Review on Machine Learning in Suspecting Money Laundering

Nazish Sana Khan [1], Arun Jhapate [1]
[1] Nazish Sana Khan, Department of Computer Science, Sagar Institute of Research & Technology, Bhopal , India
[2] Arun Jhapate, Department of Computer Science, Sagar Institute of Research & Technology, Bhopal , India

* Corresponding Author: Nazish Sana Khan

Abstract: - *Money laundering is a method used by criminals to conceal the illegal source of their revenue. Money is "cleaned" of its illicit origin and made to seem as legitimate business profits by moving it via intricate transfers and activities, or through a series of business. Large, organized criminal organizations – such as drug smuggling operations – face a severe commercial difficulty in that they end up with large sums of money that they must conceal in order to escape legal authorities conducting investigations. The beneficiaries of such vast sums of money also don't want to be forced to declare it as income, resulting in massive tax liabilities. This paper provides the comprehensive review of money laundering and techniques to prevent the problem of money laundering.*
Keywords: Money Laundering, Cryptocurrency, Transactions, etc.

I. INTRODUCTION

Money laundering is a method used by criminals to conceal the illegal source of their revenue. Money is "cleaned" of its illicit origin and made to seem as legitimate business profits by movingit via intricate transfers and activities, or through a series of business. A sort of financial crime is money laundering. It entails concealing the origins of fraudulently obtained revenues (dirty money) so that they look to come from a respectable source. Anti-money laundering (AML) refers to the efforts that financial firms do in in order to comply with legal obligations to actively monitor and report unusual behaviour [3].
Cryptocurrency, sometimes known as crypto-currency or crypto, is any type of digital or virtual currency that uses encryption to safeguard transactions. Cryptocurrencies operate without a central issuance or regulatory authority, instead relying on a decentralised system to track transactions and create new units. All cryptocurrency transaction records are irrevocable and recorded in blocks that are connected in chronological order, according to blockchain technology [4]. Cryptocurrencies transaction records comprising rich information and entire trails of financial activity are publically available thanks to the open and transparent nature of blockchain, giving researchers with unparalleled prospects for data analysis in this domain. The main benefit of analysing and mining cryptocurrency transaction data is twofold: (1) Transaction records in standard finance contexts have received little attention in previous research since they are typically not made public for reasons of security and interests. We may thoroughly study trading behaviours, distribution of wealth, and generating mechanisms of a transaction system, as well as deduce reasons for changes in the cryptocurrencies financial sector, by analyzing and mining cryptocurrencies transaction data.

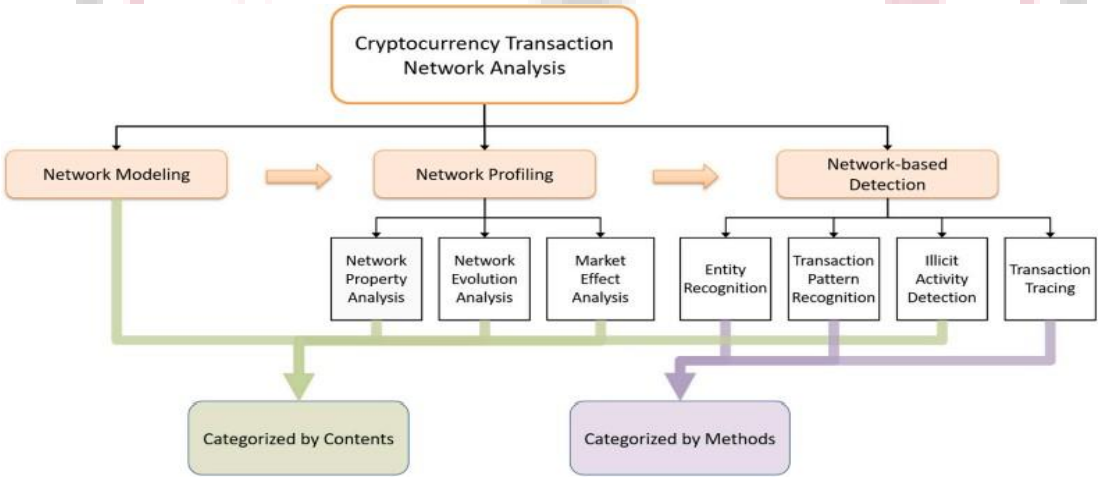

Figure 1 Cryptocurrency Transaction Network Analysis [6]

II. CRYPTOCURRENCY TRANSACTION NETWORK ANALYSIS

Following data collection, the first stage in doing bitcoin transaction network analysis is to represent transaction data in a network-structured data form, which is known as network modeling.Unlike traditional network modeling, cryptocurrencies transactional network service modeling must deal with diverse blockchain data from different sources. On the one hand, because of the data heterogeneity, the identification of node and linkages must be carefully created in compare to conventional networks with well-defined nodes and links, such as citations systems.

III. ANTI MONEY LAUNDERING

The monster known as money laundering is one of the most worrisome dangers to the global economy's security and prosperity. Certain guidelines have been put in place to stop this monsters and the horrible devastation it causes. Anti–Money guidelines are what they're termed. It has been suggested that financial institutions move away from rule-based and hazardous techniques and toward those that portray risk as their major motivator for a seamless laundering money strategy [7].

Anti-money laundering (AML) is a combination of acts, laws, processes, and regulations aimed at detecting and preventing any practices that contribute to unlawful revenue generation [1]. In the first two processes, placement and layering, the tainted money should be recognized. Otherwise, tracing its origins in the third phase will be difficult. It is imperative that automated tools be used to assist in the detection of suspicious transactions. Every financial institution is accountable for putting global regulations into effect. Various software tools are now commercially available to assist banks in detecting suspicious activities and determining which customers are safe to create a business channel together.

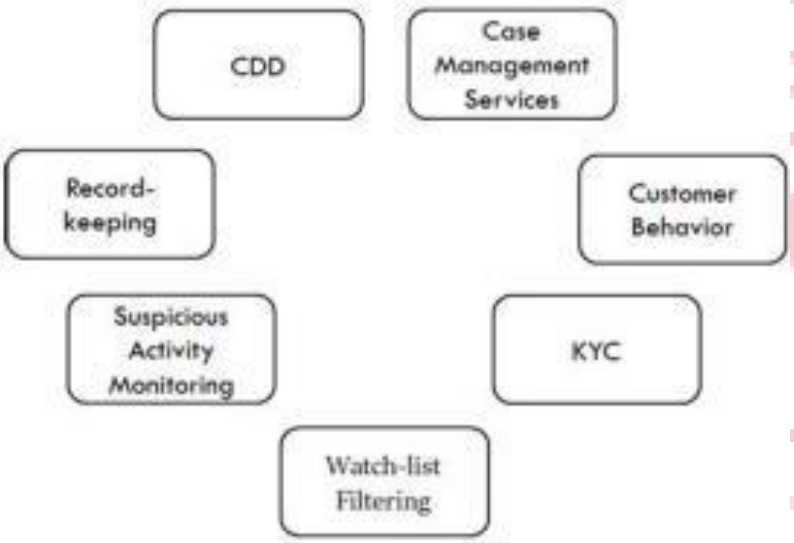

Figure 2 Key elements of anti money laundering

Money laundering is a concern to the global economy, hence Anti-Money Laundering (AML) rules have been established to mitigate the harm that this activity can do. To combat money laundering, the Financial Action Task Force began recommending that financial companies move away from rule-based systems and toward a more risk-based approach. To combat financing of terrorism and money laundering, the European Union recently approved the 'Fifth Anti-Money Laundering Directive' (5AMLD) [8]. The law covers exchanges that allow users to trade crypto-to-crypto or crypto-to-fat (or vice versa) as well as wallet service providers that retain the users' private keys. These businesses must conduct proper consumer due diligence, monitor transactions, keep track of customer information, and report any questionable transactions. The Financial Action Task Force (FATF) issued similar advice to its member nations (including the United Kingdom and the United States) to regulate cryptocurrencies marketplace. These rules are in place to keep criminals from infusing illegally obtained funds into the financial sector without being caught.

As the HSBC and Danske Bank scandals demonstrated, a lack of AML control can result in hefty fines. Because of the growing popularity of cryptocurrency, as well as the threat of criminals hiding behind the technology's completely anonymous nature, governments have begun to enact legislation to prevent financial fraud. A rule-based system is one of the most basic approaches for detecting money laundering. To detect suspicious activity, a rule-based system typically consists of a set of conditions that verify whether particular events occur or thresholds are surpassed. However, this method has downsides, one of which is that it does have a high false-positive rate and necessitates the use of subject matter experts to design rules [8]. Some of these flaws are solved through computer vision, which infers patterns from previous data, lowering the falsepositive rate while keep the false-negative rate low.

IV. MACHINE LEARNING IN AML

Machine learning is crucial in the banking business when it comes to preventing money laundering. To detect alerts and suspicious transactions detected by the internal financial system, it uses supervised machine learning, in which an ML model is trained with several sorts of data or trends. These machine learning algorithms help detect suspicious transactions, sender and beneficiaries financial data, transactional pattern based on transaction histories, and other things [9].

Machine learning techniques help AML by reducing human errors to a considerable extent. Machine learning models use a variety of tactics to tackle money laundering. Robots can comprehend human language and recognize warnings, process mortgages, screening for bad news, and screen repayments, among many other things, thanks to Natural Language Processing (NLP). Machine learning technologies also help with the identification of multiple suspicious acts as well as monitoring procedures. Machine learning (ML) teaches computers to recognize and identify transaction patterns, behaviour, suspicious users/accounts, and alert categorisation based on the risk levels such as high, medium, and low. It also keeps track of alarms, clears some warning automatically, and fully operationalizes accounts based on their behavior and paperwork.

Machines can be used to educate alerts to recognise, score, triage, enhance, close, or hibernating. People find these operations complex and time-consuming, but with the appropriate machine learning technologies, they become much easier than the previous method. Natural Language Generation (NLG) assists in the compilation and creation of Suspicious Activity Reports (SARs). This strategy can reduce the need for human operators for routine tasks, reduce the time it takes to triage alerts, and allow professionals to concentrate on more important and complex tasks.

With the inclusion of computer vision into AML TM alert triage, SAR conversion rates should rise from the current unacceptable rate of 1% in the financial sector. Machine learning is widely used in the banking and financial industries, and AML is one of the most prominent applications of machine learning.

Figure 3 Machine learning approach – model construction and training AML

Implementation of AI/ML will require:

- Recognizing and understanding the transactional patterns of customers with similar tendencies, and the unusual and unanticipated.
- Learn and understand the categories of money laundering, scams, and terrorist funding to recognise category-specific frauds.
- Finding connections between warnings that resulted in confirmed fraudulent activity and alerts that resulting in false-positive alarms.
- Regularly analyzing false-positive warnings and recognizing common predictors.

Many financial institutions have started using RPA-based automation process and see AI/ML as the next stage in their efficiency journey. AI/ML technology can aid in improving, automating, and speeding up AML processes. Moreover, to combat money laundering, these technology can scale to manage and control the huge amount, speed, and diversity of data generated by today's financial companies. Banks and financial institutions must now include AI/ML into their networks and systems. Financial crime, fraud, compliance, regulations, and money laundering are all issues that AI/ML technology can help with.

V. RELATED WORK

(Raiter, 2021) Criminals will be capable of transferring money around the world in a fraction of a second as international currency transactions become increasingly automated, while regulators will be able to inspect and monitor international currency mobility and discover strange patterns of money movement. Machine learning algorithms could be a useful supplement to the present anti- money launder efforts. Using a synthetic data that exactly approximate typical

transactional behaviour, this study empirically assessed four machine learning techniques (Logistic regression, SVM, Random Forest, and ANN). After examining the performances of several algorithms, it can be concluded that, when contrasted to the other ways, the Random Forest technique delivers the best accuracy. The Artificial Neural Network was the least accurate method (ANN).

(Chen et al., 2018) The goal of this work is to present a complete overview of machine learning algorithms and approaches for detecting fraudulent transactions. Anti-money laundered tropes, link analysis, behaviour modeling, possible risk, outlier detection, and geographic capabilities solutions, in particularly, have been found and analyzed. Existing machine learning algorithms and techniques dealt with in detail in the literature have been categorised, summarized, and compared; key steps of data preprocessing, data processing, and analytics techniques have been discussed; and extant machine learning algorithms and techniques have been categorised, summarized, and especially in comparison.

(Chahal & Gulia, 2019) This paper explains the fundamental architecture of Transfer Learning. A comparison of machine learning and deep learning is also provided in the publication, allowing researchers to gain a wide understanding of these techniques and choose which one would be the best answer for a given situation.

(Andrade et al., 2021) This study presents a machine learning-based system for determining not whether a corporation is likely to commit fraud. Four alternative classifiers – k-Nearest Neighbors, Random Forest, Support Vector Machine (SVM), and a Neural Network – were trained which is then used to detect fraud using tax and financial data from diverse organizations. With the Random Forest, the best-performing model has achieved a macro-averaged F1-score of 92.98 percent.

(Oad et al., 2021) This study provides a method for detecting abnormal behaviors using blockchain-enabled transaction scanning (BTS). The BTS technique outlines the rules for detecting outliers and transferring funds quickly, limiting abnormal transactions. The rules define the exact patterns of harmful behavior in the transaction. In addition, the BTS method's rules monitor the recent transactions and generate a list of entities that receive money suspiciously. Finally, a blockchain-based system is employed to prevent money laundering.

(Shaikh et al., 2021) The current problem in the sector is determining the relevant ties between dubious consumers and money laundering. To address this issue, this study analyzes the difficulties in recognizing associations such as business and familial affiliations, and provides a strategy for employing social network analysis to uncover links between suspicious clients (SNA). The suggested approach intends to identify different mafias and groups participating in laundering money, assisting in the fight against money laundering and potential terrorist financing. To identify suspicious clients and transaction, the suggested approach uses relationship data from customer profile and social media function indicators. A series of tests using financial data have been done, and the results suggest that the suggested framework can provide real-world benefits to financial organizations.

(Vassallo et al., 2021) This research focuses on detecting illegal activity (such as scams, terrorism financing, and Ponzi schemes) on cryptocurrencies infrastructure, at both accounts and transactions levels. Previous research has found that in this arena, imbalanced data and the dynamic environment caused by criminals' changing strategies to avoid discovery are pervasive. To better manage dynamic situations, we offer Adaptive Stacked eXtreme Gradient Boosting (ASXGB), an adaptation of eXtreme Gradient Boosting (XGBoost), and present a comparative analysis of several offline decision tree-based ensemble and heuristic-based data-sampling strategies. Our findings show that I offline decision tree-based gradient boosting algorithms outperform state-of-the-art Random Forest (RF) results at both the account and transaction levels, (ii) the data-sampling approach NCL-SMOTE improves recall at the transaction level, and (iii) proposed ASXGB successfully reduced the burden of model uncertainty while working to improve recall at the transaction level.

(Ketenci et al., 2021) In this paper, authors develop a novel feature set based on time-frequency analysis that leverages 2-D representations of monetary operations to improve detection performance of fraudulent financial monitoring devices for AML systems. As a machine learning method, random forest is used, while simulated annealing is used for hyperparameter tweaking. As a result, these features significantly increase the area under curve findings of current data scientific transactions monitoring systems (by over 1%). A false positive rate of 14.9 percent was reached using only time-frequency characteristics, with an F-score of 59.05 percent. When transactional and CRM capabilities are included, the false positive rate drops to 11.85 percent, and the F-Score rises to 74.06 percent.

(Canhoto, 2021) In this work, authors look at how predictive modeling algorithms' technological and contextual affordances might help these organizations attain that goal. We discovered that there is little opportunity for employing machine learning for laundering money techniques due to the lack of high-quality, massive trained data. On the other hand, reinforcement machine learning and, to a degree, unsupervised classification can be used to simulate aberrant financial behavior rather than actual money launder.

(Kute et al., 2021) The purpose of this work is to review the present state-of-the-art literature on DL and XAI for detecting suspect laundering money transactions, as well as to identify future research areas. The review's major finding are that researchers prefer Convolutional Neural Networks and AutoEncoder variations; graph deep learning combined with natural language is starting to emerge as an important development for AML; XAI use is not seen in the AML domain; and XAI use is not seen in the AML domain. 51 percent of AML approaches are incomprehensible, and 58 percent of studies employed a sample of old real data; significant obstacles for academics include access to recent real transaction data and a paucity of labelled training data, as well as data that is severely imbalanced. Application of XAI techniques to bring out explainability, graph machine learning using natural language processing (NLP), unsupervised and reinforcing steel learning to handle lack of labeled examples, and research collaborations programs between the research community and industry to benefit from domain expertise and controlled access to the data are some of the suggestions for future research.

VI. CONCLUSION

This article shows the work done by different researchers in machine learning techniques to identify suspicious transactions in an anti-money laundering system. Machine Learning is a prominent technique to predict in advance financial crime and lower its threat. Another benefit of adopting machine learning for AML is that it enhances the performance of the model based on learning from historical transactional data. Additionally we have discussed about key elements of anti money laundering, process of machine learning and also the cryptocurrency network transaction analysis.

References

[1] Raiter, O. (2021). Applying Supervised Machine Learning Algorithms for Fraud Detection in Anti-Money Laundering. Journal of Modern Issues in Business Research, 1(1), 14–26. https://international-journals.website/index.php/JMIB/article/view/4

[2] Chen, Z., Van Khoa, L. D., Teoh, E. N., Nazir, A., Karuppiah, E. K., & Lam, K. S. (2018). Machine learning techniques for anti-money laundering (AML) solutions in suspicious transaction detection: a review. Knowledge and Information Systems, 57(2), 245–285. https://doi.org/10.1007/s10115-017-1144-z

[3] Chahal, A., & Gulia, P. (2019). Machine learning and deep learning. International Journal of Innovative Technology and Exploring Engineering, 8(12), 4910–4914. https://doi.org/10.35940/ijitee.L3550.1081219

[4] Andrade, J. P. A., Paulucio, L. S., Paixão, T. M., Berriel, R. F., Carneiro, T. C. J., Carneiro, R. V., Souza, A. F. De, Badue, C., & Oliveira-Santos, T. (2021). A Machine Learning-based System for Financial Fraud Detection. 165–176. https://doi.org/10.5753/eniac.2021.18250

[5] Oad, A., Razaque, A., Tolemyssov, A., Alotaibi, M., Alotaibi, B., & Zhao, C. (2021). Blockchain-enabled transaction scanning method for money laundering detection.

[6] Electronics (Switzerland), 10(15), 1–18. https://doi.org/10.3390/electronics10151766

[7] Shaikh, A. K., Al-Shamli, M., & Nazir, A. (2021). Designing a relational model to identify relationships between suspicious customers in anti-money laundering (AML) using social network analysis (SNA). Journal of Big Data, 8(1). https://doi.org/10.1186/s40537-021-00411-3

[8] Vassallo, D., Vella, V., & Ellul, J. (2021). Application of Gradient Boosting Algorithms for Anti-money Laundering in Cryptocurrencies. SN Computer Science, 2(3), 1–15. https://doi.org/10.1007/s42979-021-00558-z

[9] [Ketenci, U. G., Kurt, T., Onal, S., Erbil, C., Akturkoglu, S., & Ilhan, H. S. (2021). A Time-Frequency Based Suspicious Activity Detection for Anti-Money Laundering. IEEE Access, 9, 59957–59967. https://doi.org/10.1109/ACCESS.2021.3072114

[10] Canhoto, A. I. (2021). Leveraging machine learning in the global fight against money laundering and terrorism financing: An affordances perspective. Journal of Business Research, 131(October), 441–452. https://doi.org/10.1016/j.jbusres.2020.10.012

[11] [Kute, D. V., Pradhan, B., Shukla, N., & Alamri, A. (2021). Deep Learning and Explainable Artificial Intelligence Techniques Applied for Detecting Money Laundering- A Critical Review. IEEE Access, 9, 82300–82317. https://doi.org/10.1109/ACCESS.2021.3086230

[12] Weber, M., Domeniconi, G., Chen, J., Karl Weidele, D. I., Bellei Elliptic, C., Robinson Elliptic, T., Leiserson, C. E., Bellei, C., & Robinson, T. (2019). Anti-Money Laundering in Bitcoin: Experimenting with Graph Convolutional Networks for Financial Forensics. https://doi.org/10.48550/arxiv.1908.02591

[13] Kumar, A., Das, S., Tyagi, V., Shaw, R. N., & Ghosh, A. (2021). Analysis of Classifier Algorithms to Detect Anti-Money Laundering. Studies in Computational Intelligence, 950, 143–152. https://doi.org/10.1007/978-981-16-0407-2_11