Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/43
DOI: https://doi.org/10.5281/zenodo.20717743

MACHINE LEARNING-BASED MONEY LAUNDERING DETECTION IN BLOCKCHAIN TRANSACTIONS USING GRAPH ANALYTICS

[#1]**Mr. VV Siva Prasad,** *Assistant Professor,Department of CSE,*
[#2]**Mr. B. Santhosh Kumar ,** *Assistant Professor,Department of CSE,*
[#3]**Mrs. M.Thirupathammak,** *Assistant Professor,Department of CSE,*
Sai Spurthi Institute of Technology(Autonomous), Sathupalli, Khammam, Telangana

ABSTRACT: The rapid growth of cryptocurrencies has enabled the use of decentralized networks for complex money laundering activities and the creation of new financial services. In general, the functionality of conventional rule-based monitoring systems is surpassed by contemporary money laundering techniques on blockchain networks. In order to identify anomalous or dubious patterns for widespread application, machine learning can evaluate extensive transaction data. In order to detect behavioral indicators of illicit activity, the proposed methodology implements both supervised and unsupervised models. In order to facilitate identification, it is necessary to acquire critical data, including network architecture, transaction frequency, wallet interactions, and temporal trends. In order to improve the model's stability and accuracy, we implement feature engineering and dimensionality reduction strategies. The algorithm is instructed by annotated datasets that include both legal and illicit transactions. Experimental assessment exhibits superior accuracy and improved information retention when contrasted with conventional detection methods. This method allows for the rapid identification of wallets and transaction patterns that are at a high risk.

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

1. INTRODUCTION

Cryptocurrency has revolutionized global financial transactions by facilitating the transfer and receipt of funds without the involvement of institutions. The efficacy, immutability, and transparency of numerous enterprises are improved by blockchain technology, which serves as the foundation for cryptocurrencies. The absence of genuine anonymity in bitcoin transactions enables the facilitation of illicit activities, including money laundering, fraud, terrorism financing, and ransomware payments. The security and stability of cryptocurrency ecosystems are causing a growing sense of concern among regulators, financial institutions, and law enforcement agencies as a greater number of users interact with digital assets.

Cryptocurrency money laundering is the term used to describe the intricate process of concealing the source of illicit currency through the use of multiple wallets, exchanges, and blending services. Decentralized exchanges, cryptocurrencies that obscure transactions, cross-border transfer technology, and tools that confound transaction tracing are all employed by criminals to conceal their activities. Conventional anti-money laundering (AML) techniques, which primarily rely on individual investigations,

Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/43
DOI: https://doi.org/10.5281/zenodo.20717743

customer identification, and centralized supervision, may be unable to detect intricate money laundering operations over blockchain networks. This contrast underscores the significance of automated and scalable methods for the rapid evaluation of large volumes of on-chain transaction data.

Blockchain transactions enable the public to access enormous quantities of data. This data contains critical information regarding the network's efficacy, the operation of transactions, and the communication between wallets. Transaction graphs and temporal patterns can reveal concealed connections and dubious activities, despite the anonymity of the individuals associated with wallet addresses. Manual analysis is impracticable due to the magnitude and velocity of blockchain data. Robust data-driven methodologies are indispensable for extracting valuable insights from transaction records and identifying suspicious or illicit activity within the complex and dynamic bitcoin networks.

Machine learning is an optimal tool for identifying bitcoin money laundering activities due to its capacity to accurately predict complex, non-linear patterns in extensive datasets. Supervised, unsupervised, and semi-supervised learning can be implemented to identify anomalous transactions, classify analogous wallet activities, and detect suspicious transactions. Graph-based data, transaction volume, wallet connectivity, temporal behavior, and transaction frequency can be used to develop prediction models. These algorithms can be modified to detect patterns that rule-based detection systems may fail to recognize and to address emergent methods of money laundering.

The examination of blockchain transaction networks has been made easier by recent developments in network-centric machine learning and graph analytics. The Bitcoin ecosystem can be analogized to a graph, with wallet addresses serving as nodes and transactions as edges. This enables the identification of transaction cycles that are linked to money laundering, dubious communities, and layering methodologies. Deep learning and graph neural networks have demonstrated the ability to identify temporal and structural relationships in transaction networks. As a result, the accuracy of identifying complex laundry patterns may be enhanced. These methods are capable of detecting criminal organizations and monitoring illicit financial transactions over time.

Machine learning has the potential to reduce the laundering of bitcoin, despite the existence of numerous unresolved issues. The list includes concerns regarding privacy, social inequality, inadequate data classification, and emerging methods of money laundering. The data regarding illegal transactions is frequently inaccurate and insufficient. As a result, the development of supervised models that are robust may prove to be a challenge. Models must be capable of operating efficiently in a variety of scenarios and for extended periods of time, as assailants frequently alter their strategies to evade detection. In order to address these obstacles and facilitate the detection of money laundering within Bitcoin ecosystems, it is essential to engage in rigorous feature engineering, hybrid modeling methodologies, ongoing model enhancements, and collaboration among researchers, regulatory bodies, and industry stakeholders.

JDEBM Journal of Digital Engineering and Business Management
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/43
DOI: https://doi.org/10.5281/zenodo.20717743

2. LITERATURE SURVEY

Lorenz et al. (2020) The primary focus of this investigation is the obstacles that arise when attempting to identify illicit Bitcoin transactions in the absence of classified data. In order to enhance the efficacy of detection, researchers implement supervised classifiers in conjunction with graph-based characteristics, despite the irregular distribution of classes. The method's effectiveness in detecting money laundering-related activities is illustrated by their examination of blockchain transaction graphs. In the presence of limited labels, memory can be substantially improved through semi-supervised learning.

Pettersson Ruiz & Angelis (2021) The primary objective of this project is to investigate the implementation of supervised learning in bitcoin exchanges to combat money trafficking across a variety of platforms. The model's efficacy, usability, regulatory compliance, and appropriate data utilization are evaluated by the authors. Random Forest and Gradient Boosting are more effective than alternative methods, as they produce more significant discoveries. However, the investigation underscores the challenges associated with the practical application of these technologies, particularly in terms of regulatory compliance.

Alotibi et al. (2022) The Elliptic dataset is employed by a multiplicity of individuals to compare deep learning models with conventional machine learning classifiers. They employ a variety of concepts and techniques, such as normalization and resampling, to address the issue of class disparities. These results suggest that deep learning algorithms are more effective at identifying money laundering because of their improved ability to remember illicit activities. The research suggests that deep learning is rapidly becoming an indispensable method for the identification of financial offenses.

Pocher& Romano (2023) Pocher and Romano investigate the Counter-Financing of Terrorism/Anti-Money Laundering through the use of graph-based forensic modeling. They employ fundamental AI approaches and network structures to identify complex money laundering operations. The results suggest that the norms are now more easily adhered to due to advancements in both discovery and comprehension. This initiative combines state-of-the-art technologies with formal governmental regulations.

Li & Park (2023) The application of temporal graph neural networks by Li and Park demonstrates the evolution of money trafficking within the bitcoin network over time. The modeling of temporal transaction patterns is facilitated by the incorporation of LSTM layers. Their methodology is superior to static graph models for assignments that require early identification. This investigation illustrates that AML systems are substantially affected by temporal connections.

Nguyen & Perez (2024) In order to identify illicit activities at the group level within sampled subgraphs, Nguyen and Perez implement contrastive learning. Rather than concentrating on specific transactions, their methodology aims to identify trends that are indicative of money laundering. The results suggest that it is possible to identify complex money laundering networks with increased

JDEBM
Journal of Digital Engineering and Business Managemet

Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/43
DOI: https://doi.org/10.5281/zenodo.20717743

accuracy. The necessity of replicating coordinated illicit activity is underscored by the research.

Rodríguez Valencia & Morales (2025) A thorough assessment of explainable AI and ML methodologies for bitcoin anti-money laundering is conducted by Valencia and Morales. Three criteria can be used to classify models: their comprehensibility, the extent of supervision they require, and the use of diagrams. The poll suggests that the primary concerns are compliance with regulations, insufficient data, and limited advancement prospects. There is a clear path, particularly in the context of the implementation of AML systems that are both transparent and effective within an explanation.

3. METHODOLOGICAL FRAMEWORK

Data preparation, data partitioning, analysis, and assessment are the critical phases of training and evaluating the model for the purpose of detecting money laundering transactions. The experiment's configuration is depicted in Figure 1. This chapter investigates a variety of transaction classification methodologies, including machine learning and deep learning. Among the classifiers are Naive Bayes, Random Forest, K-Nearest Neighbors, and Deep Neural Networks.

KNN: This uncomplicated instance-based learning method is employed to organize the new example in accordance with its similarity to all previous examples. To designate a class for the new instance, the most analogous class is used. For the first time, they implemented KNN based on instances.

RF: Individual trees may intermittently demonstrate increased vulnerability to training inputs and instability. In order to resolve this challenge and determine the class designation for each data point, the ensemble technique employs a compilation that aggregates and models their predictions. Decision trees are becoming more popular in data mining due to their adaptability, simplicity, and ease of comprehension, particularly when dealing with a variety of data categories. A random forest is a collection of trees that are employed for the purpose of classification or regression. The collaboration of these groups is effective when their members come from a variety of origins.

NB: Gaussian Naive Bayes, which is based on Bayes' theorem, is utilized by a multitude of Naive Bayes (NB) algorithms. The probability of an event occurring is determined by the Bayesian theorem, which is founded on prior knowledge of a pertinent condition. The objective of this approach is to identify the continuous attributes that are associated with each category and are distributed in a Gaussian fashion. The primary benefit of Naive Bayes is its effectiveness in supervised learning training. This indicates its capacity to resolve categorization obstacles that arise in practical applications. The erroneous presumption of attribute independence renders Naive Bayes intrinsically flawed. The principle of conditional independence is the foundation of Naive Bayes. The value of a class determines the independence of all aspects.

DNN: In order to accomplish hierarchical data representations, deep learning is a type of machine learning that eliminates

JDEBM
Journal of Digital Engineering and Business Managemet
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/43
DOI: https://doi.org/10.5281/zenodo.20717743

feature representations. It develops this representation autonomously by utilizing the training data. Deep neural networks (DNNs) are the foundation of this technology. They consist of critical components, including convolutions, perceptrons, and nonlinear activation functions. The organizing principle of the components is to facilitate the acquisition of increasingly complex concepts that augment students' existing knowledge. The number of layers in these strata can range from several hundred to over a thousand. Most networks establish connections between their fundamental constituents, which are vertices and edges, at the most fundamental level. The apex levels are characterized by the presence of significant attributes.

Evaluation Metrics

The model's efficacy in machine learning and deep learning is evaluated in accordance with the criteria we have established. The algorithms' efficacy is evaluated using the following metrics: F1-Score, Precision, ROC curve, and Recall. When examining non-linear datasets, these metrics are frequently implemented.

- "Precision" is the percentage of positive predictions that are accurate.
- The model's recall is a measure of its ability to quantify the number of true positive instances.

4. SYSTEM ARCHITECTURE

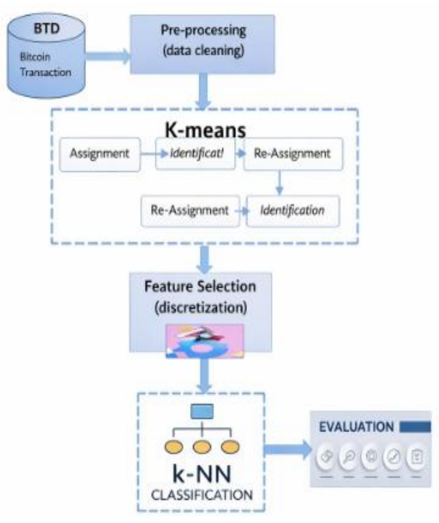

Fig1: System Architecture

The following are the components of the system's architecture:

i. Bitcoin Data
ii. Preprocessing
iii. Clustering Algorithm
iv. Feature Selection
v. Classification

Bitcoin Data: Each Bitcoin transaction is recorded on the blockchain. Transaction facts including quantity, timing, and costs, as well as sender and destination addresses, are included. Complex and difficult to understand, this knowledge is available to all. The raw data may reveal money-laundering or unethical actions. Blockchain technology is decentralized and transparent, making money transfers between addresses easy to track. However, the lack of IDs makes issue identification harder. This data underpins processing and analysis.

Preprocessing: Preprocessing Bitcoin transaction data for machine learning techniques is called preprocessing. This process may modify and delete data. This includes completing incomplete data, removing unnecessary information, and verifying that all statistics are within the same range. Removing unnecessary info

JDEBM
Journal of Digital Engineering and Business Management
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/43
DOI: https://doi.org/10.5281/zenodo.20717743

can improve readability. Structured feature vectors can be created using transaction graphs. Preparation improves model efficacy and data quality. This phase ensures clustering and classification input consistency and reliability.

Clustering Algorithm: Clustering organizes Bitcoin addresses or transactions by relationship without identifiers. This is done by observing user behavior. In transaction data, algorithms like K-means can identify hidden patterns. This level helps identify common and atypical transaction clusters. Unusual clusters may imply money laundering. Clustering helps understand large datasets by highlighting key patterns. Fundamental categories aid further research.

Feature Selection: Feature selection can determine which traits best detect illegal activities. Remove unnecessary elements to minimize size. Flow patterns, address connection, transaction diversity, and frequency are important. We now improve classification accuracy and lower computing costs. Classifying continuous variables with feature discretization is easy. With the right features, the model can better identify normal from suspicious behavior.

Classification: K-NN and other supervised algorithms may categorize new data by comparing it to examples. Transactions and addresses can be "genuine" or "suspicious." The classifier uses categorized transactions. The system has one last chance to decide. Money laundering can be detected using the model's predictions. The classification's precision, recall, and accuracy are then assessed.

5. RESULTS

Fig 2: Login Page

Fig 3: Registration Page

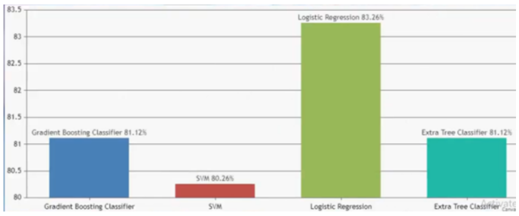

Fig 4: Classifier Accuracy Comparison

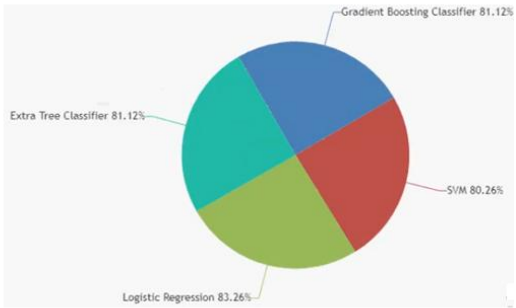

Fig 5: Accuracy Distribution Pie Chart

6. CONCLUSION

Machine learning to identify cryptocurrency money laundering is a clever and adaptive way to prevent financial crime in decentralized digital settings. Deep learning, unsupervised learning, and supervised learning models work together to detect complex money

JDEBM Journal of Digital Engineering and Business Management
Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/43
DOI: https://doi.org/10.5281/zenodo.20717743

laundering behaviors. By using graph-based feature engineering and temporal modeling, we can better analyze blockchain transactions.
Contrastive learning and graph neural networks are modern organized crime detection methods. Data imbalance and ID deficit must be addressed to develop reliable anti-money laundering (AML) systems. Explainable AI helps businesses comply with rules and make their models transparent. Data classification on bitcoin systems is simplified by preprocessing and feature selection. Experimental results show that machine learning-based AML frameworks are possible. However, models must be updated to reflect the latest money laundering strategies. When building large-scale surveillance technologies, privacy and ethics should come first. We should study hybrid models that combine representation learning and domain knowledge.

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.
11. Ouyang, S., Wang, X., & Zhang, H. (2024). Bitcoin money laundering

JDEBM
Journal of Digital Engineering and Business Management

Volume-2, Issue-2, April–June 2026
ISSN: 3107-7897 | P.ID: JDEBM/43
DOI: https://doi.org/10.5281/zenodo.20717743

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.
14. K. K. Gajula, "Leveraging Machine Learning Techniques for Earthquake Impact Analysis and Prediction," ZKG International, vol. 9, no. 2, pp. 909–913, 2024.
15. K. K. Gajula, "Deep Neural Forecasting Models for Climate Change Analysis Using Multivariate Time-Series Data," Journal of Computational Analysis and Applications, vol. 33, no. 8, pp. 6356–6368, 2024.
16. M. Shahimoon, H. Rasheed, D. R. Rao, K. R. Krishnaiah, K. K. Gajula, et al., "Evolving Predictive Models for Customer Churn in Telecommunications," in Proc. 2nd International Conference on New Frontiers in Communication Technologies, 2025.
17. K. K. Gajula, "Advanced Stress Level Forecasting Using Quantile-Optimized Regression Forest Models," 2025.