Identifying Money Laundering Risks in Digital Asset Transactions Based on AI Algorithms

Qian Yu, Zong Ke, Guofu Xiong, Yu Cheng and Xiaojun Guo

January 14, 2025

Preprint Version

Identifying Money Laundering Risks in Digital Asset Transactions Based on AI Algorithms

Qian Yu, Zong Ke*, Guofu Xiong, Yu Cheng, Xiaojun Guo

* Corresponding author: Zong Ke, a0129009@u.nus.edu

Identifying Money Laundering Risks in Digital Asset Transactions Based on AI Algorithms

Qian Yu^{1st}
Trine University, Detroit, Michigan, USA
Thea_QianYu@outlook.com

Zong Ke$^{co-first \,\&\, *}$
Faculty of Science, National University of Singapore, Singapore 119077
a0129009@u.nus.edu

Guofu Xiong^{3rd}
Ripple Labs, Inc, San Francisco, CA 94111, USA
gxiong@ripple.com

Yu Cheng^{4th}
Columbia University, New York, NY 10027, USA
yucheng576@gmail.com

Xiaojun Guo$^{5\text{th}}$
Independent Researcher, Jersey City, NJ 07302, USA
xiaojunguo2018@gmail.com

* Corresponding author: Zong Ke, a0129009@u.nus.edu

***Abstract*—This paper explores methods for detecting suspicious cryptocurrency transactions associated with money laundering, leveraging advanced AI algorithms. The study introduces a multi-model framework combining Generative Adversarial Networks (GANs), LSTM, Autoencoder-Based Anomaly Detection Model (ABAD), and other algorithms to address challenges like sample imbalance and noisy data. Graph-based feature engineering and embedding methods are utilized to construct transaction information graphs and extract meaningful patterns. The results demonstrate that the ensemble learning approach significantly outperforms individual models and traditional rule-based systems in detecting suspicious transactions. Despite its success, challenges such as imbalanced datasets, noise, and limited relational features remain. Future research is suggested to enhance model performance through graph neural networks and complex network-based methods. This work underscores the scalability and adaptability of machine learning models for addressing the evolving complexities of cryptocurrency money laundering.**

Keywords- ensemble learning; cryptocurrency; money laundering; deep learning; AI safety.

I. INTRODUCTION

Cryptocurrency, a form of digital asset derived from block-chain technology, is characterized by decentralization and anonymity. These features give cryptocurrency certain advantages, such as reducing transaction costs and improving transaction efficiency. However, these same features are exploited by criminals who use cryptocurrency as a tool for money laundering.

Currently, money laundering through blockchain typically occurs in three ways: (1) Centralized Exchange Liquidity: Some criminals use centralized exchanges to convert cryptocurrencies into fiat. However, this channel poses a risk of identity exposure, leading to the emergence of over-the-counter brokers that specialize in laundering funds. These brokers leverage their connections with exchanges to clean illicit cryptocurrencies. (2) Mixing Services: Money laundering is also conducted via specialized mixing services. These services aggregate and separate multiple unrelated transactions, creating long chains of complex transactions to obscure the source of the cryptocurrency. After the mixing process, the cleaned cryptocurrency is typically converted into fiat through centralized exchanges or OTC brokers, a common method of laundering. (3) Emerging Platforms: New decentralized platforms such as Non-Fungible Tokens (NFTs) and Decentralized Finance (DeFi) have created new channels for money laundering.

II. LITERATURE REVIEW

Among these methods, blockchain centralized exchanges' internal data, such as trading data and Know Your Customer (KYC) identity information, are commercially confidential [1]. Therefore, research on anti-money laundering (AML) in blockchain exchanges primarily focuses on public transaction records from the blockchain that are relatively easier to access [2-4].

In response, the dataset created by MIT's Weber and the blockchain data analytics company Elliptic is currently the largest publicly available dataset for blockchain fraud. They built an AML dataset based on known blockchain exchange addresses and publicly disclosed illegal money laundering addresses from the internet, and used this dataset to conduct research on exchange-based anti-money laundering [5].

However, Hu pointed out that the Elliptic dataset does not provide the real labeling process or exact feature information, making it difficult for algorithms to function effectively on other datasets [6]. Hu collected Bitcoin wallet addresses associated with money laundering services from online resources like WalletExplorer as labeled data. He extracted deepwalk and node2vec embedding features, along with 14 types of transaction graph statistics, using Adaboost as the base classifier to train multiple models, achieving good results. However, model performance is impacted by class imbalance, and the classification model only works effectively when the number of fraud labels reaches a certain threshold [7-11].

Furthermore, Oliveira proposed a graph construction technique based on illegal nodes. This method can extract new structural features, thus ameliorating Weber's work [12-16]. Experimental results show that this technique can yield an improvement of over 5% (from 91% to 97%) compared to the original features [17-19].

While these methods show some improvements, they still cannot fully address issues such as adaptability across different samples, accuracy with small datasets, and performance when datasets are rough[20-21]. Moreover, since money laundering techniques are highly diverse, the extent to which datasets align with the current real-world distribution is a critical consideration for practical applications[22].

This paper integrates multiple models to overcome issues like small sample sizes, including Random Forests, which perform well on noisy data, and Support Vector Machines, which are stable on small samples, to explore money laundering behaviors in cryptocurrency[23].

III. DATA

A. Data Sources

This paper classifies known data sources into three categories based on their functionality (as shown in Figure 1, with different data compositions): (1) On-chain transaction data, (2) Off-chain behavior data, (3) Publicly labeled fraud behavior data. Among them, the data sources of categories (1) and (2) are assumed to be unlabeled.

Specifically, on-chain transaction data: this refers to the blockchain transaction details data, contracts deployed on the chain, and their log data after being authenticated and recorded on the blockchain. Off-chain behavior data: this refers to data that occurs off-chain (usually on the Web) and contains clues about blockchain fraud activities. In existing research, this data source includes: a. Social media sites, such as Bitcoin forums and Twitter; b. Whitepaper publication sites, like Icobench; c. Exchange transaction data, such as Bitfinex and Mt.Gox. Fraud Behavior labeled data: this consists of data labeled as blockchain accounts, transactions, and contracts through complaint reports, case disclosures, online collection, and manual verification. Sources include: a. Websites disclosing illegal blockchain addresses, such as Etherscan and Cryptoscamdb; b. Websites disclosing real-world identities, such as public datasets related to WalletExplorer.

B. Public Datasets

Table 1 organizes the known public datasets. Based on existing identification methods, their data strategies are broadly divided into two categories: (1) transaction-based: Researchers primarily understand the behavior logic of both parties in blockchain transactions by analyzing transaction records. They aim to identify the transaction characteristics of target behaviors and thereby discover and identify transactions, accounts, and contracts associated with these target behaviors. (2) contract-based: The code logic in contracts provides a solid foundation for identifying fraud behaviors and ensuring their verifiability. Preemptively detecting and discovering vulnerabilities or traps in contract code can serve as a warning system for blockchain users. Table 1 lists the behaviors covered by datasets using contract-based strategies, including Ponzi schemes.

TABLE 1 PUBLIC DATASETS

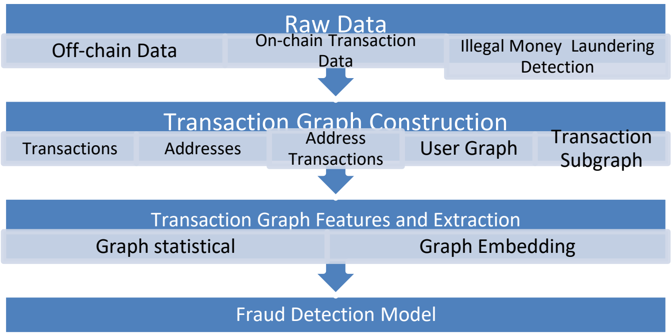
Data strategy	Open Datasets	Total Address/ Contract Count	Illegal Address /Contract Count
Trade_base	Elliptic	203769	4545
	BitcoinMixing	12360	---
	Xblock-Phishing	2973382	1157
	Bitcoin-Ponzi	32	32
	Btcransomware	4027	---
Contract_base	Contract-Ponzi	1382	131
	SADPonzi	1528	133
	HoneyBadger	857	323

C. Data Preprocessing and Feature Engineering

During exploratory analysis of the sample data, issues such as abnormal field values, incorrect values, missing data, and inconsistent data standards were found in the customer data, transaction data, behavioral data, and asset datasets. Since handling data anomalies is an engineering problem, the data quality issues from different source systems were addressed during the data ingestion process, following the standards and specifications for data warehousing.

Besides, the processing of blockchain abnormal transaction behavior data involves feature engineering preprocessing based on specific transaction rules. This includes processing data related to addresses, transactions, users, and trades. Among these, the transaction model serves as the underlying logic for constructing transaction information graphs (refer to TABLE 2 below). Two mainstream transaction models exist across various blockchain platforms: (1) The UTXO (Unspent Transaction Output) model, represented by Bitcoin, which is similar to the change-making model in real-world transactions. (2) The Account-based model, represented by Ethereum, which is akin to the bank account transaction model in the real world.

TABLE 2 DATA PREPROCESSING



D. Explanation of Different Components

Below is the construction of transaction information graphs. The address graph is used to describe the interaction information of digital currencies between addresses. The transaction graph represents the transaction flow of digital currency. The user graph shows digital currency flow between blockchain users. The transaction subgraph is a local graph within the full transaction graph, used to filter specific behavioral patterns within the transaction graph.

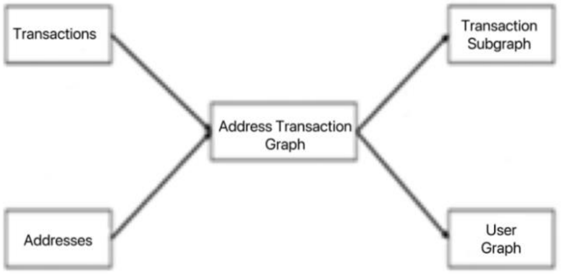


Figure 1 Relationships between transaction graphs

Figure 1 illustrates the relationships between the various transaction information graphs. Among them, the address graph and transaction graph are the most fundamental transaction information graphs, providing the relationship between addresses and transactions for the address-transaction graph.

E. Interaction Rules of Each Component

The Address Graph $G = (A, E)$ is used to represent the interaction patterns of digital currencies between addresses A and E; the Transaction Graph $G = (T, E)$ is used to represent the digital currency flow between transactions over time. The vertex set T in graph G represents the set of blockchain transactions, where each vertex t represents a blockchain transaction, identified by its transaction hash.

The Address-Transaction Graph $G = (T, A, E)$ introduces addresses into the transaction graph, making it an expanded form of the transaction graph. It describes the binary relationship between addresses and transactions. G is a heterogeneous graph containing two types of nodes: T and A. T denotes the set of transaction nodes. Each t represents a blockchain transaction identified by its transaction hash. A is the set of address nodes, where every node a represents a blockchain address. The set E represents the directed edge set of graphs G. The weight on an edge e can represent the amount of cryptocurrency transferred into an input address or the amount received by an output address, as following Figure 2.

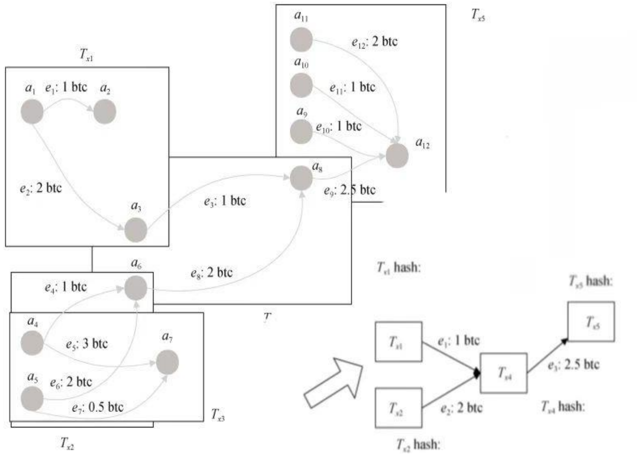


Figure 2 Trade construction process

User Graph $G = (U, E)$ is used to represent the flow of digital currency between block-chain users. U represents the set of user nodes in graph G, and E is the edge set, where each vertex u represents a blockchain user. The users here are typically abstracted from blockchain addresses using clustering or de-anonymization methods.

Transaction Subgraph: Common types of subgraphs include N-ego graphs and K-motifs graphs. Subgraph construction is typically centered around a node.

F. Feature Engineering

Building a good feature engineering framework on transaction graph structures is an essential step in downstream blockchain fraud detection algorithms. Among these, graph statistical features (attribute features) and embedding feature methods are two commonly used approaches in feature engineering. Graph Statistical Features are derived from the graph's topological structure. In blockchain fraud detection, graph statistical features introduce additional transaction attributes and behavior metrics, such as address activity days, transaction cycle, and the number of sending addresses. Graph embedding transforms the original graph data into a low-dimensional space while preserving key information, thereby improving the performance of subsequent tasks like node classification, relationship prediction, and node clustering. Graph embedding techniques can be further broadly divided into two types: node embedding and whole-graph embedding. The difference between the two is that node embedding learns a representation for a single node, as Table 3 below:

TABLE 3 COMMON GRAPH EMBEDDING TECHNIQUES

Node Embedding	GCN, Deep Walk, Random Walk, Skip-gram, node2vec
Subgraph Embedding	graph2vec, diffpool

IV. MODEL

Among the classical machine learning models widely used today, some models tend to reduce prediction variance, while others aim to reduce bias. Therefore, to mitigate the "bias-variance" trade-off, this paper adopts an ensemble learning approach to design a money laundering detection model for cryptocurrency transactions. The framework for its implementation is as follows.

A. Base Classifier Selection

Weak classifiers, such as Support Vector Machines (SVM), Random Forests (RF), GBDT, LSTM, GRU, Generative Adversarial Networks (GANs) and Autoencoder-Based Anomaly Detection Model (ABAD) often function as base classifiers of ensemble learning. To apply ensemble learning to suspicious transaction detection in cryptocurrency, this study measures the correlation of base classifiers' results using the Pearson correlation coefficient, ensuring that the selected base classifiers not only have high-performance evaluation metrics but also significant diversity.

However, in the context of cryptocurrency transaction detection where suspicious transactions are extremely imbalanced, most sample prediction probabilities are very close to each other. Therefore, the suspicious transaction detection problem can be further simplified into a binary classification problem.

$\mathrm{D} = \{(\mathrm{x}_1, \mathrm{y}_2), (\mathrm{x}_2, \mathrm{y}_2), (\mathrm{x}_3, \mathrm{y}_3), \dots\dots, (\mathrm{x}_\mathrm{m}, \mathrm{y}_\mathrm{m})\}$ is a dataset for the binary classification task $y_i \in \{-1,1\}$. a represents the number of samples predicted as positive class. a, b, c, and d satisfy the equation, $a + b + c + d = m$. The correlation measure is

$$\rho_{ij} = \frac{ab - bc}{\sqrt{(a+b)(a+c)(c+b)(b+d)}}$$

The range of ρ_{ij} is [−1, 1]. If h_i is independent of h_j, ρ_{ij} is 0. If there exists a positive correlation between h_i and h_j, ρ_{ij} is positive. Otherwise, ρ_{ij} is negative.

B. Ensemble Learning Model Construction

Derived from the method above, the base classifiers are selected, and the Bagging method is used for model training. Figure 3 shows the training process.

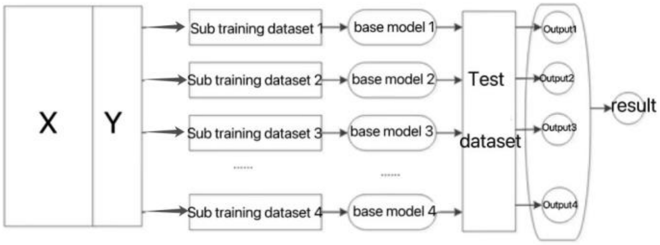


Figure 3 Training process ensemble learning

C. Assessment Approaches

TABLE 4 PREDICTION CLASSIFICATION CONFUSION MATRIX

Results	Predicted	
	1	0
1	PT	NF
0	PF	NT

In table 4, parameters are explained as below. PT, the model correctly predicts the transaction as a potentially illicit money laundering activity. PF, the model falsely predicts the transaction as a potentially illicit money laundering activity. NT, the model correctly predicts the transaction as a regular transaction. NF, the model falsely predicts the transaction as a regular transaction.

Recall denotes the percentage of accurately predicted money laundering samples out of the total actual money laundering samples. The F1-score is a comprehensive evaluation of both precision and recall. Kolmogorov-Smirnov (KS) Statistic quantifies the greatest disparity between the cumulative distributions of suspicious and normal transaction samples. AUC reflects the classifier's ability to rank samples.

D. Results

In the scenario of illicit transaction detection for anti-money laundering, the imbalance between normal and suspicious transactions is significant, making it a typical case of sample imbalance. The Area Under the Receiver Operating Characteristic Curve (AUC) is used as the evaluation metric for base classifiers. AUC reflects the classifier's ability to rank samples, and even in imbalanced class scenarios, it can objectively indicate the performance of a classifier. Therefore, considering the characteristics of anti-money laundering data in cryptocurrency transactions, AUC is selected as an performance measure for base model performance in this study.

Following the base model selection process, the models SVM, RF, GBDT, LSTM, GRU, GANs and ABAD were selected as base classifiers. These classifiers were trained on the training dataset. Table 5 shows the training results..

TABLE 5 RESULTS OF 7 BASE-MODELS

Model	**AUC**
GANs	0.8558
ABAD	0.8273
GRU	0.8159
LSTM	0.7748
SVM	0.7541
RF	0.7400
GBDT	0.7319

As the complexity of models increases, the results improve; and it can be found that unsupervised model works better. This may be driven by the fact: more than half of datasets are not labelled, making unsupervised learning fits better. Therefore, GANs, ABAD, GRU and LSTM were chosen as the base classifiers for the ensemble learning model, which was developed using the Bagging approach. To verify the performance of the ensemble learning model proposed in this paper on the cryptocurrency anti-money laundering dataset, GANs, ABAD, GRU, and LSTM were separately executed on the validation set, and their prediction outcomes were compared.

To substantiate the effectiveness of the ensemble model built in this paper, the results are shown in Figure 4. Compared to rule-based models, which fail to predict money laundering samples, machine learning models are indeed better at identifying the correspondence between money laundering features and labels, and they are proved to be effective on real-world datasets.

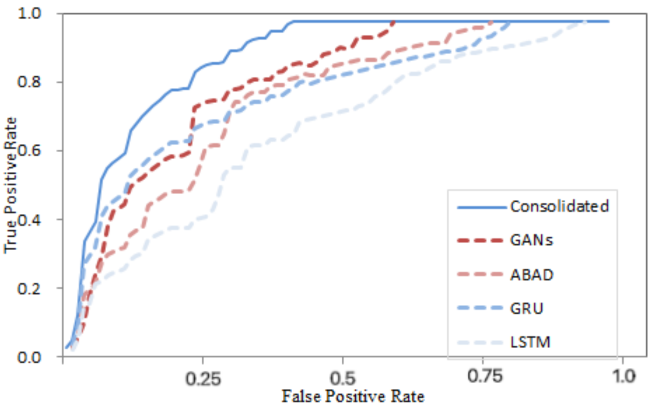


Figure 4 Performance in test datasets

V. Conclusions

Single machine learning models can, to some extent, identify suspicious money laundering transactions in cryptocurrency transactions, but ensemble learning models that integrate multiple weak classifiers significantly outperform individual models. Additionally, this model is scalable and allows flexible adjustment of base models to suit different datasets. From the experimental process and result analysis, there are still the following issues with this approach: In the experimental data used in this study, suspicious money laundering transaction samples are very rare, and the performance of individual models is poor. Second, during the feature engineering phase of this experiment, features were designed solely from a single customer perspective, without considering the relationship features between customers or the transmission of money laundering risks across relational networks. For further research on anti-money laundering modeling, it would be beneficial to explore methods based on complex networks or graph neural networks to recognize suspicious illicit financial activity transactions.

References

[1] Oliveira, C., Torres, J., Silva, M. I., Apar ício, D., Ascens ão, J. T., & Bizarro, P. (2021). GuiltyWalker: Distance to illicit nodes in the Bitcoin network. arXiv preprint arXiv:2102.05373.

[2] Lorenz, J., Silva, M. I., Apar ício, D., Ascens ão, J. T., Bizarro, P., & Ascens ão, J. T. (2005). Machine learning methods to detect money laundering in the bitcoin blockchain in the presence of label scarcity (2020). Preprint arXiv.

[3] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., ... & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484–489.

[4] P. Tasatanattakool, P. Wannapiroon and P. Nilsook, "Digital Asset Management Process using AI TRiSM," 2024 IEEE International Conference on Cybernetics and Innovations (ICCI), Chonburi, Thailand, 2024, pp. 1-6.

[5] H. Tiwari, A. Raj, U. K. Singh and H. Fatima, "Generative AI for NFTs using GANs," 2024 INDIACom, New Delhi, India, 2024, pp. 488-492.

[6] Hu, Y., Seneviratne, S., Thilakarathna, K., Fukuda, K., & Seneviratne, A. Characterizing and detecting money laundering activities on the bitcoin network. arXiv 2019. arXiv preprint arXiv:1912.12060.

[7] M. Raj, H. Khan, S. Kathuria, Y. Chanti and M. Sahu, "The Use of Artificial Intelligence in Anti-Money Laundering (AML)," 2024 3rd International Conference on Sentiment Analysis and Deep Learning (ICSADL), Bhimdatta, Nepal, 2024, pp. 272-277.

[8] Xie, T., Tao, S., Li, Q., Wang, H., & Jin, Y. (2022). A lattice LSTM-based framework for knowledge graph construction from power plants maintenance reports. Service Oriented Computing and Applications, 16(3), 167-177.

[9] K. Balaji, "Artificial Intelligence for Enhanced Anti-Money Laundering and Asset Recovery: A New Frontier in Financial Crime Prevention," 2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI), Coimbatore, 2024.

[10] Luo, D. (2024). Decentralized Energy Markets: Designing Incentive Mechanisms for Small-Scale Renewable Energy Producers.

[11] K. K. Girish and B. Bhowmik, "Money Laundering Detection in Banking Transactions using RNNs and Hybrid Ensemble," 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024, pp. 1-7.

[12] Peng, Z., Jian, J., Wang, M., Wang, Q., Boyer, T., Wen, H., ... & Chen, K. P. (2019, September). Big data analytics on fiber-optical distributed acoustic sensing with Rayleigh enhancements. In 2019 IEEE Photonics Conference (IPC) (pp. 1-3). IEEE.

[13] M. B. Jamshidi, M. Gorjiankhanzad, A. Lalbakhsh and S. Roshani, "A Novel Multiobjective Approach for Detecting Money Laundering with a Neuro-Fuzzy Technique," 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), Banff, AB, Canada, 2019, pp. 454-458.

[14] Jin, Y., Fu, G., Qian, L., Liu, H., & Wang, H. (2021). Representation and Extraction of Diesel Engine Maintenance Knowledge Graph with Bidirectional Relations Based on BERT and the Bi-LSTM-CRF Model. In 2021 IEEE International Conference on e-Business Engineering.

[15] D. V. Kute, B. Pradhan, N. Shukla and A. Alamri, "Deep Learning and Explainable Artificial Intelligence Techniques Applied for Detecting Money Laundering–A Critical Review," in IEEE Access, vol. 9, pp. 82300-82317, 2021.

[16] H. N. Mohammed, N. S. Malami, S. Thomas, F. A. Aiyelabegan, F. A. Imam and H. H. Ginsau, "Machine Learning Approach to Anti-Money Laundering: A Review," 2022 IEEE NIGERCON, Lagos, Nigeria, 2022.

[17] Sui, M., Zhang, C., Zhou, L., Liao, S., & Wei, C. (2024). An ensemble approach to stock price prediction using deep learning and time series models.

[18] M. Mahootiha, A. H. Golpayegani and B. Sadeghian, "Designing a New Method for Detecting Money Laundering based on Social Network Analysis," 2021 26th International Computer Conference, Computer Society of Iran (CSICC), Tehran, Iran, 2021, pp. 1-7.

[19] P. Y. Leonov, V. M. Sushkov, V. V. Krasinsky, V. A. Romanovsky, N. V. Kuznetsova and N. V. Akimov, "Detecting Money Laundering Patterns through Cash Flow Analysis: a Neural Network-Based Approach," 2023 IEEE XVI International Scientific and Technical Conference Actual Problems of Electronic Instrument Engineering (APEIE), Novosibirsk, Russian Federation, 2023, pp. 1390-1393.

[20] S. Marasi and S. Ferretti, "Anti-Money Laundering in Cryptocurrencies Through Graph Neural Networks: A Comparative Study," 2024 IEEE 21st Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2024, pp. 272-277.

[21] Xie, Y., Aggarwal, K., & Ahmad, A. (2023). Efficient continual pre-training for building domain specific large language models. arXiv preprint arXiv:2311.08545.

[22] G. Konstantinidis and A. Gegov, "Deep Neural Networks for Anti Money Laundering Using Explainable Artificial Intelligence," 2024 IEEE 12th International Conference on Intelligent Systems (IS), Varna, Bulgaria, 2024, pp. 1-6.

[23] M. Mehmet and D. Wijesekera, "Using dynamic risk estimation & social network analysis to detect money laundering evolution," 2013 IEEE International Conference on Technologies for Homeland Security (HST), Waltham, MA, USA, 2013, pp. 310-315.