Machine Learning for Anti-Money Laundering (AML): A Comprehensive Analysis

Pranav Khare[1], Shristi Srivastava[2]

[1,2] Independent Researcher, [1,2] Woodinville, WA, USA,

---***---

Abstract - *A major economic concern, money laundering has had far-reaching effects on both individuals and financial institutions throughout the last ten years. Because of either technology developments or the rising sophistication of criminal tactics, the industry's current anti-money laundering (AML) measures are either ineffective or both. An expanded summary of the ML techniques used in the literature to discover and evaluate money laundering detection using transaction monitoring is given in this work. This work delves into the "layering" aspect of supervised learning for money laundering fraud prediction, aiming to provide a rational and scientific approach to risk forecasting. MLP and GCNs are the names of the deep learning-based money laundering and fraud prediction models that are suggested. Based on the Elliptic dataset, the experimental results show that MLP-GCN achieved 98% and 97% accuracy compared to the AML fraud prediction model currently regarded as optimal, respectively, in comprehensive feature modelling when compared to other DL models and traditional algorithms.*

***Key Words*: Anti-Money Laundering (AML), machine learning, deep learning, Elliptic, Multi-Layer Perceptron, Graph Convolutional Networks.**

1. INTRODUCTION

The predicate or underlying crime generates the funds that will be used for money laundering, which begins with those funds and ends with them being used safely or with minimum risk [1]. Money laundering is when illicit gains are transformed into legitimate assets. There have been far-reaching effects on both society and financial institutions as a result of money laundering in the modern era [2]. Most techniques for making the source of money seem to be legitimate and concealing their origins include three steps: placement, layering, and integration [3]. The three steps provide a framework, but they are not meant to be taken as an absolute method; they are just a model [4].

The goal of AML legislation, practices, and policies is to reduce the amount of illicit funds that criminals are able to transform into legal assets. As an international institution, the FATF is devoted to combating money laundering. There is a new risk-based framework in place by the FATF that requires banks to put an emphasis on client due diligence and transaction monitoring [5]. Financial institutions are obligated to monitor patterns and anomalies in client transactions in order to continue risk assessment. Consequently, they implement transaction monitoring. Alerts are generated in a variety of systems by a variety of actions, such as unusual transaction sizes, locations, changes in client behaviour, or exaggerated wealth that differs from customer reports. When an institution believes a customer's behaviours are questionable, it files and sends suspicious activity reports (SAR) to FIU. Due to the complexity of identifying financial crime and the imposition of substantial penalties, the task of transaction monitoring for financial institutions has become more costly and challenging than ever[4]. Systems for monitoring transactions usually use principles to spot money laundering. Rules are basically if-then statementing with thresholds that are predefined and taken from past performance or business judgement. Money launderers are constantly modifying their strategies to exploit vulnerabilities and assault monitoring systems, necessitating the continual validation and calibration of the rules. There might be a labor-intensive component to this operation, which could cause a delay. Furthermore, rule-based systems may not be sufficient for recognising complex transaction patterns since they are intended to detect single transactions or simple transaction patterns. In contrast, machine learning creates models that can independently adjust to new data[6].

The difficulties with rule-based approaches have led to an increased focus on ML to reduce false positives and provide effective outcomes [7]. The institution's reputation can be maintained, operational costs can be substantially reduced, and regulators' requirements can be satisfied through the use of ML approaches. The unique features of the transactional data have a major impact on the method's efficacy, thus it's critical to recognise the benefits and drawbacks of algorithms when using them to identify money laundering. Principally, supervised and unsupervised ML methodologies have been implemented to identify instances of money laundering[8]. A labelled dataset facilitates the evaluation of supervised approaches; however, high-quality labels are necessary. Unsupervised methods are more capable of adapting to differences in customer behavior and identifying suspicious activity than supervised methods [9][10][11]. A detection of AML activities is critical in combating financial crime and maintaining the integrity of financial systems. With the advent of cryptocurrencies and the increasing complexity of financial transactions, traditional AML methods are becoming less effective. This study uses MLP and GCN models and the rich datasets from Elliptic to find out and forecast illicit transactions. The development of this research is to extend the efficiency of machine learning in detecting AML and give the financial organizations and the regulatory bodies efficient ways to accurately detect and reduce money laundering in the short time. The fact that such models have demonstrated their capacity to handle, analyze and solve large and lasting data, strengthens the understanding of their promise to revolutionize AML. The contributions of the paper as:

- The paper leverages the Elliptic Dataset for the purpose of anti-money laundering (AML) prediction. This dataset, comprising transactions from the Bitcoin blockchain, has not been thoroughly explored in machine learning literature, particularly for AML applications.
- The study conducts a comprehensive comparative analysis of two advanced machine learning algorithms: MLP and GCN. The merits and limitations of each approach in the context of AML are emphasised in the paper by employing these algorithms to the same dataset.
- The paper emphasizes the importance of feature reduction in handling the large and complex Elliptic Dataset. It details the preprocessing steps taken to ensure that the models can efficiently handle the data without compromising performance, demonstrating effective techniques for feature selection and reduction.
- A novel methodological approach is employed by splitting the dataset temporally. This means that the training and testing sets are divided based on time-steps, simulating real-world scenarios where models are trained on past data to predict future transactions.
- Both MLP and GCN models demonstrate high performance according to recall, F1-score, accuracy, and precision.
- A paper benchmarks a performance of MLP and GCN against other existing models such as MDGC-LSTM, Random Forest, C4.5, and KNN. This comparative analysis underscores a superior performance of a proposed models, providing a clear understanding of their advantages and potential for real-world AML applications.

1.1 Organization of paper

Here is a summary of the remaining portion of the paper. Section 2 provides a summary of relevant studies on the subject. Section 3 lays out the performance evaluation criteria, methodology, models, and flow diagram. Section 4 includes a comparison, the findings, and some remarks. Section 5 presents the study's conclusions and discusses potential next steps.

2. LITERATURE REVIEW

This section offers a literature overview on a topic of using DL and ML methods to combat cryptocurrency fraud and money laundering. Numerous researchers have sought to decrease a negative impact of money laundering by forecasting the risks involved with these transactions using statistical and classical ML techniques.

In, Cardoso et al., (2022) makes a novel suggestion to improve the detection of money laundering via transactions. We then use Random Forest (RF) to foretell instances of questionable activity related to money laundering. The RF model's predictive performance was 86% accurate when oversampling was used, and 92% accurate when undersampling was used [12].

In, Ketenci et al., (2021) using two-dimensional depictions of monetary transactions to provide a new set of features derived from time-frequency analysis. The effectiveness of the findings in realistic settings is shown by testing the RF algorithm on actual financial data. The F-score was 59.05% and the FPR was 14.9% when time-frequency characteristics were used independently. Incorporating transaction and CRM variables results in an enhanced F-Score of 74.06% and a FPR of 11.85%[13].

In particular, Labanca et al., (2022) comparison between the techniques implemented in this work and the latest unsupervised and supervised methods that are frequently employed in the AML domain. Our findings indicate that RF and Amaretto Isolation are the most effective methods for the given job. On average, the former has a detection rate of 0.793, which is 30% better than the latter, and its AUROC is 0.9, which is 20% better[14].

In, Chen et al., (2021) an innovative technique for determining the anomaly score threshold, called RFT, was used to create two autoencoder model variants: the single-loss and multi-loss VAE, as well as a GAN. This approach improved the performance of a model. Test outcomes shows that a FPR may decrease to 7% when using the suggested multi-loss AE model. had a successful accuracy rate of 93%, in comparison[15].

To ease the task, this study, Tai and Kan, (2019) develops a two-stage intelligent approach that utilises ML and data analysis approaches. The first stage of the built-in intelligent technique obtains a recall rate of 26.3%, according to assessments done using data supplied by Bank Sino Pac. This is three times more than the 8.6% recall rate of the Money Laundering Control Act in Taiwan. The second stage allows for an improvement of 87.04% in the precision rate[16].

In, Drezewski et al., (2012) uses analysis of bank statement cash flows to train an ANN to identify money laundering trends. The results establish the viability and reasonableness of the NN-based approach to cash flow analysis by demonstrating the ANN's high efficiency in detecting patterns of money laundering[17].

Despite advancements in machine learning for detecting money laundering, key research gaps remain. Current methods, like Random Forests and neural networks, struggle with high false positive rates and varying recall rates across different datasets. The integration of transaction and CRM features and novel thresholding methods show promise but need further validation and scalability. There's a need for more adaptive, interpretable, and universally applicable models to effectively combat evolving money laundering tactics.

Table -1: Related work for money laundering prediction

Refe.	Methodology	Achievements/results	Limitation/future work
[18]	Random Forest (RF) with oversampling and undersampling	• RF with oversampling: 86% accuracy • RF with undersampling: 92% accuracy	• Need to test on larger and more diverse datasets • Explore other resampling techniques for further improvement
[13]	Time-frequency analysis and Random Forest	• Time-frequency features: 14.9% FPR, 59.05% F-Score • Combined features: 11.85% FPR, 74.06% F-Score	• Integrate additional contextual features to enhance detection • Real-time implementation and scalability need assessment
[14]	Isolation Forest and Random Forest (Amaretto)	• Isolation Forest: AUROC 0.9 (20% better on average) • Random Forest: Detection rate 0.793 (30% better on average)	• Compare with more diverse state-of-the-art techniques • Explore hybrid models combining isolation and random forests
[15]	Autoencoder (single-loss and multi-loss VAE) and GAN with RFT	• Multi-loss AE: FPR drops to 7%, 93% accuracy	• Extend evaluation to different financial datasets • Optimize computational efficiency for real-time use
[16]	Intelligent two-stage process based on data analysis and ML	• Phase 1: Recall rate 26.3% (3x higher than current law) • Phase 2: Precision rate up to 87.04%	• Improve first phase precision without sacrificing recall • Enhance model adaptability to evolving money laundering tactics
[17]	Artificial Neural Network (ANN) analyzing cash flows	• High efficiency in identifying money laundering patterns	• Validate on larger and more varied datasets • Investigate interpretability of ANN decisions and outputs

3. METHODOLOGY FOR ANTI-MONEY LAUNDERING

We discuss the experiments that were conducted using the Elliptic dataset in this section. We begin by investigating the categorization of these datasets after appropriate feature reductions using a variety of ML methods, such as GCNs and MLPs for AML prediction.

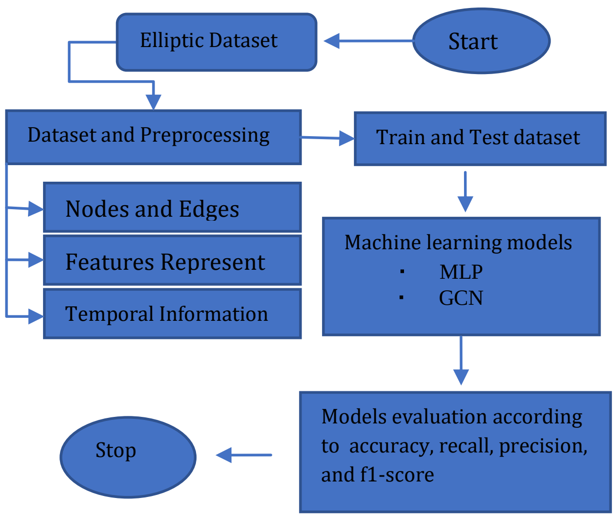

Fig -1: Flowchart of methodology

An above figure 1 shows each step of methodology for anti-money laundering prediction. The detailed analysis of each phase discussed below:

3.1 Dataset and Preprocessing

The Elliptic Dataset[1] is used for the purpose of evaluating the algorithms' effectiveness in identifying transactions connected to money laundering. This is due to the fact that it is the most comprehensively sampled publicly available dataset that contains information regarding the legitimacy of transactions, and it has not yet been extensively investigated in the ML-literature. Additionally, Elliptic, a company that specializes in the detection of financial crime in cryptocurrencies, has released the dataset, which enhances the data's reliability by collaborating with financial institutions and law enforcement. Close inspection reveals that the information comprises 49 graphs culled by a Bitcoin blockchain at different points in time. On display here are directed acyclic graphs covering two weeks' worth of transaction data. Additionally, every graph incorporates following linked blockchain transactions, beginning with a single transaction. As seen graphically in Figure 2:

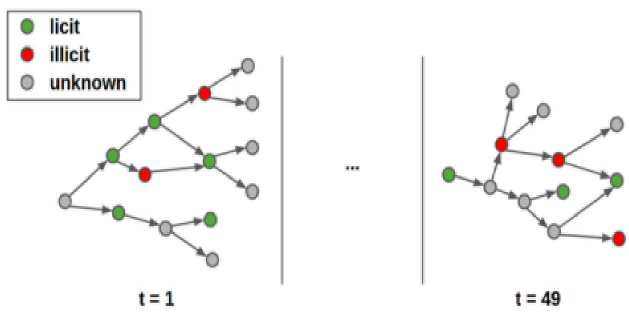

Fig -2: Structure of the dataset

[1] Available at https://www.kaggle.com/ellipticco/ellipticdataset

Each node in the network, as shown in Figure 2, displayed a Bitcoin transaction, and edges show a movement of Bitcoins. Figure 2 illustrates the division of the dataset into three categories: licit, illegal, and unknown. However, the last group is not included while developing or evaluating the ML-model. There are two main causes behind this. First, there must be a ground truth for every data point as all the algorithms used in the study relied on supervised learning. Some transactions do not have an appropriate label, rendering unsupervised learning useless. Secondly, the creator does not have access to the computing power or other hardware resources that would be required to make it functional. Figure 2 therefore displays the modified dataset, which has been stripped of the unfamiliar transactions.

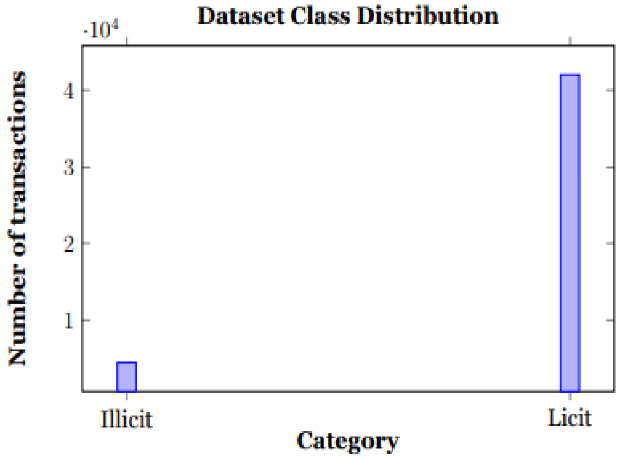

Fig -3: Distribution of the nodes in the adapted dataset

There were 46,564 transactions in the modified Elliptic Bitcoin Dataset, with 4545 being illicit and 4,2019 being licit, as shown in Figure 3.

3.1.1 Nodes and Edges

There are 203,769 transactions at nodes and 234,355 edges in a graph network that show how payments are made between nodes based on elliptic data. The data set is classified as having 2% illegal transactions (4,545 nodes) and 21% legal transactions (42,019 nodes), as shown in Table 2. The characteristics of uncertain labels accompany the surviving nodes. The complete Bitcoin transaction graph on the Blockchain is considered to include a graph network containing elliptic data as a sub-graph.

Table -2: Elliptic data set description

Transactions	Licit	Illicit	Unknown
Train set	26432	3462	106371
Test set	15587	1083	50834
Total	42019	4545	157205

3.1.2 Features

There are 166 attributes connected with each node in the utilised transaction graph. The initial 94 features refer to the local characteristics of a Bitcoin data, like fees, volume, and timestamps. The remaining features reflect aggregated numbers, such as the average amount of Bitcoin received or spent by the inputs and outputs.

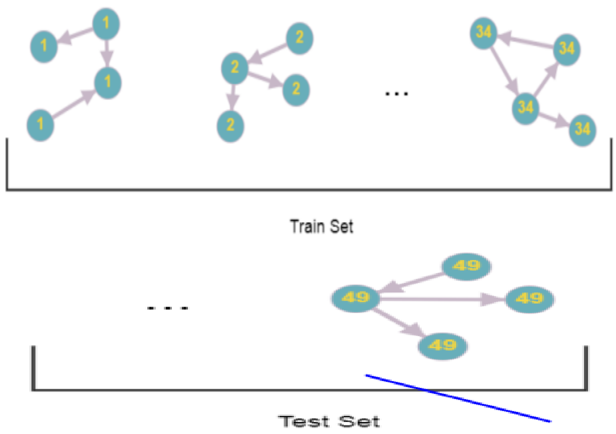

Fig -4: Elliptic data set with the train and test sets

Elliptic data set with the train and test sets separated, as shown abstractly in Figure 4. The labelled nodes stand in for the time steps. There is a distinct time-step for every linked graph.

3.1.3 Temporal Information

Each of the 49 time-steps in the elliptic data set is associated with a different transactional network. The Bitcoin blockchain records transactions every three hours, and these transactions are represented by time-steps, which form a connected graph network. A two-week period is often used to separate these time-steps. Furthermore, there's not a single edge connecting the graphs of any two different time-step pairs. The train and test sets are divided in this experiment according to a certain time frame. The data from the remaining time-steps is included in the test set, which is represented by the labels of the nodes in Figure 4, while the train set, which includes the data from the first 34 time-steps, is shown. In addition, the five most recent time steps by a training set make up a validation set.

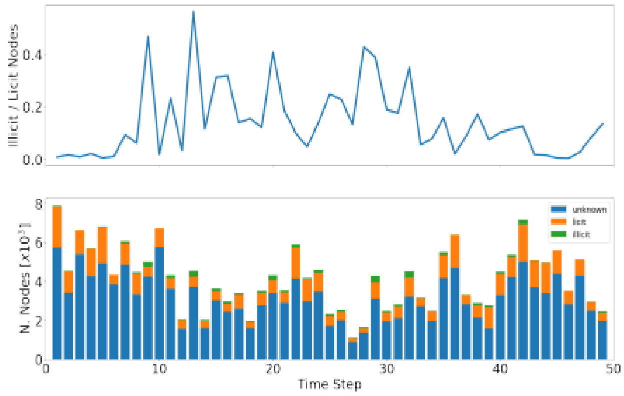

Fig -5: (Top) A percentage of nodes at various time steps in the dataset that are illegal compared to those that are licit. (Bottom) Relationship between node count and time step.

Every time step has a relatively constant number of nodes, somewhere from 1,000 to 8,000. See Figure 5.

3.2 Benchmark Machine learning Methods

In this paper have used two machine learning algorithms (MLP and GCN) on Elliptic data for anti-money laundering prediction.

3.2.1 Multi-Layer Perceptron (MLP)

Figure 6(a) shows the most famous neural network architecture, which is the multi-layered perceptron (MLP). Nodes in this design are grouped into layers and linked to one another across adjacent levels. In particular, a k-th layer's calculations may be expressed as [19]:

$$x^{(k)} = \sigma\left(W^{(k)}x^{(k-1)} + b^{(k)}\right) \dots \dots (1)$$

The input of layer k is represented by x(k) ∈ Rp(k), which is equivalent to the p(k) nodes of the NN. Full connection among a (k − 1)-th and k-th layers is made possible by the weights W(k) ∈ Rp(k) × p(k−1) and a bias vector b(k) ∈ Rp(k). The output becomes non-linear due to the activation function σ, which allows for complex representations. Specifically, for a sample in D, the input feature is x(0) [8].

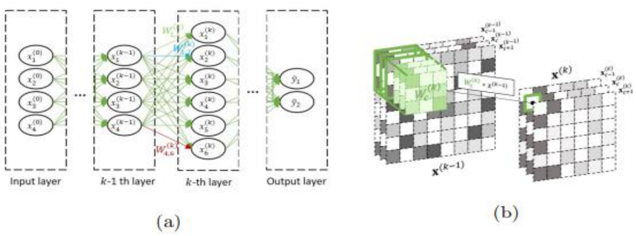

Fig -6: (a) A three-layer model of MLPs, including an input layer, hidden layers, and an output layer. (b) Input x (k-1) is sent via the convolution filter W(k) c, which then produces the c-th channel of x (k).

A fully connected (FC) layer is another name for the layer in Eq. (1). Neural networks may build more intricate representations of the input by stacking more FC layers. There is a forward propagation stage that runs calculations by a first (input) layer to a final (output) layer in order to get predictions. A neural network's parameters Θ = {W(k), b (k)} may be optimized by updating them in the inverse order during a backward propagation step.

3.2.2 Graph Convolutional Networks (GCN)

A growing number of researchers are interested in deep learning's application to graph-structured data [20][21]. There has been a lot of progress in solving the scalability problems caused by the combinatorial complexity of graph structures, which is an issue for real-world applications [22][23][24]. More especially, GCNs are taken into account. Similar to a perceptron, but with the addition of a neighborhood aggregation step inspired by spectral convolution, a GCN is composed of several layers of graph convolution.

N represents a collection of transactions at the nodes, while E represents a set of edges that indicate a flow of Bitcoin. Therefore, the Bitcoin transaction graph by an Elliptic Data collection can be represented as G = (N, E). Inputs to the l-th GCN layer include an adjacency matrix A and a node embedding matrix H(l). Outputs, which are weight matrices W(l), update the node embedding matrix to H (l+1). When we do the math's, we write[25] (2):

$$H^{(l+1)} = \sigma\left(\hat{A}H^{(l)}W^{(l)}\right) \dots \dots \dots \dots (2)$$

where $\hat{A}$ is defined as (3) and represents a normalization of A:

$$\hat{A} = \widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}, \widetilde{A} = A + I, \widetilde{D} = \mathrm{diag}(\sum_j \widetilde{A}_{ij}) \dots \dots \dots \dots (3)$$

The activation functions, usually ReLU, are implemented in all levels except an output layer using σ. A first embedding matrix, H(0) = X, is given by the characteristics of the nodes. Assume that there are L convolutional layers in the graph. The softmax, with H (L) representing prediction probabilities, is the output layer for node categorization.

Adam is an optimizer that we use in our experiments; it is a method for minimizing a loss function that combines momentum with adaptive learning rates. Weight decay is a regularization approach that we use as a parameter optimization tool. It helps limit the model complexity and reduces a danger of overfitting. Finally, we activate a model using a ReLU. The incorporation of nonlinearity and the enhancement of the capacity to capture complex correlations within data are common goals of this method when using DL

4. RESULTS ANALYSIS AND DISCUSSION

This section shows the result of applying both machine learning algorithms to the study on anti-money laundering. Each algorithm experimental results are discussed below.

4.1 Evaluating Model Performance

The variety of performance measures available makes classification model evaluation a challenging task. The current investigation depends on a binary classification model that uses the numbers 1 (money laundering) and 0 (no money laundering) to forecast whether a transaction will include money laundering or not. One of numerous possible outcomes may be used to classify a given observation. Equation (4-7) sets forth the computational formulae for various assessment metrics:

Accuracy: One simple metric for accuracy is a proportion of accurate classifications (TP + TN) to a total number of observations.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \dots (4)$$

Precision: Accurate positive predictions inside the positive class are measured as precision. Precision is best understood as follows, according to mathematical equation 3:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{TN}} \dots (5)$$

Recall: The recall metric quantifies the quantity of real positive data points that the model accurately predicted. Recall is shown by the following mathematical equation 4:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FP}} \dots (6)$$

F-measure: The recall and precision values are used to generate the F1-score. The idea of the F1-score is demonstrated as follows in mathematical equation 5:

$$\text{F} - \text{measure} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \dots (7)$$

A total of seven formulas are provided, where n displayed a number of categories, TP represents a number of positively predicted samples within a category, TN represents a number of negatively predicted samples within a category, FP represents a number of incorrectly forecast positive samples within a category, and FN displayed a number of positively predicted samples within a category.

A fundamental assessment metrics used in this research are f1-score, recall, accuracy, and precision. The reason for this decision is the very unequal distribution of the Elliptic dataset's samples.

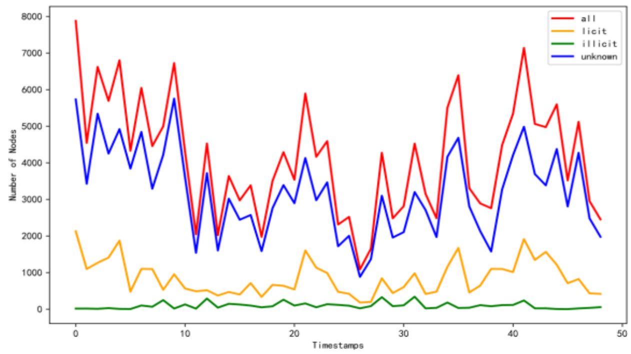

Fig -7: Distribution of money laundering transaction nodes across dynamic timestamps.

Figure 7 shows that there are many less nodes linked to illegal money laundering in an Elliptic dataset compared to the number of nodes linked to real nodes. In this case, models have a tendency to do better on categories with more samples and worse on categories with less samples. There needs to be a heavy focus on the relatively small criminal transaction samples since our study is primarily concerned with forecasting illicit money laundering activities.

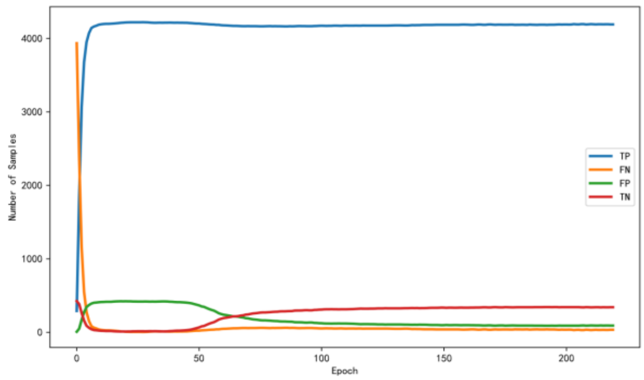

Fig -8: Trends of different elements in confusion matrix over

Figure 8's confusion matrix shows that the model is now producing almost perfect predictions for real-world transaction data. On the other hand, the TN values are almost nil, suggesting that it has a very limited capacity to identify illicit money laundering activities.

Table -3: Classification results of model for anti-money laundering detection

Parameters	Accuracy	Precision	Recall	F1 Score
MLP	98	97.3	99.5	98.4
GCN	97.3	97.5	99.4	98.4
C4.5 [26]	92	92	92	92
KNN[27]	92	97	97	97

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 03 | AUG 2023 www.irjet.net p-ISSN: 2395-0072

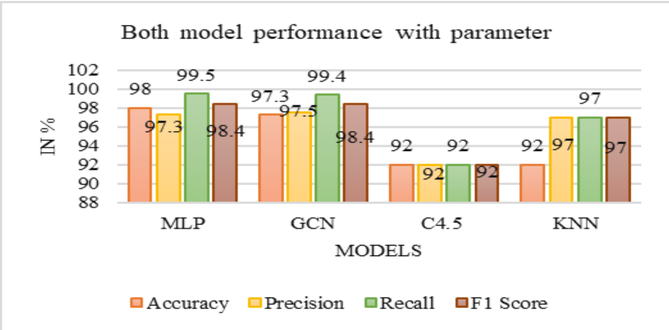

Fig -9: Bar graph of Both models' performance for anti-money laundering detection

The research on AML applied two ML algorithms, MLP and GCN, to evaluate their effectiveness. The results, as summarized in Table 3, demonstrate that both models exhibit high performance across various metrics. In particular, the MLP attained a 98.4% F1 score, a 99.5% recall, a 97.3% precision, and a 98% accuracy rate. Along with a high accuracy of 97.3%, the suggested GCN model also has a precision of 97.5%, recall of 99.4%, and an F1 score of 98.4%. These outcomes reveal that both models are fairly precise in identifying money laundering activities with slight differences in accuracy, and recall of MLP compared to GCN, and precision of GCN compared to MLP. The achieved F1 scores in both models are high which proves the equal balance of their precision and recall. The above bar graph (Figure 9) also depicts these metrics, and even though the performance of both models differs slightly, they are quite similar, which strengthens their performance in this application.

Table -4: Comparison between base and proposed models for anti-money laundering detection

Parameters	Accuracy	Precision	Recall	F1 Score
MLP	98	97.3	99.5	98.4
GCN	97.3	97.5	99.4	98.4
C4.5 [26]	92	92	92	92
KNN[27]	92	97	97	97

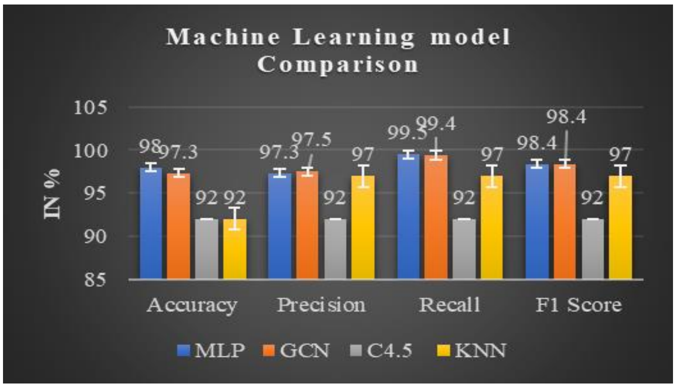

Fig -10: Bar graph of ML models performance comparison for anti-money laundering detection

Table 4 shows that the suggested MLP and GCN models outperform the current modern ML models for AML. The MLP model achieved the highest accuracy (98%), recall (99.5%), and a high F1 score (98.4%), indicating its excellent overall performance. In addition, the GCN model did quite well; its accuracy was 97.3%, precision was 97.5%, recall was 99.4%, and F1 score was 98.4%. A 77% Figure 10 visually depicts these performance metrics, clearly demonstrating the MLP and GCN models' superior capability in accurately and reliably detecting money laundering activities compared to the existing models.

5. CONCLUSION AND FUTURE WORK

We provide MLP, a model for predicting money laundering based on DL, to manage the spatiotemporal complexity of data related to money laundering transactions and so mitigate the risks associated with these types of financial transactions. In this research, we used the Elliptic Bitcoin Dataset to effectively identify instances of money laundering by using ML methods, particularly GCN and MLP. Both models demonstrated high performance, with MLP achieving 98% accuracy and GCN achieving 97.3%, alongside balanced precision, recall, and F1 scores. These models were able to detect illegal transactions in a dataset that was otherwise completely valid, demonstrating their resilience and practical use. The lack of such data has hindered our capacity to verify and improve our approaches in a wider range of circumstances, despite the approach's partial effectiveness in handling money laundering transaction data. However, a possible application case for our suggested approach is any transaction scenario that has distinct transaction characteristics and transparent transaction flows. Hence, in order to test and improve our techniques and make them more applicable to a broader variety of real-world uses, future research endeavours need to go deeper into transactions and larger datasets. Future work will focus on advanced feature engineering, exploring more sophisticated models, expanding dataset, developing real-time detection capabilities, operational testing in financial environments, and addressing integration of transactions with unknown labels to further enhance the detection of financial crimes.

REFERENCES

[1] F. M. J. Teichmann, "Twelve methods of money laundering," *J. Money Laund. Control*, 2017, doi: 10.1108/JMLC-05-2016-0018.

[2] A. Bahulkar, N. O. Baycik, T. Sharkey, Y. Shen, B. Szymanski, and W. Wallace, "Integrative Analytics for Detecting and Disrupting Transnational Interdependent Criminal Smuggling, Money, and Money-Laundering Networks," 2018, doi: 10.1109/THS.2018.8574121.

[3] Heinrich-Böll-Stiftung, R. Schönenberg, and A. von Schönfeld, "Transnational Organized Crime:Analyses

of a Global Challenge to Democracy," *CrimRxiv*, 2014, doi: 10.21428/cb6ab371.67b10530.

[4] B. Oztas, D. Cetinkaya, F. Adedoyin, and M. Budka, "Enhancing Transaction Monitoring Controls to Detect Money Laundering Using Machine Learning," 2022, doi: 10.1109/ICEBE55470.2022.00014.

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

[6] Y. Zhang and P. Trubey, "Machine Learning and Sampling Scheme: An Empirical Study of Money Laundering Detection," *Comput. Econ.*, 2019, doi: 10.1007/s10614-018-9864-z.

[7] A. I. Canhoto, "Leveraging machine learning in the global fight against money laundering and terrorism financing: An affordances perspective," *J. Bus. Res.*, 2021, doi: 10.1016/j.jbusres.2020.10.012.

[8] V. Rohilla, S. Chakraborty, and R. Kumar, "Deep learning based feature extraction and a bidirectional hybrid optimized model for location based advertising," *Multimed. Tools Appl.*, 2022, doi: 10.1007/s11042-022-12457-3.

[9] V. Rohilla, S. Chakraborty, and M. Kaur, "An Empirical Framework for Recommendation-based Location Services Using Deep Learning," *Eng. Technol. Appl. Sci. Res.*, 2022, doi: 10.48084/etasr.5126.

[10] S. Mann, A. Balyan, V. Rohilla, D. Gupta, Z. Gupta, and A. W. Rahmani, "Artificial Intelligence-based Blockchain Technology for Skin Cancer Investigation Complemented with Dietary Assessment and Recommendation using Correlation Analysis in Elder Individuals," *Journal of Food Quality*. 2022, doi: 10.1155/2022/3958596.

[11] A. Rath, A. Das Gupta, V. Rohilla, A. Balyan, and S. Mann, "Intelligent Smart Waste Management Using Regression Analysis: An Empirical Study," 2022, doi: 10.1007/978-3-031-07012-9_12.

[12] M. Cardoso, P. Saleiro, and P. Bizarro, "LaundroGraph: Self-Supervised Graph Representation Learning for Anti-Money Laundering," 2022, doi: 10.1145/3533271.3561727.

[13] U. G. Ketenci, T. Kurt, S. Onal, C. Erbil, S. Akturkoglu, and H. S. Ilhan, "A Time-Frequency Based Suspicious Activity Detection for Anti-Money Laundering," *IEEE Access*, 2021, doi: 10.1109/ACCESS.2021.3072114.

[14] D. Labanca, L. Primerano, M. Markland-Montgomery, M. Polino, M. Carminati, and S. Zanero, "Amaretto: An Active Learning Framework for Money Laundering Detection," *IEEE Access*, 2022, doi: 10.1109/ACCESS.2022.3167699.

[15] Z. Chen, W. M. Soliman, A. Nazir, and M. Shorfuzzaman, "Variational Autoencoders and Wasserstein Generative Adversarial Networks for Improving the Anti-Money Laundering Process," *IEEE Access*, 2021, doi: 10.1109/ACCESS.2021.3086359.

[16] C. H. Tai and T. J. Kan, "Identifying Money Laundering Accounts," 2019, doi: 10.1109/ICSSE.2019.8823264.

[17] R. Drezewski, J. Sepielak, and W. Filipkowski, "System supporting money laundering detection," *Digit. Investig.*, 2012, doi: 10.1016/j.diin.2012.04.003.

[18] H. S. Assumpcao, F. Souza, L. L. Campos, V. T. De Castro Pires, P. M. L. De Almeida, and F. Murai, "DELATOR: Money Laundering Detection via Multi-Task Learning on Large Transaction Graphs," 2022, doi: 10.1109/BigData55660.2022.10021010.

[19] A. Chatterjee, J. Saha, and J. Mukherjee, "Clustering with multi-layered perceptron," *Pattern Recognit. Lett.*, 2022, doi: 10.1016/j.patrec.2022.02.009.

[20] Y. Li, R. Zemel, M. Brockschmidt, and D. Tarlow, "Gated graph sequence neural networks," 2016.

[21] L. McInnes, J. Healy, N. Saul, and L. Großberger, "UMAP: Uniform Manifold Approximation and Projection," *J. Open Source Softw.*, 2018, doi: 10.21105/joss.00861.

[22] N. Shibu, sini sunny, A. Rajkumar, A. Sayed, P. Kalaiselvi, and R. Namakkal-Soorappan, "N-Acetyl Cysteine Administration Impairs EKG Signals in the Humanized Reductive Stress Mouse," *Free Radic. Biol. Med.*, 2022, doi: 10.1016/j.freeradbiomed.2022.10.122.

[23] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2017.

[24] R. Ying, R. He, K. Chen, P. Eksombatchai, W. L. Hamilton, and J. Leskovec, "Graph convolutional neural networks for web-scale recommender systems," 2018, doi: 10.1145/3219819.3219890.

[25] M. Weber *et al.*, "Anti-Money Laundering in Bitcoin: Experimenting with Graph Convolutional Networks for Financial Forensics," no. 10, 2019.

[26] S. Kaur and N. K. Sandhu, "Novel Technique of Anti-Money Laundering Using Deep Learning," vol. 9, no. 5, 2022.

[27] J. Alotibi, B. Almutanni, T. Alsubait, H. Alhakami, and A. Baz, "Money Laundering Detection using Machine Learning and Deep Learning," *Int. J. Adv. Comput. Sci. Appl.*, 2022, doi: 10.14569/IJACSA.2022.0131087.