Proceedings of the 3rd International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/138/2025.21382

Enhancing Anti-Money Laundering Detection with Self-Attention Graph Neural Networks

Tingyu Yuan[1,a,*], Sizhe Wang[2,b], Yixin Tao[3,c]

[1]School of Economics and Management, Shanghai Ocean University, Shanghai, China
[2]Business School, Lehigh University, Dallas, TX, USA
[3]Accounting/Finance Department, Heritage Provider Network, Northridge, CA, USA
a. 1401945702@qq.com, b. siw422@lehigh.edu, c. yixintao@usc.edu
**corresponding author*

Abstract: Money laundering remains a significant global issue, undermining financial stability and security. This study introduces a Self-Attention-GNN Model enhanced with a self-attention mechanism to improve the detection of money laundering activities in a large, imbalanced dataset of financial transactions. The dataset, covering 97 days and including approximately 180 million transactions, contains 223,000 labeled laundering cases. By representing financial transactions as a graph—where entities such as accounts and banks are nodes, and transactions are edges—the model captures intricate relational and structural dependencies within the transaction network. The addition of the self-attention mechanism enables the model to dynamically adjust feature aggregation, focusing on the most relevant nodes and edges, which significantly improves the model's ability to identify laundering activities. Despite the challenges posed by class imbalance, the model achieves robust performance in detecting illicit transactions while reducing false positives. The paper also discusses potential strategies for further optimizing precision and recall, such as advanced graph architectures, oversampling methods, and enhanced node embedding techniques. Overall, this research highlights the power of graph-based deep learning approaches for anti-money laundering (AML) applications, demonstrating how structural and relational dependencies within financial networks can be leveraged to enhance detection accuracy.

Keywords: Graph Neural Networks, Anti-money laundering, Financial transactions, Self-attention

1. Introduction

Money laundering remains a major global challenge, threatening economic stability and the integrity of financial systems. With the rise of digital transactions, detecting illicit activities has become more complex. Traditional anti-money laundering (AML) methods often generate high false positive rates, misclassifying legitimate transactions, while false negatives allow illicit activities to go undetected. This study leverages Graph Neural Networks (GNNs) with self-attention mechanisms to enhance money laundering detection, addressing these critical issues.

The dataset used in this study represents a diverse financial ecosystem, covering transactions among individuals, companies, and banks across various activities, such as salary payments, loan repayments, and purchases. A small subset of these transactions involves money laundering, making

DOI: 10.54254/2755-2721/138/2025.21382

detection particularly challenging. Graph-based modeling allows for a deeper understanding of transactional relationships, revealing suspicious activities that may not be evident when analyzing individual transactions in isolation.

GNNs are particularly effective for relational and structural data analysis, making them ideal for financial transaction networks. In this framework, accounts, banks, and individuals are represented as nodes, while transactions serve as edges. Suspicious transactions often emerge when viewed within a broader network—such as frequent transactions among a small group of nodes, potentially indicating smurfing or circular layering. Unlike conventional machine learning models, GNNs excel at detecting hidden patterns in these interactions, providing a more powerful tool for AML analysis.

This study applies self-attention-enhanced GNNs to predict money laundering activities using a labeled dataset. The dataset includes variables such as transaction networks, amounts, and payment methods, alongside binary labels distinguishing legitimate and laundering transactions. The primary objectives are to analyze relational patterns, enhance detection accuracy, and reduce false alarms. By integrating self-attention mechanisms, the model dynamically adjusts feature aggregation based on the importance of neighboring nodes, improving its ability to detect money laundering. The proposed approach improves accuracy while reducing false positives and negatives, enhancing the practical usability of AML models in financial institutions. This study contributes to the growing body of machine learning research in AML, demonstrating the potential of graph-based deep learning for financial crime detection.

2. Literature Review

The increasing sophistication of money laundering schemes has driven the application of advanced machine learning and deep learning techniques in anti-money laundering (AML) systems. Recent studies have explored diverse approaches, including graph-based models, recurrent architectures, and hybrid frameworks, to enhance the detection of laundering activities. This review synthesizes insights from the literature and situates this study within the broader research context.

Graph-based models have gained prominence in AML research for their ability to capture the relational and temporal dynamics of financial transactions. Wan, Fei, and Li introduced a hybrid model combining Dynamic Graph Convolutional Neural Networks (DGCNN) with Long Short-Term Memory (LSTM) networks, demonstrating enhanced prediction accuracy for laundering activities by leveraging both topological and sequential data [1]. Similarly, Wei et al. proposed a dynamic graph convolutional network to analyze evolving transaction networks, which proved effective in identifying complex laundering patterns [2]. Alarab and Prakoonwit extended this line of research by integrating Temporal Graph Convolutional Networks (TGCNs) with LSTM, achieving high performance on Bitcoin transaction datasets, a domain marked by pseudonymity and rapid transaction dynamics [3]. These studies underscore the potential of graph-based frameworks in modeling the intricate interdependencies among entities and transactions.

Recurrent Neural Networks (RNNs) and their variants have also been widely applied in AML. Girish and Bhowmik (2024) explored a hybrid ensemble model combining RNNs with other classifiers, reporting improvements in recall for minority classes, such as laundering activities. Their findings align with this study, which utilized RNNs to capture temporal dependencies in transaction sequences, achieving strong recall but struggling with precision due to the dataset's imbalance. Additionally, Yang, Liu, and Li emphasized the role of intelligent algorithms, including RNNs, in enhancing supervisory capabilities for AML [4]. These studies highlight the strengths of recurrent architectures in processing sequential data but also point to the challenges posed by imbalanced datasets and high false positive rates.

Beyond model-specific approaches, broader reviews have provided critical insights into the state of AML research. Kute et al. and Han et al. examined the application of deep learning and explainable

DOI: 10.54254/2755-2721/138/2025.21382

artificial intelligence (XAI) techniques in AML [5-6]. They emphasized the trade-offs between model accuracy and interpretability, a challenge echoed in Jensen and Iosifidis's work, which proposed methods to qualify and prioritize AML alarms using deep learning [7]. These studies stress the importance of balancing technical performance with operational feasibility, particularly in regulatory environments that demand transparency and accountability [8].

3. Data introduction

3.1. Data Description

Money laundering detection faces challenges with high false-positive and false-negative rates. This study utilizes the HI-Large dataset, covering 97 days and 180 million transactions, including 223,000 laundering-related cases. The dataset captures the full laundering cycle: Placement, Layering, and Integration, allowing comprehensive analysis of financial interactions between individuals, companies, and banks. By examining transaction patterns across institutions, this study provides insights into illicit financial activities, enhancing the accuracy of money laundering detection and contributing to improved financial security measures.

3.2. Descriptive statistical analysis

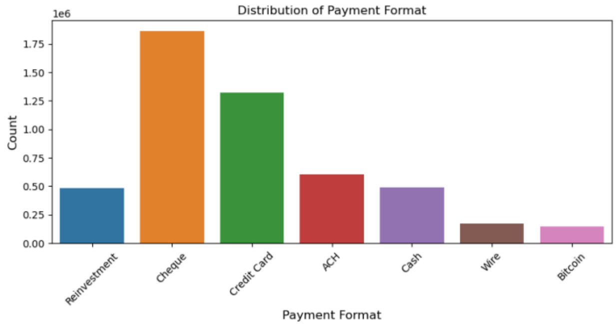

Figure 1: Distribution of Payment Format

The bar chart illustrates the distribution of transaction payment formats within the dataset. The most prevalent payment method is "Cheque," followed by "Credit Card," which also shows a significant count. Other payment formats, such as "ACH," "Cash," and "Reinvestment," appear less frequently but still have notable representations. In contrast, "Wire" and "Bitcoin" transactions are relatively rare, indicating that traditional financial methods dominate the dataset, while digital currencies play a smaller role in the recorded transactions.

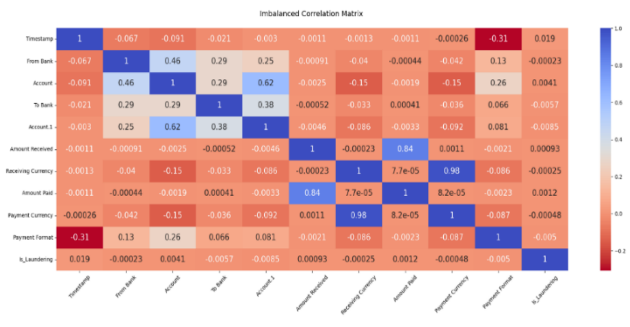

Figure 2: Imbalanced Correlation Matrix

DOI: 10.54254/2755-2721/138/2025.21382

The heatmap depicts the correlation coefficients among the numerical and categorical variables in the dataset. Most correlations are relatively low, indicating weak linear relationships between the majority of variables. However, a strong positive correlation exists between "Amount Received" and "Amount Paid," reflecting expected consistency in financial transactions. Moderate correlations are observed between account identifiers, such as "Account" and "Account.1." The "Payment Format" variable has a notable negative correlation with "Timestamp" and a slightly weaker negative relationship with "Is Laundering," suggesting some temporal and laundering-related patterns. Overall, the matrix highlights minimal multicollinearity, ensuring the dataset is suitable for predictive modeling.

4. Self-attention-GNN Model Introduction

4.1. Model structure

Adding self-attention mechanism to the existing GNN structure can effectively improve the innovation and performance of the model. Self-attention mechanism can help the model better capture the complex relationship and importance between nodes, thus improving the ability to identify money laundering activities.

1. The volume layer:

$$[H^{(l+1)} = \sigma\left(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}H^{(l)}W^{(l)}\right)$$

First of all, the model uses graph volume to extract the initial features of nodes. The graph volume accumulation layer generates a new node representation by aggregating the information of each node and its neighbors. This process can capture the global structure information between nodes.

2. Self-attention layer:

$$[e_{ij} = \text{LeakyReLU}\left(a^T\left[Wh_i \,\square\, Wh_j\right]\right)$$

After the scroll is stacked, the self-attention mechanism is introduced into the model. The self-attention layer dynamically adjusts the way of feature aggregation by calculating the attention weight between each node and its neighbors. In this way, the model can identify which neighbor nodes are most important for the feature update of the current node. The introduction of self-attention mechanism makes the model pay more attention to the key nodes and edges in the trading network, thus improving the ability to identify money laundering activities.

In order to optimize the model, we use binary cross entropy loss function. The loss function can effectively measure the difference between the predicted results of the model and the real labels, and guide the training process of the model.

Finally, the model uses the updated node embedding to predict the edge level. By combining the characteristics of two connecting nodes, the model can judge whether the transaction is money laundering.

5. Self-attention-GNN model analyse

The GNN model's results highlight a significant disparity in performance between the two classes of legitimate transactions and laundering activities. For legitimate transactions (class 0), the model achieves an exceptional precision of 1.00, meaning all predicted legitimate transactions are correct, and a recall of 0.94, successfully identifying the majority of such cases. This leads to a high F1-score of 0.97, demonstrating a balanced performance for this majority class. However, for laundering activities (class 1), while the recall is impressively high at 0.93—ensuring most illicit transactions are detected—the precision is only 0.02, indicating that the vast majority of flagged laundering transactions are false positives. The overall F1-score for this minority class is merely 0.04, emphasizing the challenges posed by the highly imbalanced dataset.

Table 2: Model result

	precision	recall	f1-score	support
0	1.00	0.94	0.97	9986
1	0.02	0.93	0.04	14
accuracy			0.94	10000
macro avg	0.51	0.94	0.51	10000
weighted avg	1.00	0.94	0.97	10000

The overall accuracy of 94% reflects strong classification performance for the dataset as a whole, though this metric is largely dominated by the majority class. Macro averages, which weigh both classes equally, reveal lower performance with an F1-score of 0.51, highlighting the imbalance. Weighted averages, skewed toward the majority class, inflate performance measures, masking the weaknesses in classifying laundering activities.

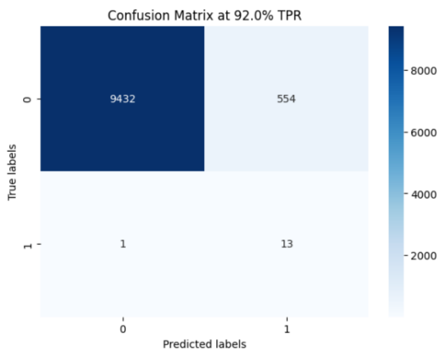

Figure 3: Confusion Matrix

The confusion matrix at a 92% true positive rate (TPR) offers more nuanced insights into the model’s performance. Of the 10,000 transactions evaluated, the model correctly classified 9432 legitimate transactions (true negatives) and 13 laundering transactions (true positives). However, it incorrectly flagged 554 legitimate transactions as laundering activities (false positives) and missed only 1 laundering transaction (false negative). These results highlight the model’s excellent recall for laundering activities.

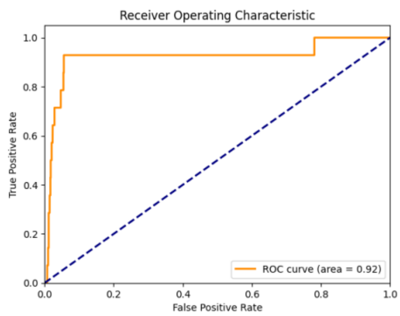

Figure 4: ROC curve

The ROC curve further underscores the model's potential, achieving an area under the curve (AUC) score of 0.92. This indicates a strong ability to differentiate between legitimate and laundering

DOI: 10.54254/2755-2721/138/2025.21382

transactions overall. However, while the ROC AUC score suggests robust discriminative power, its utility is limited by the imbalance in the dataset. A high AUC in this context may overestimate the model's practical effectiveness, especially considering the significant trade-offs between precision and recall for laundering transactions.

In summary, while the Self-attention-GNN model demonstrates strong recall for detecting laundering activities and overall accuracy, its precision for identifying such transactions is extremely low. This trade-off between recall and precision is evident in the high number of false positives, which could hinder practical deployment. The ROC AUC score indicates strong overall discriminatory power but is insufficient alone to gauge effectiveness in an imbalanced setting.

Addressing this imbalance through techniques such as oversampling, graph-level loss balancing, or graph augmentation methods could enhance the model's precision while maintaining high recall. These adjustments would make the model more practical for anti-money laundering applications, balancing detection accuracy with operational feasibility.

6. Conclusions and Suggestions

This study leverages a Self-attention-GNN model to address the complexities of money laundering detection in a highly imbalanced dataset of financial transactions. By representing transactions as a graph, the Self-attention-GNN model effectively captures structural and relational patterns among entities (e.g., individuals, accounts, and banks). The analysis demonstrates that the Self-attention-GNN is particularly adept at uncovering illicit activity within the transaction network, achieving high recall rates for laundering activities.

However, the results also reveal a notable trade-off: while recall is high, precision for identifying laundering transactions remains low, resulting in a high false positive rate. This highlights the inherent difficulty of applying machine learning models to imbalanced datasets, where the minority class—laundering activities—is heavily outnumbered by legitimate transactions.

Although the overall accuracy and strong ROC AUC score suggest robust model performance from a global perspective, these metrics fail to adequately capture the model's challenges in correctly classifying the minority class. These findings underscore the importance of evaluating machine learning models with metrics tailored to imbalanced datasets, ensuring practical applicability in domains like anti-money laundering.

Future research should focus on testing the model with real-world transactional data from financial institutions to evaluate its practical effectiveness. Incorporating dynamic features, such as account behavior trends over time, could also enrich the model's predictive power and provide deeper insights into laundering patterns. These refinements would help build a more balanced, accurate, and operationally viable system for anti-money laundering applications.

References

[1] *Wan, Fei, and Li. "A Novel Money Laundering Prediction Model Based on a Dynamic Graph Convolutional Neural Network and Long Short-Term Memory." Symmetry 16.3 (2024): 378.*

[2] *Wei, Tianpeng, et al. "A Dynamic Graph Convolutional Network for Anti-money Laundering." International Conference on Intelligent Computing. Singapore: Springer Nature Singapore, 2023.*

[3] *Alarab, Ismail, and Simant Prakoonwit. "Graph-based lstm for anti-money laundering: Experimenting temporal graph convolutional network with bitcoin data." Neural Processing Letters 55.1 (2023): 689-707.*

[4] *Girish, K. K., and Biswajit Bhowmik. "Money Laundering Detection in Banking Transactions using RNNs and Hybrid Ensemble." 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2024.*

[5] *Kute, Dattatray Vishnu, et al. "Deep learning and explainable artificial intelligence techniques applied for detecting money laundering–a critical review." IEEE access 9 (2021): 82300-82317.*

DOI: 10.54254/2755-2721/138/2025.21382

[6] Han, J., Huang, Y., et al. "Artificial intelligence for anti-money laundering: a review and extension." Digital Finance 2.3 (2020): 211-239.

[7] Jensen, Rasmus Ingemann Tuffveson, and Alexandros Iosifidis. "Qualifying and raising anti-money laundering alarms with deep learning." Expert Systems with Applications 214 (2023): 119037.

[8] Yang, Guangyi, aoxing Liu, and Beixin Li. "Anti-money laundering supervision by intelligent algorithm." Computers & Security 132 (2023): 103344.