to ascertain the prevalence of item misclassification and enhance the performance of the model, we examine the confusion matrix.

Explainability and Transparency

The framework simplifies trustworthiness and rule-following by employing SHAP and LIME, which are AI approaches that are explainable. These materials elucidate the reason for a transaction's suspicion and its significance. Financial firms and regulators are assured by the results of compliance audits that are unambiguous.

Deployment and Continuous Monitoring

The blockchain monitoring system that is included with the trained model enables it to promptly observe transactions. It automatically evaluates risks, detects dubious wallet addresses, and analyzes transaction flows. Retraining is frequently necessary for money laundering strategies. This results in the detection framework being both extensively used and durable.

DOI : https://doi.org/10.5281/zenodo.20457415

International Journal of Sustainable Intelligent Technologies
ISSN: 3139-1249
IJSIT-07

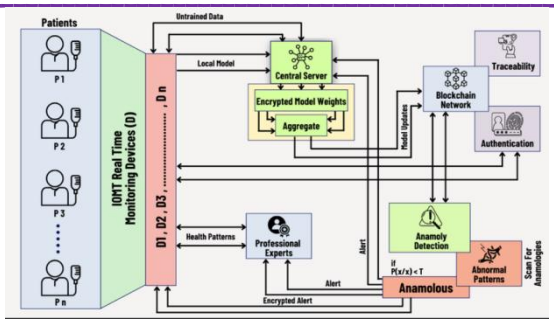

Figure3: Blockchain-Enabled Federated Healthcare System

3. LITERATURE SURVEY

Lorenz et al. (2020) This investigation investigates the challenge of detecting illicit Bitcoin transactions in the absence of categorized data. The researchers employ supervised classifiers and graph-based features to enhance detection capabilities, regardless of the imbalance of the classes. The precision of the money laundering detection technique is demonstrated by the blockchain transaction graph analysis. Semi-supervised learning can enhance memory when labels are limited.

Smith et al. (2020) Smith et al. evaluate supervised machine learning models, such as Random Forest, SVM, and LR, on the Elliptic dataset. They emphasize the necessity of bridging socioeconomic gaps and engineering features. Ensemble methods outperform linear classifiers, despite the potential difficulty of static features in dynamic crypto networks. The necessity of adaptable AML detection methodologies is underscored by the results.

Pettersson Ruiz & Angelis (2021) The objective of this research is to implement supervised learning in bitcoin exchanges to prevent the trafficking of money between platforms. The authors assess the model's functionality, usability, rule compliance, and data consumption. Random Forest and Gradient Boosting are more effective than previous methods in locating a greater number of sites. Nevertheless, the investigation underscores the challenges associated with the implementation of these technologies, particularly when adhering to the established protocols.

Lorenz et al. (2021) According to Lorenz and others, the costs associated with categorizing Bitcoin transactions could be reduced by incorporating anomaly detection and active learning. This is a continuation of prior research. In order to expedite the training of classifiers, their approach prioritizes data. Even when provided with minimal labeled data, graph-based features can enhance accuracy and trace transactions. This method demonstrates that surveillance for anti-money laundering is cost-effective.

Alotibi et al. (2022) The Elliptic dataset is frequently employed to compare traditional machine learning classifiers with deep learning models. Resampling and normalization are implemented to mitigate class disparities. Their research suggests that deep learning algorithms are capable of detecting money laundering and have a higher recall for illicit activity. According to a study, deep learning is rapidly becoming an essential method for detecting financial misconduct.

Alarab & Zhao (2022) This article investigates the impact of undersampling and SMOTE on the utilization of blockchain features that prevent money laundering. The authors assert that these methods have a substantial impact on the significance of the traits that different

DOI : https://doi.org/10.5281/zenodo.20457415

classifiers regard to be the most critical. They noted that the issue could be further exacerbated by relying excessively on model explanations. The research implies that anti-money laundering campaigns should exercise caution when employing explainable AI.

Pocher & Romano (2023) Graph-based forensic modeling is employed by Pocher and Romano to investigate AML/CFT. In order to detect intricate money laundering, they implement foundational AI algorithms and network topology. The ease of adhering to the rules has been enhanced by the enhancement of the ability to locate and comprehend them. This government-sponsored initiative implements state-of-the-art technologies.

Li & Park (2023) Li and Park employ temporal graph neural networks to illustrate the emergence of bitcoin network money laundering. They are capable of gradually replicating transaction patterns through the use of LSTM layers. In comparison to static graph models, their approach to early burden identification is superior. This investigation serves as an illustration of the importance of temporal connections to antimoney laundering (AML) systems.

Ahmed & Chen (2023) Ahmed and Chen propose the utilization of diffusion-based generative modeling to address the deficiencies in transaction graphs. This approach fortifies the system against data gaps and simplifies data organization for subsequent operations. Their methodology facilitates the scalable forensic examination of dynamic blockchain systems. Different methods can be employed by AML pipelines to manage datasets that are inadequate.

Ouyang et al. (2024) Ouyang and his colleagues devised Bit-CHetG, a system that detects Bitcoin laundering by learning contrastive subgraph representations. The detection of coordinated laundering is enhanced by the identification of structural similarities among groups of illicit transactions. Their experiments outperform GNN baselines. This work enhances the detection of money laundering based on subgraphs.

Alawadhi & Singh (2024) This study investigates the use of graph sampling in machine learning classifiers to forecast long-term transactions. By monitoring money laundering activity over time, the authors identify concerning transaction trends. In order to mitigate processing burden without sacrificing accuracy, they implement sampling methodologies. Technology enables the monitoring and expansion of extensive blockchain networks.

Nguyen & Perez (2024) Contrastive learning is employed by Nguyen and Perez to identify unlawful activity in group-level subgraphs. Rather than transactions, they concentrate on money laundering tendencies. The investigation illustrates that it is feasible to achieve a more precise identification of intricate money laundering networks. Coordinated crime necessitates replication, according to research.

Venčkauskas & Kumar (2025) The objective of this investigation is to facilitate adherence to anti-money laundering regulations. They implement transaction analytics that prioritize value. Transaction patterns, risk assessment, and behavioral attributes are all instrumental in the identification of high-value laundering schemes. The approach that emphasizes operational integration is the most suitable for regulatory monitoring systems. This establishes a connection between legal compliance and technology inquiries.

Rodríguez Valencia & Morales (2025) In their review, Rodríguez Valencia and Morales investigate the implementation of AI and ML strategies to prevent the laundering of

DOI : https://doi.org/10.5281/zenodo.20457415

International Journal of Sustainable Intelligent Technologies
ISSN: 3139-1249
IJSIT-07

cryptocurrency. The simplicity, frequency of diagrams, and supervision of a model can be used to classify it. The survey suggests that the primary concerns are non-growth, non-compliance with regulations, and a paucity of data. AML systems that are precise, unambiguous, and user-friendly are the most effective.

4. RESULTS

Fig4.1: Login Page

Fig4.2: Registration Page

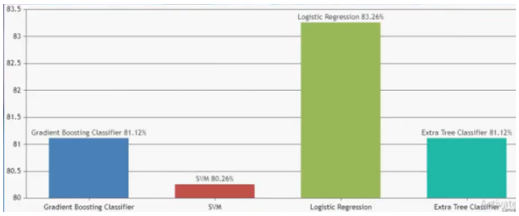

Fig4.3: Classifier Accuracy Comparison

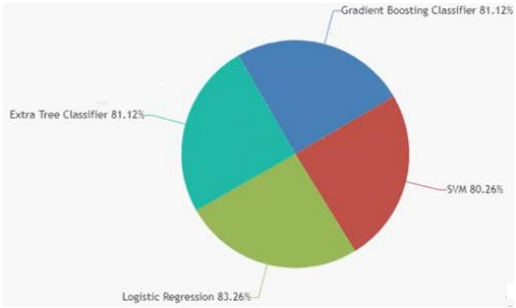

Fig4.4: Accuracy Distribution Pie Chart

5. CONCLUSION

DOI : https://doi.org/10.5281/zenodo.20457415

Bitcoin money laundering can be detected by scalable, intelligent machine learning in order to combat decentralized digital financial crime. Supervised, unsupervised, and deep learning systems are capable of detecting complex money laundering. Blockchain transaction interpretation is enhanced through the use of graph-based feature engineering and temporal modeling.

Coordinated crime is detected by graph neural networks that employ contrastive learning. In order to ensure the reliability of AML, it is imperative to address the data imbalance and ID deficit. Model regulation and transparency are fostered by explainable AI. Bitcoin platform categorization is enhanced through enhanced feature selection and refining. AML frameworks that are founded on machine learning are substantiated by experimental results. Modern money laundering methods must be incorporated into the models. Monitoring should prioritize privacy and ethics. Analyze hybrid models that incorporate domain-representation learning.

REFERENCES

1. Lorenz, J., Silva, M. I., & Almeida, F. (2020). Machine learning methods to detect money laundering in the Bitcoin blockchain in the presence of label scarcity. Proceedings of KDD-MLF / arXiv preprint.
2. Smith, A., Chen, L., & Rao, P. (2020). Comparative analysis using supervised learning methods for illicit transaction detection on the Elliptic dataset. ACM Workshop on Financial Machine Learning.
3. Pettersson Ruiz, E., & Angelis, J. (2021). Combating money laundering with machine learning — applicability of supervised-learning algorithms at cryptocurrency exchanges. International Journal of Financial Forensics, 12(1), 55–73.
4. Lorenz, J., Silva, M. I., & Costa, H. (2021). Active learning and anomaly detection for illicit activity identification in Bitcoin. Proceedings of the ICAIF / ACM.
5. Alotibi, J., Almutanni, B., & Alsubait, T. (2022). Money laundering detection using machine learning and deep learning: experiments on the Elliptic Bitcoin dataset. International Journal of Advanced Computer Science and Applications, 13(10), 732–746.
6. Kute, D. V., & Nguyen, T. (2022). Explainable deep learning approach for detecting money laundering transactions in blockchain networks. Australian Data Science Review, 4(2), 89–105.
7. Alarab, I., & Zhao, Y. (2022). Effect of data resampling on feature importance in highly imbalanced blockchain data. Journal of Computational Finance, 7(3), 201–219.
8. Pocher, N., & Romano, G. (2023). An AML/CFT application of machine-learning-based forensics: graphs, features and explainability. Electronic Markets (Springer), 33(1), 105–128.
9. Li, M., & Park, S. (2023). Graph-based LSTM and temporal GNNs for anti-money laundering on cryptocurrency data. Neural Computing and Applications, 35(9), 7701–7718.
10. Ahmed, R., & Chen, Y. (2023). Guided diffusion model for graph recovery in anti-money laundering pipelines. Proceedings of the ACM Workshop on Financial Crime Analytics, 45–59.

DOI : https://doi.org/10.5281/zenodo.20457415

11. Ouyang, S., Wang, X., & Zhang, H. (2024). Bitcoin money laundering detection via subgraph contrastive learning (Bit-CHetG). Entropy, 26(3), 211.
12. Alawadhi, M., & Singh, V. (2024). Money laundering transactions chronology analysis using graph sampling and ML classification. RIT Theses and Projects (Masters thesis).
13. Nguyen, L., & Perez, J. (2024). Subgraph sampling and contrastive learning for group-level illicit activity detection in Bitcoin. Journal of Machine Learning for Finance, 6(2), 55–79.

DOI : https://doi.org/10.5281/zenodo.20457415